Wavelets for Urban Spatial Feature Discrimination: Comparisons with Fractal, Spatial Autocorrelation, and Spatial Co-Occurrence Approaches

Soe Win Myint, Nina S.-N. Lam, and John M. Tyler

Abstract

Traditional image processing techniques have proven inadequate for urban mapping using high spatial resolution remote-sensing images. This study examined and evaluated wavelet transforms for urban texture analysis and image classification using high spatial resolution ATLAS imagery. For the purpose of comparison and to evaluate the effectiveness of the wavelet approaches, two different fractal approaches (isarithm and triangular prism), spatial autocorrelation (Moran's I *and Geary's* C*), and spatial co-occurrence matrix of the selected urban classes were examined using 65 × 65, 33 × 33, and 17 × 17 samples with a pixel size of 2.5 m. Results from this study suggest that a multi-band and multi-level wavelet approach can be used to drastically increase the classification accuracy. The fractal techniques did not provide satisfactory classification accuracy. Spatial autocorrelation and spatial co-occurrence techniques were found to be relatively effective when compared to the fractal approaches. It can be concluded that the wavelet transform approach is the most accurate of all four approaches.*

Introduction

The traditional classification approach has been criticized for not being able to classify complex urban features accurately when fine spatial resolution imagery is used (Cushnie, 1987; Gong and Howarth, 1990; Fung and Chan, 1994). The higher the spatial resolution of remotely sensed image data, the more detailed features and smaller objects (e.g., single-family vs. multi-family homes, roads, trees, shrubs, grass land, pools, footpaths, driveways, parking lots) can be detected. Consequently, the spectral signatures of urban features as a whole become more heterogeneous. Moreover, urban features are composed of spectrally diverse materials concentrated in a small area (e.g., plastic, metal, rubber, glass, cement, wood, shingle, sand, gravel, brick, stone, soil, vegetation, water). In fact, the high frequency spatial appearance or complex nature of urban areas may be the major limitation in urban mapping using high spatial resolution image data (Myint *et al.*, 2002).

Unfortunately, common image processing algorithms do not take the local structure or the spatial arrangement of neighborhood pixels into consideration in their classification. The traditional classification method, such as the maximum likelihood classifier, uses spectral information (pixel values) as a basis to analyze and classify remote sensing images. To extract the heterogeneous nature of urban features in high-resolution images, we need to consider the spatial arrangements of neighborhood pixels, instead of an individual pixel value. Traditional spectral based algorithms use the individual pixel value and ignore any spatial information which may be crucial in urban land use and land cover mapping. This is especially true when most of the urban land-cover classes generally contain a number of spectrally different pixels or objects. For example, roads, houses, grasses, trees, bare soil, shrubs, swimming pools, and footpaths, each of which may have a completely different spectral response are together considered as a residential class. Moreover, the training statistics of urban classes may exhibit a very high standard deviation and violate one of the basic assumptions of the widely used maximum-likelihood decision rule, namely, that the pixel values follow a multivariate normal distribution (Sadler *et al.*, 1991).

There have been efforts to increase the accuracy of the spectral analysis of remotely sensed data by using texture transforms in which some measure of variability in digital number (DN) values is estimated within local windows; e.g., contrast between neighboring pixels (Edwards *et al.*, 1988), the standard deviation (Arai, 1993), or local variance (Woodcock and Harward, 1992). A number of researchers used variograms to describe the spatial patterns manifested in remote sensing imagery (e.g., De Jong and Burrough, 1995; Woodcock *et al.*, 1988). Emerson *et al.* (1999) analyzed the fractal dimension using the isarithm method and the spatial autocorrelation of satellite sensor imagery using Moran's *I* and Geary's *C* to observe the differing spatial structures of the smooth and rough surfaces in remotely sensed images. Lam and Quattrochi (1992) demonstrated that the fractal dimension of remote sensing data could yield quantitative insight on the spatial complexity and information content contained within these data. Quattrochi *et al.* (1997) used a software package known as the Image Characterization and Modeling Systems (ICAMS) to explore how fractal dimension is related to surface texture using the isarithm method (Lam and De Cola, 1993), the variogram (Mark and Aronson, 1984), and the triangular prism methods (Clarke, 1986).

These methods alone may not provide satisfactory accuracy when applied to fine-resolution remotely sensed images for urban land-use and land-cover mapping. Recent development of the wavelet transform could be used to extract different

S.W. Myint is with the Department of Geography, University of Oklahoma, 100 East Boyd Street, Norman, Ok 73019 (swmyint@ou.edu).

N. Lam is with the Department of Geography, 227 Howe/Russell Geoscience Complex, and J.M. Tyler is with the Department of Computer Science, 290 Coates Hall; both at Louisiana State University, Baton Rouge, LA 70803 (nlam@lsu.edu; tyler@lsu.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 70, No. 7, July 2004, pp. 803–812.

0099-1112/04/7007–0803/$3.00/0

texture features at different scales, thus providing a useful alternative for urban land cover classification.

Multiresolution Wavelet Decomposition

Mallat (1989) developed the multiresolution analysis theory using an orthonormal wavelet basis. The multiresolution wavelet transform decomposes a signal into a low frequency approximation and its high frequency detail information at a coarser spatial resolution. In satellite image analysis using 2-D wavelet transform techniques, rows and columns of image pixels are considered signals. The approximation (A) and details (Ds) of a two-dimensional image $f(x, y)$ at spatial resolution 2^j can be defined by the coefficients computed by the following convolutions:

$$A^d_{2^j} f = ((f(x, y) * \phi_{2^j}(-x)\phi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in Z^2} \quad (1)$$

$$D^1_{2^j} f = ((f(x, y) * \phi_{2^j}(-x)\Psi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in Z^2} \quad (2)$$

$$D^2_{2^j} f = ((f(x, y) * \Psi_{2^j}(-x)\phi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in Z^2} \quad (3)$$

$$D^3_{2^j} f = ((f(x, y) * \Psi_{2^j}(-x)\Psi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in Z^2} \quad (4)$$

where integer j is a decomposition level, m, n are integers, $\phi(x)$ is a one-dimensional scaling function, and $\Psi(x)$ is a one-dimensional wavelet function. In general, $\phi(x)$ is a smoothing function, which provides low frequency information (low-pass filter), and $\Psi(x)$ is a differencing function, which provides high frequency information (high-pass filter). $A^d_{2^{j+1}} f$ can be perfectly reconstructed from $A^d_{2^j} f$, $D^1_{2^j} f$, $D^2_{2^j} f$, $D^3_{2^j} f$. The expressions 1 through 4 show that in two dimensions, $A^d_{2^j} f$ and $D^k_{2^j} f$ are computed with separable filtering of the signal along the abscissa and ordinate. The wavelet decomposition can thus be interpreted as signal decomposition in a set of independent, spatially oriented frequency channels (Mallat, 1989).

Dilation equation $\phi(x)$ and wavelet equation $\Psi(x)$ can be defined as:

$$\phi(t) = \sqrt{2}\sum_k c(k)\phi(2t - k) \quad (5)$$

$$\Psi(t) = \sqrt{2}\sum_k d(k)\phi(2t - k). \quad (6)$$

For example, Haar, the simplest wavelet transform, has coefficients: $c(0) = c(1) = \frac{1}{\sqrt{2}}$, $d(0) = \frac{1}{\sqrt{2}}$, and $d(1) = -\frac{1}{\sqrt{2}}$. Thus, its dilation equation and wavelet equation can be expressed as:

$$\phi(t) = \phi(2t) + \phi(2t - 1) \quad (7)$$

$$\Psi(t) = \phi(2t) - \phi(2t - 1) \quad (8)$$

(Strang and Nguyen, 1997).

The approximation and detail coefficients can be computed with a pyramid algorithm based on convolutions with the above mentioned two one-dimensional parameter filters. Figure 1 illustrates the application of the filters and the down sampling procedure for computing the approximation and detail coefficients. Approximation of a signal $A^d_{2^j} f$, also known as *trend*, can be obtained by first convolving the rows of the input signal $A^d_{2^{j+1}} f$ with the low pass filter (L). Next, the filtered signals are down sampled by keeping one column out of two. Then, the columns of the resulting signals are convolved with another one-dimensional low pass filter, retaining every other row.

To obtain a horizontal detail image, first the rows of the input image are convolved with a low pass filter L, and the filtered signals are down sampled by keeping one column out of two, as we do in processing approximation images. However, for the next stage, the columns of the signals are convolved with a high pass filter H, and again every other row is retained. Following the above procedure, for the vertical details, the original signals are convolved first with a high pass filter H and then with a low pass filter L. For the diagonal detail image, the same down sampling procedure is carried out, using two high pass filters, consecutively. Figure 2 represents a standard orthonormal wavelet decomposition with two levels of an image. The image is a randomly selected training sample of a residential area generated from the ATLAS image of Baton Rouge, Louisiana, the study area. The pyramid decomposition can be applied continuously to the approximation image until the desired level of spatial resolution $2^{-j}(-1 \geq j \geq -J)$ is reached.

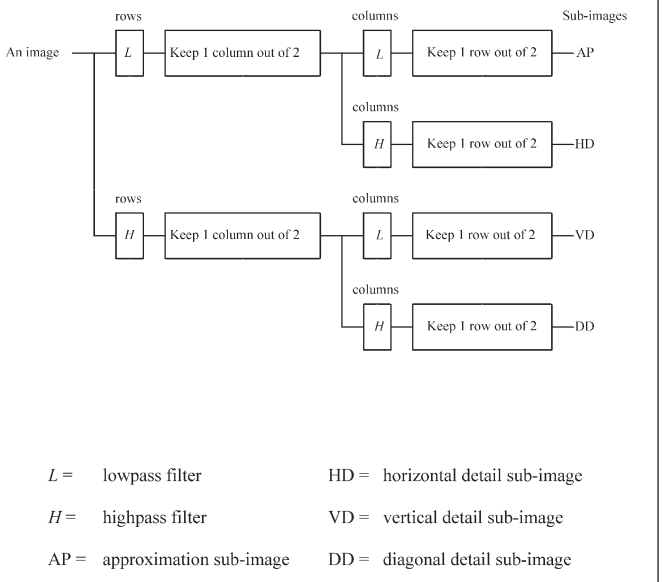

Figure 1. Decomposition procedure of an image by the multiresolution analysis (Modified from Myint *et al.*, 2002).

Data and Study Area

Advanced Thermal Land Application Sensor (ATLAS) image data at 2.5 m spatial resolution acquired with 15 channels (0.45 μm to 12.2 μm) were used for this study. The data were collected by a NASA Stennis LearJet™ 23 flying at 6,600 feet over Baton Rouge, Louisiana, on 07 May 1999 (Figure 3). Lo *et al.* (1997) developed a spatial model of warming and cooling characteristics of commercial, residential, agricultural,

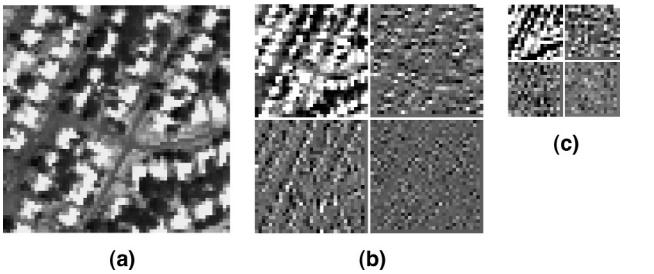

Figure 2. Multiresolution wavelet decomposition of an image (Residential area—Thermal band). (a) original image, (b) wavelet representation at level 1, (c) wavelet representation at level 2. Upper left sub-image of (b) or (c) corresponds to the lowest frequencies (approximation), upper right sub-image gives the horizontal high frequencies (horizontal detail), lower left the vertical high frequencies (vertical detail), and lower right the high frequency in both directions (the diagonal detail).

Figure 3. A subset of Baton Rouge area (Near Infrared Band 0.76 μm–0.90 μm).

vegetation, and water features using a GIS approach. It is important for city planners and environmental officers to observe the relationship among urban land-use land-cover classes associated with surface vegetation, water availability, and associated temperature fluctuation within an urban area. These are crucial information for city officials and environmental agencies to develop a better infrastructure management plan to avoid environmental degradation caused by increase in air pollution, noise pollution, traffic congestion, urban heat island effect, chemical contamination, and soil loss. Following Lo *et al.* (1997), six urban land-use and land-cover features with different textural appearances were selected: single-family homes with less than 50 percent tree canopy (residential-1), single-family homes with more than 50 percent tree canopy (residential-2), commercial, woodland, agriculture, and water body (Figure 4). Band-2 (0.52 μm to 0.60 μm), band-6 (0.76 μm to 0.90 μm), and band-12 (9.60 μm to 10.20 μm) were selected for wavelet based texture analysis and classification in this study. In addition to these three bands, principal component analysis (PCA), a technique that compresses redundant data and produces new bands that are often more interpretable than the original data, was applied to all available bands in the ATLAS data. PCA1 band was the only band selected to examine if a composite band could produce an increased accuracy, since it contains the largest percentage of the total scene variance. Figure 5 shows the differences in texture features of the same area in four different bands: band-2, band-6, band-12, and PCA1 band.

Two segmented regions containing each land-use and land-cover class were identified visually with the help of local area knowledge, ground information collection, and existing maps. Five training pixels were then selected randomly from each region leading to a sample of 10 pixels. Unlike the previous studies where 65 × 65 and 33 × 33 local window sizes were used (Myint, 2001; Myint, *et al.*, 2002), a 17 × 17 square window size (corresponding to 42.5m × 42.5m) centered at the randomly selected pixels in each region was used to subset training samples (Figure 6). The purpose is to examine if a smaller local window, with the use of the wavelet transform techniques, produces a satisfactory accuracy. The use of 17 × 17 pixels still exceeds the minimum distance required to cover individual texture features containing each land-cover class. It should be noted that smaller window does not cover as much spatial/texture information as larger window does. On the other hand, if the window size is too large, too much information from other land-cover features could be included, and hence the algorithm might not produce accurate results.

To be consistent with previous studies and for comparison purposes, we selected all 17 × 17 samples from the previously selected larger homogeneous texture samples (i.e., 33 × 33, 65 × 65). All 17 × 17 samples were subsets of the respective 33 × 33 and 65 × 65 samples. We define a homogeneous texture sample as a sample without having two or more visually

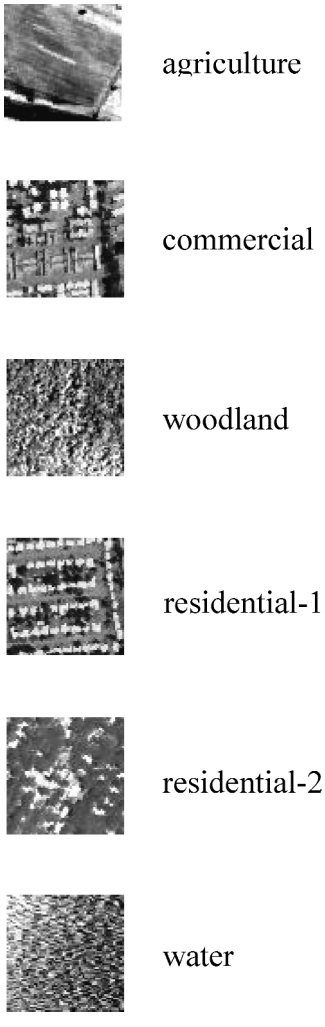

Figure 4. Samples of six urban texture features (PCA1 band).

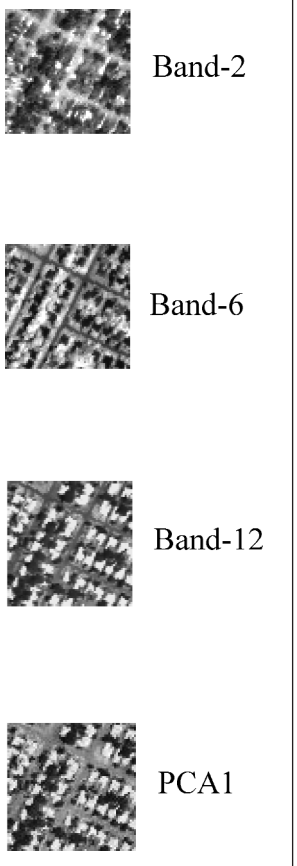

Figure 5. Different texture appearances of a sample (Residential-1) shown in each of the four selected bands: Band-2 (0.52 μm–0.60 μm), Band-6 (0.76 μm–0.90 μm), band-12 (9.60 μm–10.20 μm), and PCA1.

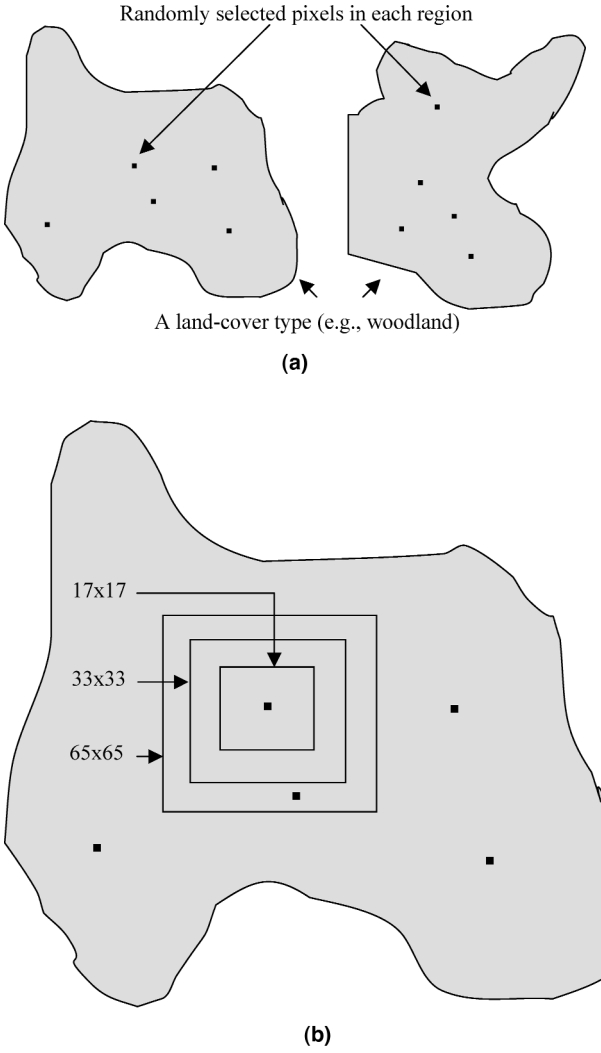

Figure 6. (a) Two visually identified segmented regions of a land-cover type (e.g., Woodland) with randomly selected points/pixels; (b) Diagram of three different sizes of samples (17 × 17, 33 × 33, 65 × 65) centered at the randomly selected pixels from a visually identified region.

different texture features of land-use or land-cover classes. In other words, texture appearances of samples in a land-use or land-cover class are visually similar to each other. Criteria, for the selection of sample size, include image spatial resolution, minimum mapping unit size, and the nature of the classes (minimum distance to completely cover a class, repetition of spatial objects and features of a class, size of the regions, classes selected, level of heterogeneity, and orientation of features) to be identified. This paper reports the classification accuracy of all four approaches (wavelets, spatial co-occurrence, spatial autocorrelation, fractals) for the selected window sizes (i.e., 17 × 17, 33 × 33, 65 × 65). Moreover, based on previous results which show that band-2, band-6, and band-12 produced an increased accuracy, this study will therefore examine the performance of all approaches using these three bands both separately and in combination.

Analysis Procedure

Wavelet Transform and Its Textural Measures

In general, further decomposition is performed in the low frequency channels. This is known as the standard wavelet decomposition. However, decomposition can also be done with the high frequency sub-images. Myint *et al*. (2002) demonstrated the effectiveness of four different wavelet decomposition procedures with 33 × 33 samples: (1) standard decomposition, (2) horizontal decomposition, (3) vertical decomposition, and (4) diagonal decomposition. Figure 7a illustrates the first level decomposition of an original image and Figure 7b through 7e represent four different decomposition procedures with three levels of an original image. It was demonstrated in the previous analyses that with 33 × 33 samples, the standard decomposition was the most accurate among all decomposition approaches. For this reason, the standard decomposition was the only approach employed in this analysis to evaluate the discriminant power of texture measures using 17 × 17 samples. To evaluate the power of the wavelet transform technique for texture classification, a linear discriminant analysis approach was used. The texture measures (e.g., entropy of the decomposed sub-images) for the sample generated above were subjected to discriminant analysis using the Minitab© software package. The procedure generates a discriminate function (or, for more than two groups, a set of discriminant functions) based

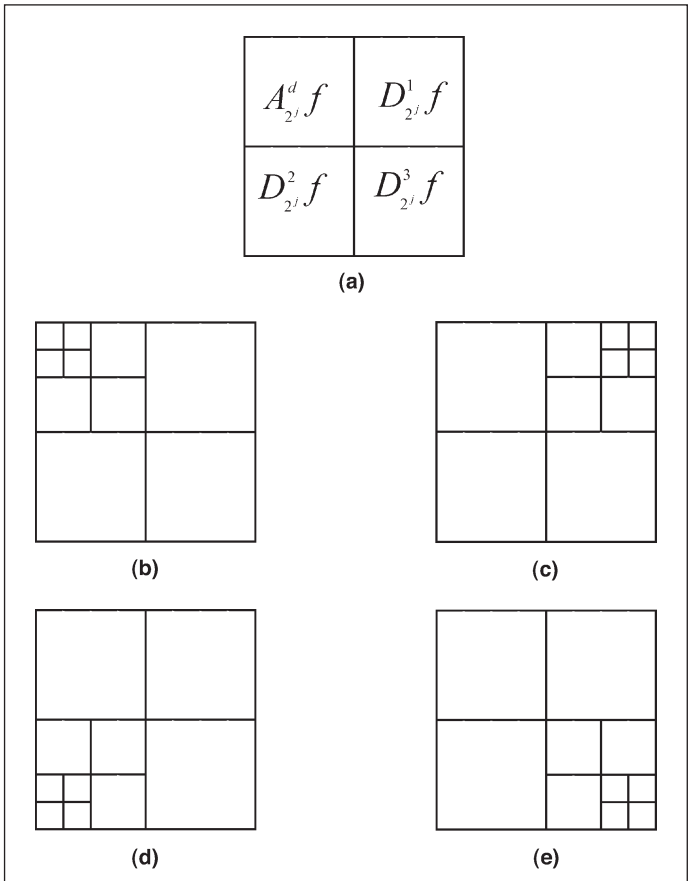

Figure 7. Four different decomposition procedures with three levels of an original image; (a) decomposition at the first level, (b) standard, (c) horizontal, (d) vertical, and (e) diagonal decompositions (Adapted from Myint *et al.*, 2002).

on linear combinations of the predictor variables which provide the best discrimination between the groups.

Haralick *et al.* (1973) proposed a variety of measures to extract textural information from the gray level co-occurrence matrices (GLCM), also known as the spatial co-occurrence matrix. Gong and Howarth (1992) demonstrated the efficiency of occurrence frequency methods with the use of several measures: mean, standard deviation, skewness, kurtosis, range, and entropy measures. Zhu and Yang (1998) used entropy as a measure to identify texture features in twenty five types of aerial relief samples selected from remote sensing images. Albuz *et al.* (1999) used the sum of squares of the wavelet coefficients of each sub-band for their image retrieval system. Sheikholeslami *et al.* (1999) calculated the mean and variance of wavelet coefficients to represent the contrast of the image. In this study, we used the same four texture measures as in the previous studies (Myint, 2001; Myint *et al.*, 2002): log energy (*LOG*), Shannon's index (*SHAN*), entropy (*ENT*), and angular second moment (*ASM*) to characterize the texture of the land-cover classes. The four measures are defined as follows:

$$LOG = \sum_{i=1}^{K} \sum_{j=1}^{K} \log(P(i, j)^2) \tag{9}$$

$$SHAN = -\sum_{i=1}^{K} \sum_{j=1}^{K} P(i, j) * \log(P(i, j)) \tag{10}$$

$$ASM = \sum_{i=1}^{K} \sum_{j=1}^{K} P(i, j)^2 \tag{11}$$

$$ENT = -\sum_{i=1}^{K} \sum_{j=1}^{K} Q(i, j) * \log|Q(i, j)|;$$

$$Q(i, j) = \frac{|P(i, j)|^2}{\sqrt{\sum_{i,j} |P(i, j)|^2}} \tag{12}$$

where $P(i, j)$ is the $(i, j)th$ pixel wavelet coefficient value of a decomposed image at a particular level. The Matlab© software package was used to compute the above mentioned texture measures in the analysis. The computed texture feature values of the decomposed sub-images of each original sample at different levels were used to discriminate samples of urban classes.

Fractal, Spatial Autocorrelation, and Spatial Co-occurrence Matrix

For comparison purposes and to better evaluate the efficiency of wavelet approaches in image classification, two different fractal approaches (isarithm and triangular prism), spatial autocorrelation techniques (Moran's *I* and Geary's *C*), and spatial co-occurrence matrix of the selected features were examined in this study. The discriminant analysis was carried out to discriminate between classes of urban land use land cover on the basis of these texture measures.

Fractal Analysis

There is voluminous literature on the concepts and uses of fractals since Mandelbrot coined the term in 1975 (Mandelbrot, 1983). A potential use of fractals is the analysis of image texture (De Jong and Burrough, 1995). In these situations, it is commonly observed that the degree of roughness or large brightness differences in short spatial intervals in an image or surface is a function of scale, and not of experimental technique. A number of fractal algorithms have been programmed into a software package known as the Image Characterization and Modeling System (ICAMS) (Quattrochi *et al.*, 1997; Lam *et al.*, 1998) which was used in this study to explore the texture features of the study area. ICAMS provides the ability to calculate the fractal dimension of remotely sensed images using the isarithm method (Lam and De Cola, 1993), the variogram (Mark and Aronson, 1984), and the triangular prism methods (Clarke, 1986). For detailed description of the fractal approaches, see Jaggi *et al.* (1993). The variogram method was not included in this analysis since it tends to be unstable with non-stationary signals or features (e.g., remotely sensed images) (Lam *et al.*, 2002).

Spatial Autocorrelation

ICAMS also contains modules for analyzing the spatial autocorrelation of images. Moran's *I* and Geary's *C* are two indices of spatial autocorrelation which reflect the differing spatial structures of the smooth and rough surfaces. Moran's *I* is calculated from the following formula:

$$I(d) = \frac{n \sum_{i}^{n} \sum_{j}^{n} w_{ij} z_i z_j}{W \sum_{i}^{n} z_i^2}$$

where w_{ij} is the weight at distance d so that $w_{ij} = 1$ if point j is within distance d from point i; otherwise, $w_{ij} = 0$; z's are deviations (i.e., $z_i = y_i - y_{\text{mean}}$ for variable y), and W is the sum of all the weights where $i \neq j$. Moran's *I* varies from +1.0 for perfect positive correlation (a clumped pattern) to −1.0 for perfect negative correlation (a checkerboard pattern).

Geary's *C* contiguity ratio, another index of spatial autocorrelation, is similar to Moran's *I,* but uses the formula:

$$C(d) = \frac{(n-1) \sum_{i}^{n} \sum_{j}^{n} w_{ij} (y_i - y_j)^2}{2W \sum_{i}^{n} z_i^2}.$$

With the same terms listed above. Geary's *C* normally ranges from 0.0 to 3.0, with 0.0 indicating positive correlation, 1.0 indicating no correlation, and values greater than 1.0 indicating negative correlation.

Spatial Co-occurrence Matrix and Its Textural Measures
One commonly applied statistical procedure for interpreting texture uses image spatial co-occurrence matrix (SCM), which is also known as gray level co-occurrence matrix (GLCM) (Franklin *et al.*, 2000; Pesaresi, 2000). The use of SCM or GLCM in texture analysis is also referred to as the spatial gray level dependence method. Spatial co-occurrence matrix is a widely used texture and pattern recognition technique in the analysis of remotely sensed data, and it has been successful to a certain extent. There are a number of texture measures which could be applied to spatial co-occurrence matrices for texture analysis (Haralick *et al.*, 1973; Corners and Harlow, 1980; Peddle and Franklin, 1991).

The construction of the four directional spatial co-occurrence matrices for a 3×3 window from an example image normalized to four gray levels (0 to 3) is illustrated in Figure 8. The final matrix for a given point location in the image contains the number of times each possible pair of pixel values occurred in the selected orientation (e.g., horizontal 0 degrees, vertical 90 degrees, left diagonal 135 degrees, and right diagonal 45 degrees) within the specified neighborhood surrounding that point (Figure 8c). In this study, five texture measures based on the co-occurrence matrix were used: angular second moment or Energy (*ASM*); entropy (*ENT*); homogeneity or inverse difference moment (*IDM*); contrast or Inertia (*CON*); and correlation (*COR*). The formulae used to compute each measure from the spatial co-occurrence matrix and their performance in general are as follows:

Angular Second Moment (or) Energy (*ASM*) $= \sum_i \sum_j \left\{ \frac{P(i,j)}{R} \right\}^2$,

Entropy (*ENT*) $= \sum_i \sum_j \left(\frac{P(i,j)}{R} \right) * \log\left(\frac{P(i,j)}{R} \right)$,

Inverse Difference Moment (or)

Homogeneity (*IDM*) $= \sum_i \sum_j \frac{1}{1+(i-j)^2} \left(\frac{P(i,j)}{R} \right)$,

Contrast (or) Inertia (*CON*) $= \sum_i \sum_j \frac{(i-j)^2 P(i,j)}{R}$,

Correlation (*COR*) $= \sum \sum \frac{(i-\mu_x)(j-\mu_y)P_{ij}}{\sigma_x \sigma_y}$.

The computed texture feature values of the samples were used to discriminate selected urban land-use and land-cover classes.

Results and Discussion

Multi-band Approach with 17 × 17 Samples

Table 1 shows the accuracy of the urban land cover classes by linear discriminant analysis on the decomposed sub-images at two separate levels and combination of two levels for 17×17

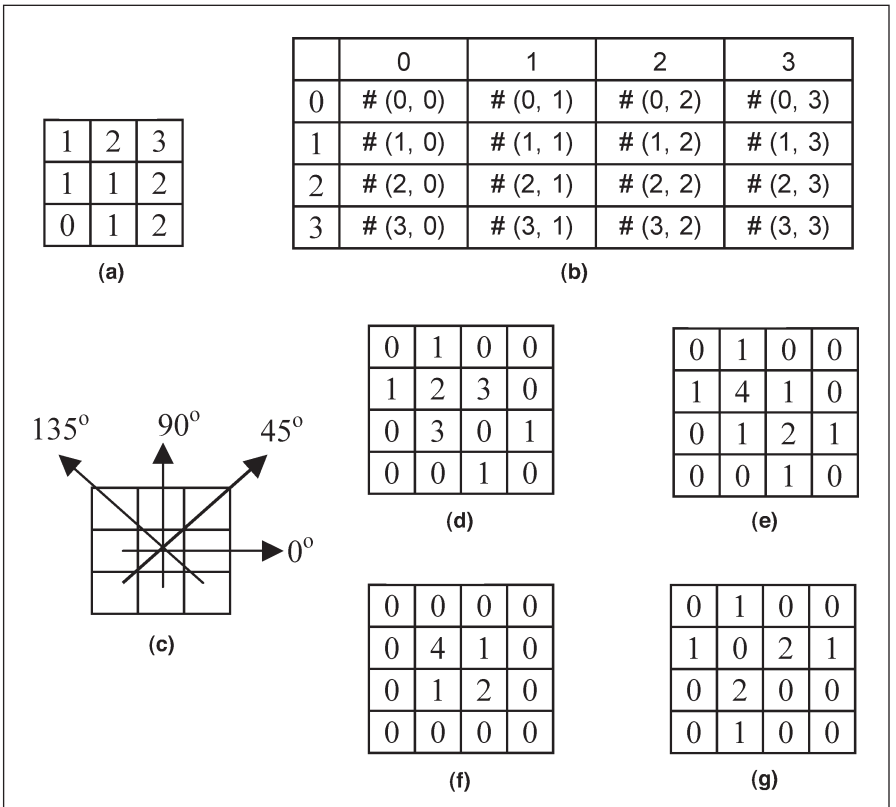

Figure 8. (a) 3×3 window with gray tone range 0 to 3; (b) General form of any spatial co-occurrence matrix for window with gray tone range 0 to 3. Number (*i*, *j*) represents number of times gray tones *i* and *j* were neighbors; (c) selected orientations (e.g., horizontal-0 degrees, vertical-90 degrees, left diagonal-135 degrees, and right diagonal-45 degrees); (d)–(f) spatial co-occurrence matrices derived for four angular orientations (Peddle and Franklin, 1991).

TABLE 1. CLASSIFICATION OF 17 × 17 SAMPLES USING THE STANDARD DECOMPOSITION PROCEDURE

		Overall Accuray (%)				
		(Measures)				
Bands	L	NSI	*LOG*	*SHAN*	*ENT*	*ASM*
Band-2	l_1	4.0	48.3	68.3	65.0	70.0
(visible)	l_2	4.0	36.7	48.3	40.0	50.0
	l_1–l_2	8.0	60.0	76.7	71.7	76.7
Band-6	l_1	4.0	61.7	65.0	56.7	65.0
(near infrared)	l_2	4.0	58.3	58.3	56.7	50.0
	l_1–l_2	8.0	78.3	73.3	73.3	75.0
Band-12	l_1	4.0	55.0	66.7	56.7	66.7
(thermal)	l_2	4.0	46.7	45.0	51.7	43.3
	l_1–l_2	8.0	63.3	75.0	66.7	73.3
PCA1	l_1	4.0	51.7	66.7	63.3	66.7
	l_2	4.0	46.7	45.0	51.7	45.0
	l_1–l_2	8.0	73.3	68.3	71.7	73.3

L = wavelet decomposition level; NSI = number of sub-images.

samples using the *LOG*, *SHAN*, *ENT*, and *ASM* measures. As expected, Table 1 shows that using a combination of two decomposition levels would achieve higher accuracy than using only one level. Also, higher levels produce lower accuracies. band-6 was found to be the most effective band in characterizing urban texture features since 17 × 17 samples of band-6 with combination of two levels using the *LOG* measure reaches an overall accuracy as high as 78 percent. This result is consistent with what was reported earlier (Myint, 2001) with the use of 65 × 65 samples.

Linear discriminant analysis with the combination of two bands (i.e., band-6 and band 12) was carried out to increase the classification accuracy of 17 × 17 samples. Table 2 shows that the multi-band approach dramatically increase the overall accuracy. In the case of combining band-6 and band-12, the highest accuracy (92 percent) achieved in the classification was produced by a combination of two levels with the *LOG* or *ENT* measure. In the combination of all three bands, the highest accuracy (98 percent) achieved in the classification was produced by combining the two levels with the *SHAN* measure. Clearly, a combination of three bands significantly increases the overall accuracy. The second highest accuracy (97 percent) was produced by the combination of two levels with the *LOG, ENT* or *ASM* measure. The overall accuracy produced by the multi-band approach with the combination of two levels exceeds the standard acceptable accuracy 85 percent (Townshend, 1981) for all measures (Figure 9).

TABLE 2. CLASSIFICATION OF 17 × 17 SAMPLES GENERATED FROM MULTI-BAND APPROACH USING THE STANDARD DECOMPOSITION PROCEDURE

		Overall Accuray (%)				
		(Measures)				
Bands	L	NSI	*LOG*	*SHAN*	*ENT*	*ASM*
Band-6 + 12	l_1	4.0	80.0	83.3	78.3	83.3
	l_2	4.0	65.0	58.3	65.0	60.0
	l_1–l_2	8.0	91.7	86.7	91.7	86.7
Band-2 + 6 + 12	l_1	4.0	88.3	91.7	88.3	93.3
	l_2	4.0	73.3	71.7	75.0	68.3
	l_1–l_2	8.0	96.7	98.3	96.7	96.7

L = wavelet decomposition level; NSI = number of sub-images.

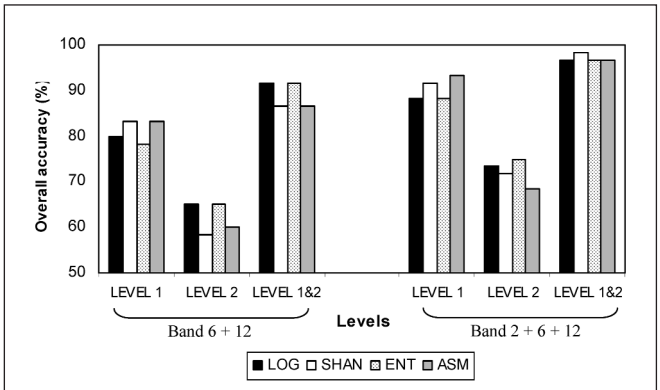

Figure 9. Overall accuracies for *SHAN*, *ENT*, and *ASM* plotted against different band and level combinations (17 × 17 samples).

Comparison with Fractal, Spatial Autocorrelation, and Spatial Co-occurrence Matrix Using 65 × 65, 33 × 33, and 17 × 17 Samples

Figures 10 to 12 present the overall classification accuracies for 65 × 65, 33 × 33, and 17 × 17 samples produced by wavelet transforms using log energy (W-LOG), Shannon's index (W-SHAN), entropy (W-ENT), and angular second moment (W-ASM); fractal analysis using isarithm (Isarithm) and triangular prism (Triangular) methods; spatial autocorrelation using Moran's *I* (Moran) and Geary's *C* (Geary); and spatial co-occurrence matrix using angular second moment or Energy (C-ASM), entropy (C-ENT), homogeneity or inverse difference moment (C-IDM), contrast or Inertia (C-CON), and correlation (C-COR). Table 3 also presents the overall classification accuracies produced by different algorithms for the three window sizes: 65 × 65, 33 × 33, and 17 × 17.

No results were available for the fractal approaches for 33 × 33 and 17 × 17 samples since ICAMS does not recommend the computation of fractal estimates for small window sizes. It may be inappropriate to apply fractal approaches in small local window sizes since self-similarity of features, shapes, areas, and distances need to be observed at different scales (2^N), and a regression analysis between them needs to be performed to estimate the fractal dimension value. For

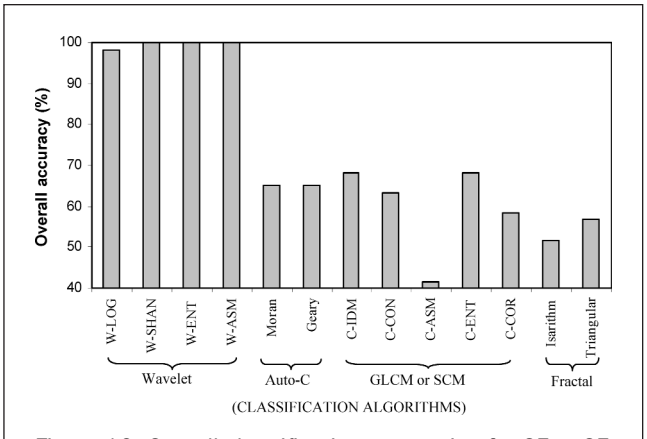

Figure 10. Overall classification accuracies for 65 × 65 samples generated from band-6 produced by wavelet, spatial autocorrelation, spatial co-occurrence matrix, and fractal analysis approaches.

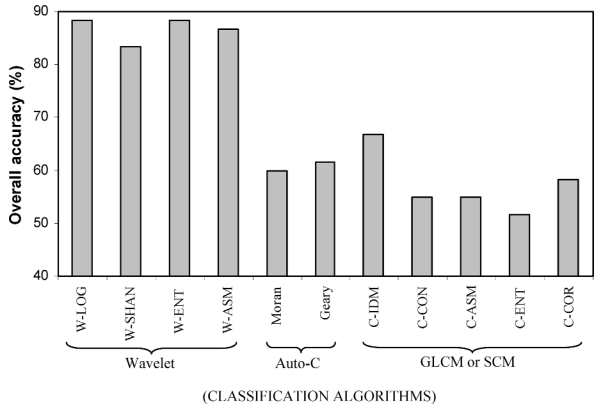

Figure 11. Overall classification accuracies for 33 × 33 samples generated from band-6 produced by wavelet, spatial autocorrelation, spatial co-occurrence matrix, and fractal analysis approaches.

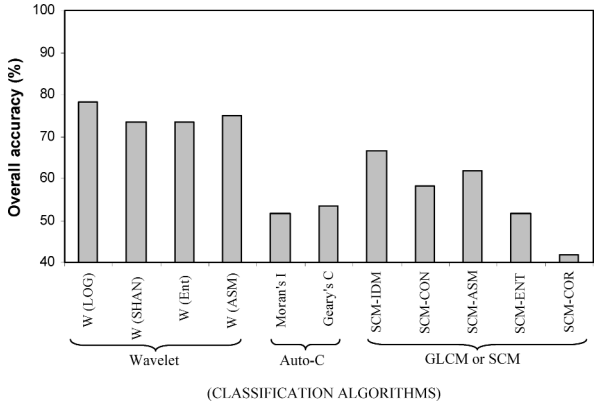

Figure 12. Overall classification accuracies for 17 × 17 samples generated from band-6 produced by wavelet, spatial autocorrelation, spatial co-occurrence matrix, and fractal analysis approaches.

TABLE 3. OVERALL CLASSIFICATION ACCURACIES FOR BAND-6 FOR THREE WINDOW SIZES: 65 × 65, 33 × 33, 17 × 17 PRODUCED BY DIFFERENT ALGORITHMS

		Overall Classification Accuracy (Sample Size)		
Techniques	Measures	65 × 65	33 × 33	17 × 17
Wavelet	W-LOG	98.3	88.3	78.3
	W-SHAN	100.0	83.3	73.3
	W-ENT	100.0	88.3	73.3
	W-ASM	100.0	86.7	75.0
Spatial Autocorrelation	Moran	65.0	60.0	51.7
	Geary	65.0	61.7	53.3
Spatial Co-occurrence Matrix	C-IDM	68.3	66.7	66.7
	C-CON	63.3	55.0	58.3
	C-ASM	41.7	55.0	61.7
	C-ENT	68.3	51.7	51.7
	C-COR	58.3	58.3	41.7
Fractal	Isarithm	51.7	—	—
	Triangular	56.7	—	—

example, we can obtain only a few steps or observations (e.g., $2^5 = 32$) from a 33 × 33 sample to perform a regression analysis for its fractal estimate. Figures 10 to 12 compare the accuracies produced by different textural approaches for band-6. Table 3 shows that fractal analysis yields the lowest accuracy. For fractal analysis a local window, the size needs to be sufficiently large since all fractal approaches require a regression on the logarithm of number of edges, total surface areas, or variance against the logarithm of the cell size to perform the regression analysis that determines the fractal dimension value. Hence, an appropriate window size is very important for more accurate fractal dimension estimates (Myint, 2003). However, smaller window size is generally thought to yield higher accuracy in image classification because if the window is too large, much spatial information from other land-cover features could be included, and this will create a mixed-boundary feature problem. This is especially true in a very complex urban environment; it could be very difficult for a visual interpreter to identify and allocate a particular sample to a class, especially when dealing with the very highly heterogeneous nature of mixed urban features. In this application, a very accurate algorithm using a large local window may not produce satisfactory accuracies.

From Figures 10 through 12, it can be observed that a wavelet analysis technique is far more accurate than other well-known and widely accepted methods. It was reported earlier (Myint, 2001) that the wavelet transforms approach with three different measures: ENT, ASM, and SHAN obtained the highest accuracy, 100 percent for the 65 × 65 samples (Table 3). The same approach with the LOG measure yielded 98 percent accuracy which is still much higher than accuracies produced by any other widely accepted approach. The spatial co-occurrence matrix with the IDM or ENT measure achieved the next higher accuracy (68 percent). It is noted that there is a large gap between the lowest accuracy (98 percent) achieved by wavelets versus the highest accuracies (68 percent) achieved by the other advanced texture methods (fractal, spatial autocorrelation, and spatial co-occurrence matrix approaches).

In general, the highest accuracies for the 65 × 65 samples produced by different texture methods are the wavelet transform (100 percent), spatial autocorrelation (65 percent), spatial co-occurrence (68 percent), and fractal analysis (57 percent). The highest accuracies achieved by different approaches for the 33 × 33 samples are the wavelet transform (88 percent), spatial autocorrelation (62 percent), and spatial co-occurrence (67 percent). The lowest accuracy of each method was produced by W-SHAN (83 percent), Moran (60 percent), and C-ENT (52 percent). It is difficult to determine which approach is more efficient between the spatial autocorrelation and spatial co-occurrence matrix since the C-IDM approach produced more accurate results than both autocorrelation methods, whereas the remainder of the co-occurrence approaches produced lower accuracy than that of Moran's *I* and Geary's *C*. However, it is obvious that the accuracies for 33 × 33 samples achieved by all wavelet transform approaches are much higher than the other three texture methods.

The overall classification accuracies for the 17 × 17 samples are relatively low, ranging from 73 percent to 78 percent for the wavelet approach, autocorrelation from 52 percent to 53 percent, and co-occurrence from 42 percent to 67 percent. It is difficult to make a conclusion on which approach is more accurate between spatial autocorrelation and spatial co-occurrence techniques, since the latter produced the highest and lowest accuracies, and the former is in between. Among the five texture measures of the spatial co-occurrence matrix, C-IDM appears stable since it consistently produced better accuracy for all samples. In general, all accuracies obtained for the 17 × 17 samples were relatively low. However, the overall

accuracies were significantly increased in wavelet analysis with 17 × 17 samples when it included a multi-band approach. As demonstrated earlier (Table 2), the combination of all three bands: band-2, band-6, and band-12 with the SHAN measure achieved the highest accuracy (98 percent) for 17 × 17 samples in the standard decomposition procedure. All other measures (LOG, ENT, and ASM) of the same wavelet approach produced 97 percent accuracy for the same samples.

Conclusion

Results from this study confirm that the wavelet approach is the most accurate of all approaches considered. Regarding the power of spectral bands, band-6 (near infrared) was found to be the most efficient band for 17 × 17 samples since it produced the highest accuracy (78 percent). This result is partly consistent with the results from the previous studies, where band-6 was found to be the most accurate (100 percent) from using 65 × 65 samples, and PCA1 and band-12 produced the highest accuracy (93 percent) when 33 × 33 samples were used. Therefore, it may be difficult to make a general conclusion on which band performed most effectively among band-6, band-12, and PCA1. The 65 × 65 samples gave the highest accuracy among the three window sizes, implying that a window size of 162.5 m by 162.5 m may be appropriate for urban land cover classification. This study confirms that classification accuracy decreases with increasing level of decomposition, but the combination of more levels effectively increases the accuracy. Furthermore, a multi-band approach also drastically increases the overall accuracy.

In this research, the samples were selected randomly from the homogeneous texture regions. It is obvious that they are truly representative samples of the selected classes. Hence, it is recommended that, in addition to the number of training samples, more detailed evaluations should be carried out so that an optimal window size can be determined for different land cover types in different remote sensing imagery. Future research should be done to develop operational algorithms to perform wavelet analysis using different local moving window sizes. Moreover, other types of wavelet transforms will need to be tested. There are a number of other wavelet transforms that may perform differently than the Haar transform used in this study. Examination of these new transforms and how they perform with different window sizes and land cover types will be important areas of research for further increasing the performance of wavelet transform based texture analysis and classification. Furthermore, the potential of the wavelet analysis proposed in this study should also be examined in other environment and application areas.

References

Albuz, Elif, E. Kocalar, A.A. Khokhar, 1999. Vector-wavelet based scalable indexing and retrieval system for large color image archives, *IEEE Transactions on Geoscience and Remote Sensing*, 44:3021–3024.

Arai, K., 1993. A classification method with a spatial-spectral variability, *International Journal of Remote Sensing*, 14:699–709.

Clarke, K.C., 1986. Computation of the Fractal Dimension of Topographic Surfaces Using the Triangular Prism Surface Area Method, *Computers and Geosciences* 12(5):713–722.

Conners, R.W., and C. Harlow, 1980. A theoretical comparison of texture algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(3):204–222.

Cushnie, J.L., 1987. The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies, *International Journal of Remote Sensing*, 8:12–29.

De Jong, S.M., and P.A. Burrough, 1995. A Fractal Approach to the classification of Mediterranean Vegetation Types in Remotely Sensed Images, *Photogrammetric Engineering & Remote Sensing*, 61:1041–1053.

Emerson, C.W., N.S.N Lam, and D.A. Quattrochi, 1999. Multi-Scale Fractal Analysis of Image Texture and Pattern, *Photogrammetric Engineering & Remote Sensing*, 65(1):51–61.

Edwards G., R. Landary, and K.P.B. Thomson, 1988. Texture analysis of forest regeneration sites in high-resolution SAR imagery, *Proceedings of the International Geosciences and Remote Sensing Symposium* (IGARSS 88), ESA SP-284 (Paris: European Space Agency), pp. 1355–1360.

Franklin, S.E., R.J. Hall, L.M. Moskal, A.J. Maudie, and M.B. Lavigne, 2000. Incorporating texture into classification of forest species composition from airborne multispectral images, *International Journal of Remote Sensing*, 21(1):61–79.

Fung, T., and K. Chan., 1994. Spatial composition of spectral classes: A structural approach for image analysis of heterogeneous land-use and land-cover types, *Photogrammetric Engineering & Remote Sensing*, 60(2):173–180.

Gong, P., 1994. Reducing boundary effects in a kernel-based classifier, *International Journal of Remote Sensing*, 15:1131–1139.

Gong, P., and P.J. Howarth, 1990. The use of structural information for improving land cover classification accuracies at the rural urban fringe, *Photogrammetric Engineering & Remote Sensing*, 56(1): 67–73.

_____. 1992. Frequency based contextual classification and gray level vector reduction for land use identification, *Photogrammetric Engineering & Remote Sensing*, 58(4):423–437.

Haralick, R.M., K. Shanmugan, and J. Dinstein, 1973. Textural features for image classification, *IEEE Transaction on Systems, Man, and Cybernetics*, SMC-3(6):610–621.

Hodgson, M.E., 1998. What size window for image classification? A cognitive perspective, *Photogrammetric Engineering & Remote Sensing*, 64(8):797–807.

Jaggi, S., D.A. Quattrochi, and N.S.N. Lam, 1993. Implementation and operation of three fractal measurement algorithms for analysis of remote-sensing data, *Computer and Geosciences*, 19(6):745–767.

Lam, N.S.N., and D.A. Quattrochi, 1992. On the Issues of Scale, Resolution, and Fractal Analysis in the Mapping Sciences, *Professional Geographer*, 44(1):88–97.

Lam, N.S.N., and L. De Cola, 1993. Fractal Simulation and Interpolation, *Fractals in Geography* (N.S.N. Lam and L. De Cola, editors), Prentice Hall, Englewood Cliffs, New Jersey, pp. 56–74.

Lam, N.S.N, D. Quattrochi, H. Qui, and W. Zhao, 1998. Environmental assessment and monitoring with image characterization and modeling system using multiscale remote sensing data, *Applied Geographic Studies* 2(2):77–93.

Lam, N.S.-N., H.-L. Qiu, D.A. Quattrochi, and C.W. Emerson, 2002. An evaluation of fractal methods for characterizing image complexity, *Cartography and Geographic Information Science* 29(1): 25–35.

Lo, C.P., D.A. Quatrochi, and J.C. Luvall, 1997. Application of high-resolution thermal infrared remote sensing and GIS to assess the urban heat island effect, *International Journal of Remote Sensing*, 18(2):287–304.

Mallat, S.G., 1989. A Theory for multi-resolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11:674–693.

Mandelbrot, B., 1983. *The Fractal Geometry of Nature*, Freeman and Co., New York.

Mark, D.M., and P.B. Aronson, 1984. Scale Dependent Fractal Dimensions of Topographic Surfaces: An Empirical Investigation with Applications in Geomorphology and Computer Mapping, *Mathematical Geology*, 16:671–683.

Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi, 1996. *Wavelet Toolbox*, The Math Works, Inc., Massachusetts.

Myint, S.W., 2001. A Robust Texture Analysis and Classification Approach for Urban Land-Use and Land-Cover Feature Discrimination, *GeoCarto International*, 16(4):27–38.

Myint, S.W., N.S.N. Lam, J. Tyler, 2002. An Evaluation of Four Different Wavelet Decomposition Procedures for Spatial Feature Discrimination Within and Around Urban Areas, *Transactions in GIS*, 6(4): 403–429.

Myint, S.W., 2003. Fractal Approaches in Texture Analysis and Classification of Remotely Sensed Data: Comparisons with Spatial Autocorrelation Techniques and Simple Descriptive Statistics, *International Journal of Remote Sensing*, 24(9):1925–1947.

Peddle, D.R., and S.E. Franklin, 1991. Image Texture Processing and Data Integration for Surface Pattern Discrimination, *Photogrammetric Engineering & Remote Sensing*, 57:413–20.

Pesaresi, M., 2000. Texture Analysis for Urban Pattern Recognition Using Fine-resolution Panchromatic Satellite Imagery, *Geographical and Environmental Modeling*, 4(1):43–63.

Quattrochi, D.A., N.S.N. Lam, H. Qiu, and W. Zhao, 1997. Image Characterization and Modeling System (ICAMS): A Geographic Information System for the Characterization and Modeling of Multiscale Remote Sensing Data, *Scale in Remote Sensing and GIS* (D.A. Quattrochi and M.F. Goodchild, editors), CRC Press, Boca Raton, Florida, pp. 295–308.

Sadler, G.J., M.J. Barnsley, and S.L. Barr, 1991. Information extraction from remotely-sensed images for urban land analysis, *Proceedings of the Second European Conference on Geographical Information Systems (EGIS'91)*, Brussels, Belgium, April, EGIS Foundation, Utrecht, pp. 955–964.

Sheikholeslami, G., A. Zhang, and L. Bian, 1999. A Multi-Resolution Content-Based Retrieval Approach for Geographic Images, *Geoinformatica*, 3(2):109–139.

Strang, G., and T. Nguyen, 1997. *Wavelets and Filter Banks*, Wellesley-Cambridge Press, revised edition, 1997, Wellesley, Massachusetts, 520 p.

Townshend, J.R.G., 1981. *Terrain Analysis and Remote Sensing*. George Allen and Unwin: London: 272 p.

Woodcock, C., and V.J. Harward, 1992. Nested-hierarchical scene models and image segmentation, *International Journal of Remote Sensing*, 13:3167–3187.

Woodcock, C.E., A.H. Strahler, and D.L.B. Jupp, 1988. The use of variograms in remote sensing: I. Scene models and simulated images, *Remote Sensing of Environment*, 25:323–348.

Zhu, C., and X. Yang, 1998. Study of remote sensing image texture analysis and classification using wavelet, *International Journal of Remote Sensing*, 13:3167–3187.

(Received 29 November 2002; accepted 05 June 2003; revised 15 July 2003)