Spatial and Grayscale Metadata for Similarity Searches of Image Databases

Charles W. Emerson[1]
Department of Geography, Western Michigan University, Kalamazoo, Michigan 49008-5424

Sivagurunathan Chinniah
PSMA Australia Limited, Griffith ACT 2603, Australia

Nina Siu-Ngan Lam
Louisiana State University, Department of Environmental Studies, Baton Rouge, Louisiana 70803

Dale A. Quattrochi
National Aeronautics and Space Administration, Earth Science Office, VP61, Marshall Space Flight Center, Huntsville, Alabama 35812

Abstract: This paper presents a content-based image retrieval process wherein the user identifies a feature of interest using a region quadtree decomposition of the image, spatial statistics and histograms of the grayscale values of the feature definition are calculated, and the result is compared to a database of these same calculations that have been performed on similar images. The sum of squared differences between the indices calculated for the quads that form the feature of interest and corresponding quads in the database yields a ranked list of matching image tiles. In an analysis of Landsat 7 imagery of North Georgia and an IKONOS panchromatic image of Kalamazoo, Michigan, we found that the retrieval success rate depends on the spatial and spectral characteristics of the feature of interest and the configuration of quads used to define the feature.

INTRODUCTION

Satellite and aircraft-borne remote sensors have gathered huge volumes of data over the past 30 years. Presently, earth resources sensors are generating terabytes of image data every day. As the geographical and temporal coverage, the spectral and spatial resolution, and the number of individual sensors increase, the sheer volume

[1]charles.emerson@wmich.edu

GIScience & Remote Sensing, 2007, 44, No. 2, p. 182–201.

and complexity of available data sets will complicate management and use of the rapidly growing archive of earth imagery. This problem is a subset of the much larger challenge of organizing and indexing all types of digital data. The vast amount of information on the World Wide Web would be of little use without a means to locate information on topics selected by a user. Search engines that rely on keyword matches between the query and Web page titles or other indexed data are essential for successful use of this resource. However, indexing multimedia data, such as imagery, videos, and audio files have proven to be problematic (Paquet, et al., 2000).

The many existing and potential uses for remotely sensed imagery make accessing images suited to a particular user's needs extremely complex and difficult. A single scene often covers a large part of the Earth's surface, so it may also take a lengthy manual search to find other occurrences of a feature of interest in a particular image. Even seemingly simple searches for images depicting a particular location involve time-consuming analyses of the many individual scenes that have been gathered over the past 30 or more years, each having different sensor platforms, levels of quality (due to cloud cover, illumination, etc.), dates, and pre-processing.

Metadata schemes such as the Earth Observing System Data and Information System (EOSDIS) Core Metadata Model (http://ecsinfo.gsfc.nasa.gov) address this to some extent by specifying location, lineage (including image processing and projection information), sensor type, and other identifying characteristics to aid searches for images of specific areas at specific times. Ohm et al., (2000) characterize these as "high-level descriptors," which are generated when raw imagery is prepared for release. Mid-level descriptors include rule-based semantic identification of features within a scene such as lakes, mountains, and vegetated areas. By their nature, the mid-level descriptors are often user-specific, and it would not be practical to add all of this information as formal metadata, since a typical satellite image contains a large number of identifiable features, and it is impossible to anticipate all uses to which an image may be applied. However, low-level descriptors (image characteristics such as shape, color, pattern, and texture) can provide useful metadata on an image's content if it is indexed and organized for efficient retrieval.

Content-based image retrieval is the process of selecting images from an archive based on semantic and visual contents. This necessarily involves high- to low-level characteristics of the image itself (Smeulders et al., 2000). Most of the applications reported in the pattern recognition literature (Datcu, et al, 2003; Yao and Chen, 2003; Li and Narayanan, 2004), generally use some type of supervised or unsupervised image classification technique to assist image retrieval, although Manjunath and Ma (1996) demonstrate a purely texture-based approach. Image classification generally requires some sort of human intervention, such as the selection of training sites for supervised techniques or the linking of cluster signatures to land cover classes in the case of unsupervised classification. However, the approach presented here is more automated, requiring human intervention only in the feature identification stage.

Because a digital image is a raster characterization of a continuous surface of reflected or emitted light energy, large discrete objects such as buildings or water bodies are simply groups of homogeneous pixels set against a background of heterogeneous pixels that may represent mixtures of small features or land covers that form a complex mosaic of different materials. Traditional unsupervised and supervised image classification procedures consider only spectral characteristics on a pixel-

by-pixel basis, and therefore do not consider spatial characteristics of the features depicted in an image. Texture indices usually involve the use of a moving window or co-occurrence matrix of a defined dimension, and thus they explicitly measure spatial relationships. Object-oriented image classification procedures (Baatz and Schape, 2000; Blaschke, et al., 2000) consider both spectral and spatial characteristics. However, the segmentation of an image into homogeneous groups of pixels (or "objects") involves the specification of shape and scale parameters that limit the utility of this technique in a metadata context, because the varying size and shape of potential features of interest determine the degree of segmentation and customized hierarchical classification that needs to be performed for a given query. The approach presented here relies on a regular spatial decomposition of an image, thus requiring the user to approximate the characteristic texture, tone, shape, and size of a feature of interest within the hierarchy of this regular decomposition. This feature approximation allows the development of a smaller searchable metadata table that can be accessed more quickly than a list of object characteristics.

Because most current earth-imaging satellites (Landsat 7, SPOT, Ikonos, Quickbird, IRS) include a higher resolution panchromatic image along with lower resolution multispectral images, this pilot-scale study uses the spatial characteristics and grayscale statistics derived from panchromatic images as indices of image content. More powerful and successful image retrieval would inevitably result from the inclusion of color criteria, as measured by the spectral and spatial characteristics of individual spectral bands (at the cost of increased image index database size). This capability is planned for future versions of the software application.

METHODOLOGY

Software

The Image Characterization and Modeling System (ICAMS) (Quattrochi, et al., 1997; Lam et al., 2002) was developed primarily as a test-bed for evaluating the performance of various spatial analytical methods on remotely sensed images and as such, it is intended to work with other commercially available image analysis and GIS software packages. ICAMS includes utilities for contrast stretching, edge detection, wavelet decomposition, and Fourier transforms. It also computes fractal dimension using the box counting, triangular prism, and isarithm methods (Jaggi, et al, 1993), and it measures lacunarity, a scale-dependent measure of the gaps in an imaged pattern (Dong, 2000). The application also includes utilities for measuring Moran's I and Geary's C indices of spatial autocorrelation (Cliff and Ord, 1973). In addition to these global (whole image or user-defined region of interest) measures, ICAMS includes local measures of triangular prism fractal dimension (Clarke, 1986), Moran's I, Geary's C, and Getis' G and G* indices of spatial autocorrelation (Getis and Ord, 1992). These local measures compute the values of these indices in a moving window, thus producing an output image that shows the differences in spatial complexity across a scene. The ICAMS content based image retrieval utility is the subject of this paper and is explained in the section entitled "Similarity Search."

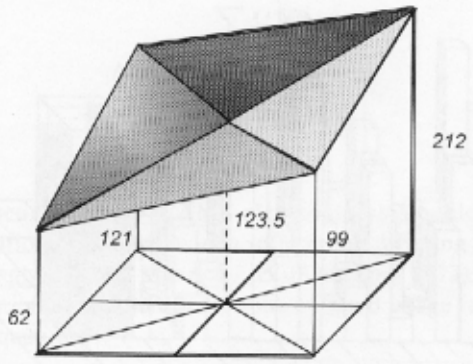

Fig. 1. Triangular prism.

Analytical Methods

Fractal Dimension. Fractal analysis (Mandelbrot, 1983) provides tools for measuring the geometric complexity of imaged features. Fractal dimension (D) is a non-integer value that, in Mandelbrot's (1983) definition for fractals, exceeds the Euclidean topological dimension as the form of a point pattern, a line, or an area feature grows more geometrically complex. The triangular prism method (Clarke, 1986) of estimating the fractal dimension constructs triangles by averaging the z-values (which in this case are the digital numbers) for sets of four adjacent pixels (Fig. 1). The z-values for each pixel are used to establish heights at each corner, and triangles are formed by connecting these corner values to the mean value of the four pixels at the center of the array. The areas of the triangular "facets" of the prisms are then summed to represent the total step 1 surface area. The algorithm then steps to 3×3 arrays of pixels, with the center height corresponding to the average of the four corners. The algorithm continues to increase the pixel size and compute the triangular prism areas until the entire surface is calculated as a single composite array. The logarithm of the total of all the prism facet areas at each step is plotted against the logarithm of the pixel dimension at each step. The fractal dimension is calculated by performing a least squares regression on the surface areas and pixel sizes. The regression slope B is used to determine the fractal dimension D, where:

$$D = 2 - B. \tag{1}$$

In the box counting fractal dimension measurement method (Liebovitch and Toth, 1989; Sarkar and Chaudhuri, 1992), the image is considered to be a three-dimensional space, with x and y being the column and row of the pixels, and the z-value corresponding to the 8-bit grayscale value of the pixel. If we stack a series of three-dimensional boxes of pixel dimension r, so that the maximum z-values are covered by the boxes at each x,y location, we can then determine the number of boxes needed to span the grayscale values from the minimum to the maximum values (Fig.

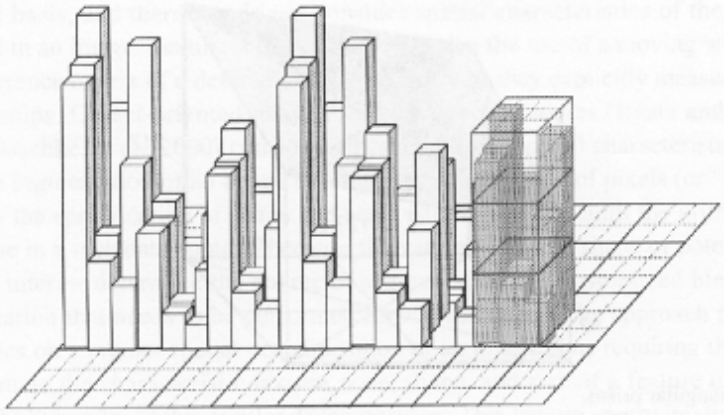

Fig. 2. Box Counting method of measuring fractal dimension.

2). The total count of boxes needed to span all of the grayscale values throughout the image (N_r) is summed in Equation 2 as:

$$N_r = \sum_{i,j} n_r(i,j) \tag{2}$$

where n_r is the number of boxes needed to span the grayscale values at row, column location (i,j). We compute d, the capacity, using a range of values of r and Equation 3:

$$d = \frac{log(N_r)}{log(1/r)} \tag{3}$$

The fractal dimension is computed using Equation 1 and the slope of the least squares line B that best fits the computed values of d for a range of box sizes.

Lacunarity. Lacunarity is a scale-dependent measure that is related to fractal dimension, but instead focuses on the distribution of gaps in a pattern. Low-lacunarity geometric features are homogeneous, with similar gaps occurring at regular intervals if they occur at all. High-lacunarity features have an irregular arrangement of gaps. Since homogeneous patterns can appear heterogeneous over large spatial extents, lacunarity is scale-dependent. The method for measuring lacunarity in ICAMS-Java is derived from the gliding box algorithm proposed by Allain and Cloitre (1991), and developed by Plotnick et al. (1996). Dong (2000) extended the gliding box technique to grayscale images.

The gliding box method is similar to the box counting method described above, except in the case of lacunarity, $Q(M,r)$, the probability function for the number of boxes of size r (called the mass, M) is calculated and used in Equation 4 to estimate the lacunarity Λ:

$$\Lambda(r) = \frac{\sum_M M^2 Q(M, r)}{\left[\sum_M MQ(M, r)\right]^2} \tag{4}$$

Spatial Autocorrelation. ICAMS also contains modules for analyzing the spatial autocorrelation of images, which reflects the differing spatial structures of the smooth and rough surfaces. Moran's I (Cliff and Ord, 1973) is currently the only spatial autocorrelation index used in the content based image retrieval utility and is calculated using Equation 5:

$$I = \frac{n\sum_i^n \sum_j^n w_{ij} z_i z_j}{(n-1)\sum_j^n w_{ij}} \tag{5}$$

where w_{ij} is a binary connectivity measure ($w_{ij} = 1$ if pixel j adjoins pixel i, otherwise $w_{ij} = 0$), user selectable at either rook's case (four pixel neighbors), or queen's case (eight neighbors). The z-values are standardized grayscale deviations (i.e., $z_i = x_i - x_{\text{mean}}$ for variable x). Moran's I varies from +1.0 for perfect positive correlation (a clumped pattern) to −1.0 for perfect negative correlation (a checkerboard pattern).

Wavelet Energy Signatures. ICAMS includes a utility for multi-resolution decomposition of an image using wavelets, or small oscillatory components that operate locally. The general form of a continuous wavelet transform of one-dimensional data W_f can be expressed as (Daubechies, 1991):

$$W_f(a, b) = |a|^{1/2} \int_{-\infty}^{0} f(x)\psi\left(\frac{x-b}{a}\right)dx = \langle f, \psi_{a,b}(x)\rangle \tag{6}$$

where

$$\psi_{a,b}(x) = |a|^{1/2}\psi\left(\frac{x-b}{a}\right) \tag{7}$$

and where a and b are the scale and translation parameters, respectively; and $\Psi(x)$ is a window function, referred to as the mother wavelet. The mother wavelet is a prototype to generate wavelets with variable parameters a and b to scale and shift the mother wavelet so that the original variation in grayscale values is approximated. The scale index a indicates the wavelet's width, and the location index b gives its position.

The wavelet transform starts with the choice of a fixed mother wavelet. There are many types of wavelets, ranging from the simple Haar wavelet to more complex

forms. ICAMS presently uses a relatively simple Daubechies DB4 wavelet (Daubechies, 1988). The continuous wavelet transform is computed by changing the scale of the wavelet, shifting the wavelet in space, multiplying by the signal, and integrating over all locations. The results of the wavelet transform are wavelet coefficients, which measure the variation of the signal in a neighborhood b whose size is proportional to a. In other words, this definition of the continuous wavelet transform shows that the wavelet coefficient is a measure of similarity between the wavelets and the frequency content of the image. The calculated continuous wavelet transform coefficients refer to the closeness of the image to the wavelet at the current scale.

Mallat (1989) proposed a framework for the multi-resolution decomposition of the discrete wavelet transform. Multi-resolution decomposition is normally accomplished by digital filtering techniques in a dyadic fashion, and the scale a is given by $a = 2^{-j}$, where $j = 1, 2, 3 \ldots$ etc., and integer j is a decomposition level. It decomposes an image into a coarser resolution representation that consists of the low-frequency approximation information and the high-frequency detail information. Let A_0 be the image, and H and L be one-dimensional high-pass and one-dimensional low-pass filters, respectively. They are the conjugate mirror filters associated with the wavelet. With the separability of the wavelet basis and the convolution formula, we can obtain:

$$\begin{aligned} A_{j+1} &= A_j * HH \\ D^1_{j+1} &= A_j * LH \\ D^2_{j+1} &= A_j * HL \\ D^3_{j+1} &= A_j * LL, \end{aligned} \tag{8}$$

where A_j is an approximate image of the image A_0 at a spatial resolution j. A_j is decomposed to an approximate image A_{j+1} and three detail images D^1_{j+1}, D^2_{j+1}, and D^3_{j+1}, and in three orientations (horizontal, vertical, and diagonal) at a lower spatial resolution $j+1$.

The approximate and detail coefficients can be calculated with a pyramid algorithm based on convolutions with the two one-dimensional filters H and L. The result in the output of the low-pass filter L represents the data's low-frequency approximation. The outputs of the high-pass filter H are referred to as the data's high-frequency detail information. The pyramid decomposition can be continuously applied to the approximation image until the desired coarser resolution 2^{-j} ($j \geq 0$, the maximum decomposition level) is reached.

Grayscale Statistics. A five-bin histogram of the grayscale values contained in each quad is stored in five fields in the index database. In order to facilitate comparisons between different levels of the quadtree structure with the different numbers of pixels in each level's quads, the histogram was normalized by converting each frequency bin to a percentage of total pixels in the quad. The five-bin histogram then was considered as a group of values in the ranking of results. The mean and standard

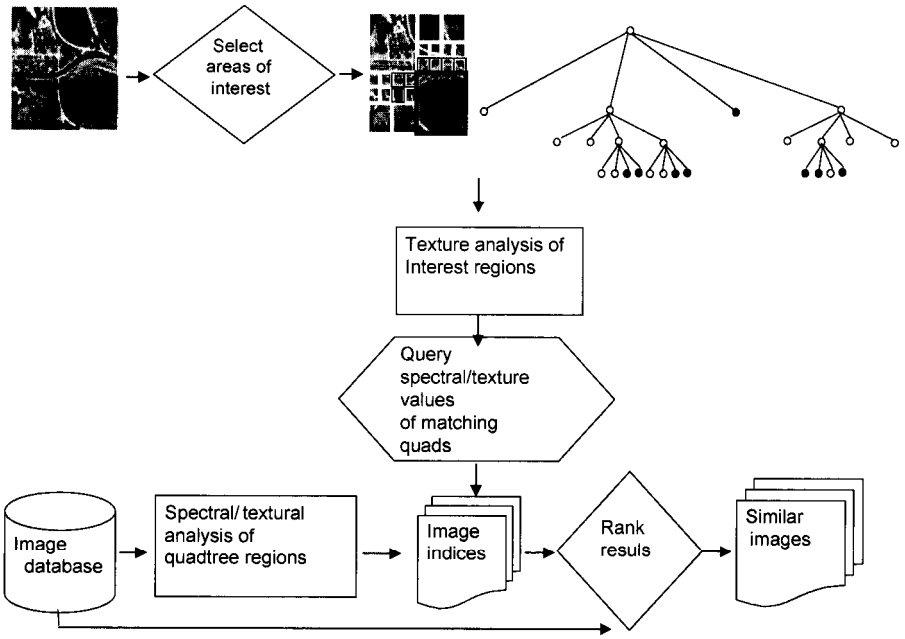

Fig. 3. Content-based image retrieval process.

deviation of the grayscale values for each quad were also computed and stored in the metadata table.

Similarity Search

ICAMS includes a similarity search capability in which the user defines a feature of interest, the application computes the spatial and spectral characteristics of this feature, and then searches a metadata table for similar characteristics that indicate the presence of matching features. Figure 3 shows the procedure used in this example. First, a separate application creates the metadata table of spatial and grayscale statistics of 512 × 512 pixel tiles derived from a larger image scene. Fractal dimension computed by the box counting and triangular prism methods, Moran's *I*, lacunarity, Daubechies' DB4 wavelet energy, and mean, standard deviation, and a normalized five-bin histogram of the grayscale values of each of these 512 × 512 pixel tiles are analyzed in a recursive region quadtree structure (Samet, 1984). The results from progressively smaller square quadtree regions (quads) are stored in a relational database of tile indices. Each 512 × 512 pixel tile in the image index database has one record with the spatial and grayscale statistics for the whole tile, four records representing the first level of quadtree decomposition, 16 records for the second level, 64 records for the third level, 256 records for the fourth level, and 1096 records for the fifth level. Five levels were chosen as the highest level of decomposition, because the resulting quads have a dimension of 16 × 16 pixels, and this was thought to be the minimum support necessary for the fractal dimension calculations.

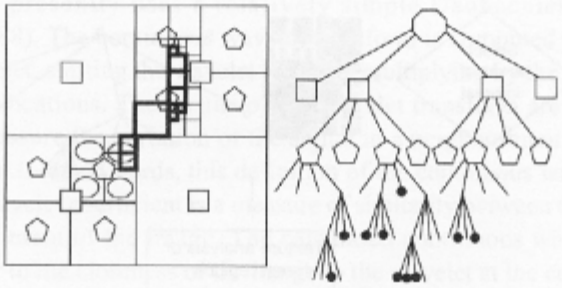

Fig. 4. Region quadtree map and tree structure for user-defined feature.

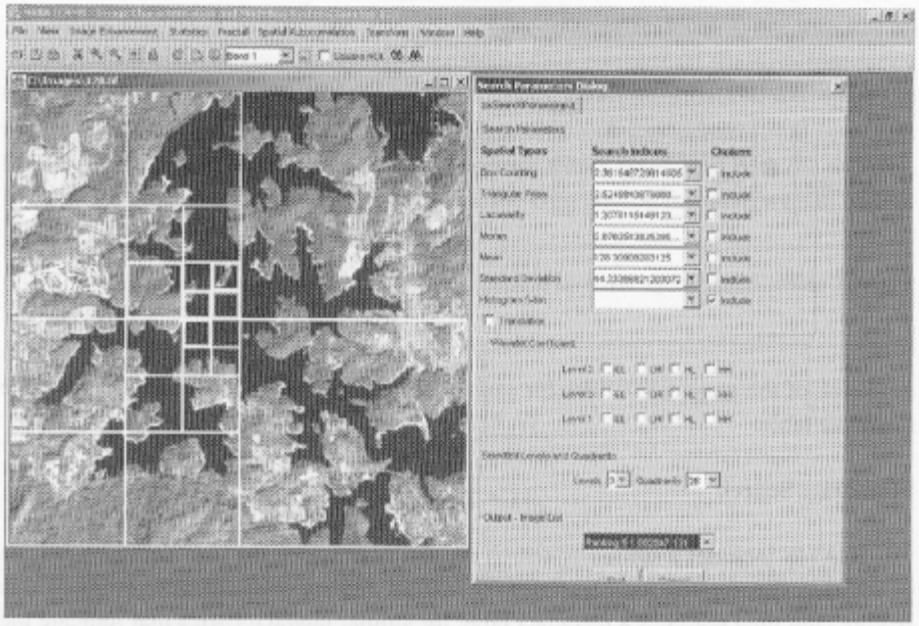

Fig. 5. Region quadtree feature selection in ICAMS.

Figure 4 shows how quads that correspond to a feature outline are indexed within the region quadtree structure. The user clicks the left mouse button on the tile to break it down into progressively smaller segments. If the user's concept of a feature of interest involves a characteristic size and shape, then the outline of the feature is approximated by selecting the appropriate quads. Figure 5 shows an example feature of interest that has been defined in this fashion, with the quads corresponding to the feature selected by a right-click (indicated by thicker lines). The feature definition is constructed by computing the selected spatial or gray statistics for the ensemble of quads that define the feature, and this set of indices is used to find matching features.

In static feature matching, the simplest form of image retrieval, the tile index database is queried for the same quads in the same orientation and location as the user-defined definition of the feature of interest. Corresponding quads in each of the

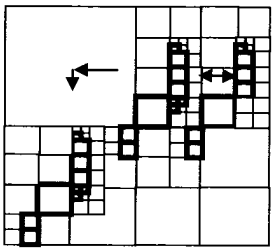

Fig. 6. Translation of user-defined feature.

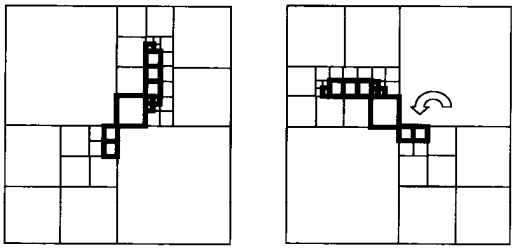

Fig. 7. Rotation of user-defined feature.

database tiles are compared to the feature definition, and the results are ranked using a least-squares difference calculation. Future versions of ICAMS will include other distance measures such as Mahalanobis distance, if they can be shown to improve retrieval performance. Static feature matching is only appropriate for databases of images such as portraits that have similar feature sizes and orientations.

In the context of satellite remote sensing, content-based image retrieval poses challenges, such as translation of the x,y location of matching features in other images, rotation of the feature, scaling up or down in size, isolation of the feature from other patterns and textures, and 3-D orientation of the feature with respect to the observer (termed "object pose" in the pattern recognition literature). The region quadtree structure was used in this application because it allows at least a partial ability to minimize some of these problems. In the current implementation of ICAMS, translation is taken into account by moving the user-defined set of quads around the entire range of possible locations of the largest quad in the feature definition (Fig. 6). As compared to static searches, translation increases the search time, especially if the feature definition does not include any large quads. 90 degree rotations and reflections can be calculated within the quadtree feature definition by simply changing the numbering of the nodes, although other types of rotations pose a difficult challenge (Fig. 7).

Scaling at integer multiples defined by the quadtree structure can be incrementally performed by jumping up or down within the quadtree (Fig. 8). However, lacunarity, fractal dimension calculations, and other spatial indices may change with the size of the support, or number of pixels used in the calculation. For vertical aerial and satellite imagery, the object pose problem is generally not severe, although adjustments have to be considered for oblique images and features significantly

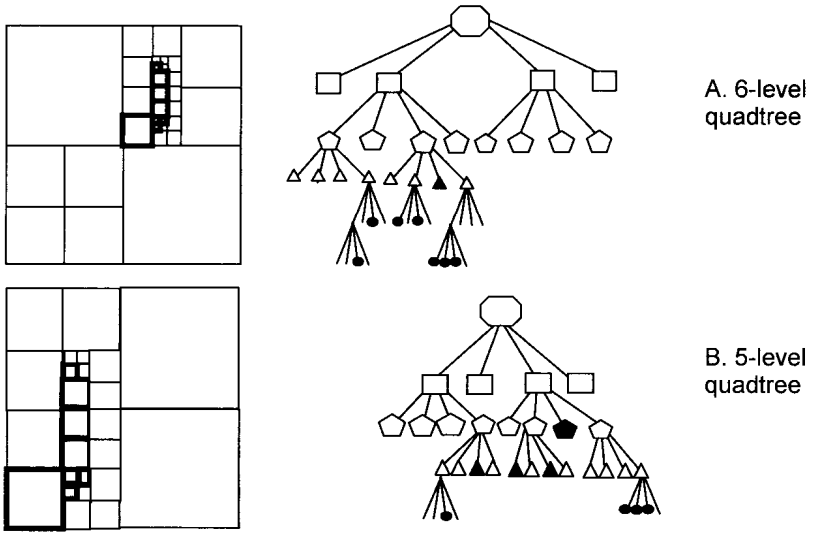

Fig. 8. Scaling up user-defined feature according to quadtree hierarchy.

off-nadir. Rotation and scaling have not yet been incorporated into the ICAMS content-based image retrieval utility.

RESULTS

Landsat 7 Image

A Landsat 7 ETM+ panchromatic image obtained on October 24, 2000 was subsetted into 506 512 × 512 pixel tiles to form an example tile database. The Landsat scene (Path 19, Row 36) includes the northern half of the Atlanta, Georgia metropolitan area, plus the southern end of the Appalachian Mountains in northern Georgia, Alabama, North Carolina, and Tennessee (Fig. 9). The scene includes a range of land covers, including extensive areas of forested mountains and the northern suburbs of Atlanta, some large lakes, and smaller areas of pasture lands.

In this example, two different features were identified: (1) a collection of quads that define a lake shore or river bank with a land/water contrast, and (2) an extensive residential area. A quantitative determination of the success or failure of a content-based image retrieval process necessarily involves a degree of subjectivity, since success equates to returning a tile that contains one or more features that match the semantic definition (as in land/water contrast) that the analyst had in mind when they were defining the feature. In this case, the database of 506 subsetted tiles was visually inspected for land/water contrast areas, and it was found that 62 of the 506 tiles contained one or more matching features.

Figure 10 shows the original query tile of an area with a characteristic land/water contrast and the feature definition for a five-bin grayscale histogram analysis. The tiles with the top five closest matches to the query tile are contained in Figures 10B–10F. The five-bin histogram returned matches in all five of the top ranked land/water

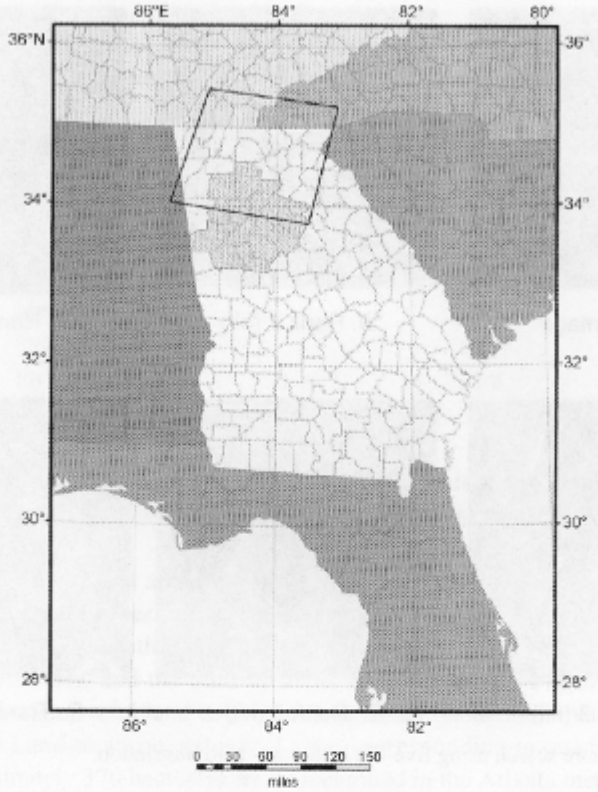

Fig. 9. Footprint of Landsat image over the Atlanta metropolitan area.

Table 1. Comparison of Indices for the Top Five Ranked Land/Water Contrasts

Index	Matches
Box counting	1/5
Triangular prism	1/5
Lacunarity	1/5
Moran's *I*	1/5
Mean	2/5
Std. deviation	2/5
5-bin histogram	5/5

contrast tiles. Table 1 shows that for this type of feature, the five-bin histogram outperformed the other indexes at image retrieval. A Mann-Whitney test was performed on the ranks of all remaining 505 tiles in the database. Matches from the visual inspection of the database were coded as a binary grouping variable (0 = no land/water, 1 = land/water interface). Tables 2 and 3 show that the results of this analysis

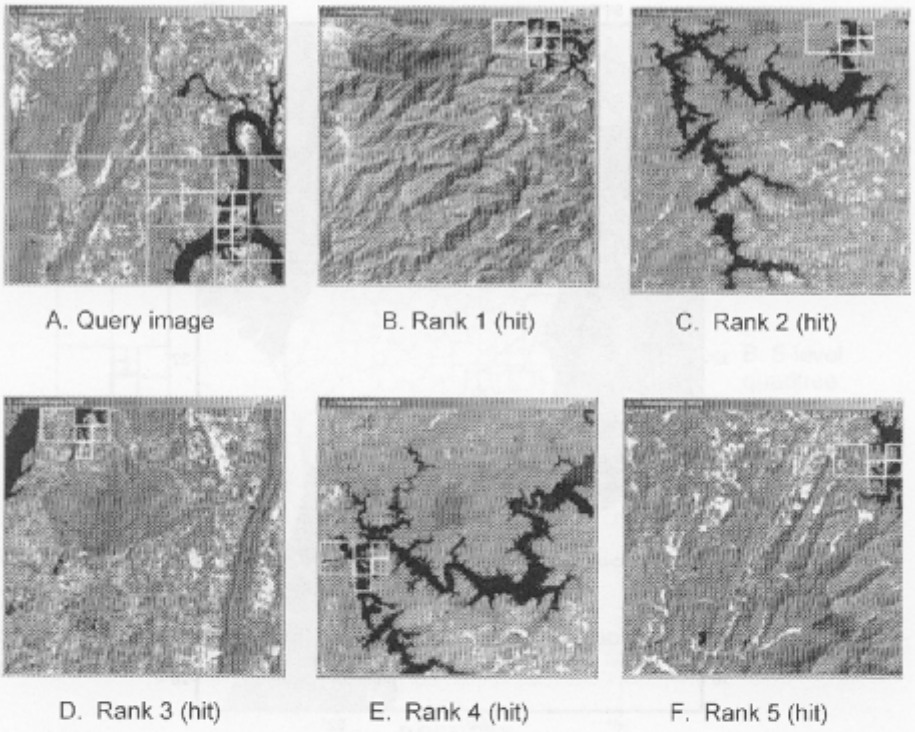

Fig. 10. Lake shore search using five-bin histogram with translation.

Table 2. Ranks of Five-Bin Histogram Analysis of Land/Water Contrast

Group	*N*	Mean rank	Sum of ranks
No land/water	444	273.39	121,385.00
Land/water	61	104.59	6,380.00
Total	505		

Table 3. Mann-Whitney Test Statistics of Land/Water Analysis

Statistic	Value
Mann-Whitney U	4489.00
Wilcoxon W	6380.00
Z	−8.471
Asymp. sig. (2-tailed)	0.000

Table 4. Comparison of Indices for the Top Five RankedResidential Areas

Index	Matches
Box counting	3/5
Triangular Prism	2/5
Lacunarity	4/5
Moran's *I*	0/5
Mean	0/5
Std. deviation	0/5
Five-bin histogram	2/5

yield Mann-Whitney U, Wilcoxon W, and *z*-score statistics that show the overall rankings from the five-bin histogram were not likely to resemble those derived from a chance arrangement.

Images of urban areas are often more complex than natural landscape images, with a number of land covers closely interspersed. Residential areas are particularly difficult to pick out from other land covers, because lawns and street trees resemble grasslands and forests, and the built up areas are often difficult to distinguish from more intensive commercial land uses. Although small residential areas are dispersed throughout the Landsat scene, extensive areas corresponding to quads of 128 × 128 pixels (approximately 370 hectares) are concentrated in the Atlanta metropolitan area. From visual inspection, 69 of the 506 tiles had extensive residential areas. The heterogeneity of urban land covers makes summary statistics of grayscale values of relatively little use, since a diverse range of land cover types may have similar grayscale distributions. Table 4 shows that lacunarity yielded the highest number of residential matches in the top five ranked tiles, followed by box counting and triangular prism fractal dimension. Figure 11 shows that four of the top five tiles in the lacunarity analysis matched extensive residential areas. The mismatch occurred in a mountainous area with small water bodies, cleared areas, shadows, and other features that form a complex land cover that has a texture similar to an urban area. The retrieval efficacies of lacunarity and box counting fractal dimension were also compared using the Mann-Whitney test. Tables 5 and 6 show that both methods were significantly different from random rankings, with the box counting fractal dimension having a slightly lower mean rank of residential area matches.

IKONOS Image

An IKONOS panchromatic image obtained on October 24, 2002 was subsetted into 380 512 × 512 pixel tiles to form an example tile database. The image covers the western portion of Kalamazoo, Michigan, including the Western Michigan University campus. The scene includes a range of land covers such as residential and commercial areas, golf courses, forested wetlands, and small lakes. Analysis proceeded as before with the translation capability enabled.

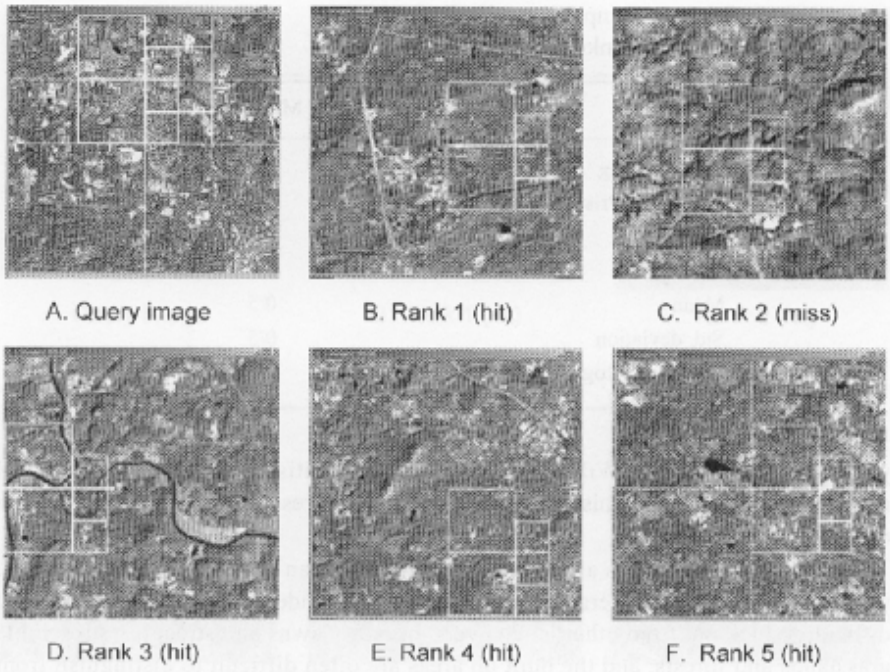

Fig. 11. Residential area search using lacunarity.

Table 5. Ranks of Box Counting Fractal Dimension and Lacunarity Analysis of Residential Areas

Method	Grouping	*N*	Mean rank	Sum of ranks
Lacunarity	No residential	438	278.10	121,807.00
	Residential	67	88.93	5,958.00
	Total	505		
Box counting	No residential	438	278.26	121,878.99
	Residential	67	87.85	5,886.00
	Total	505		

Nine of the 380 tiles contained all or part of a golf course. Figure 12 shows the selected quads that form the feature definition and the top five tiles retrieved from the database. As compared to the rest of the image, the most distinguishable property of the golf course area is the lighter tone of the grass and sand traps as compared to the surrounding trees and urban land covers. It is therefore not surprising that the indices that achieved the best results were the five-bin histogram and the mean/standard deviation of the grayscale values. Table 7 summarizes the results of the golf course image retrieval process.

Table 6. Mann-Whitney Test Statistics of Residential Area Analysis

Statistic	Lacunarity	Box counting
Mann-Whitney U	3680.00	3608.00
Wilcoxon W	5958.00	5886.00
Z	−9.882	−9.947
Asymp. sig. (2-tailed)	0.000	0.000

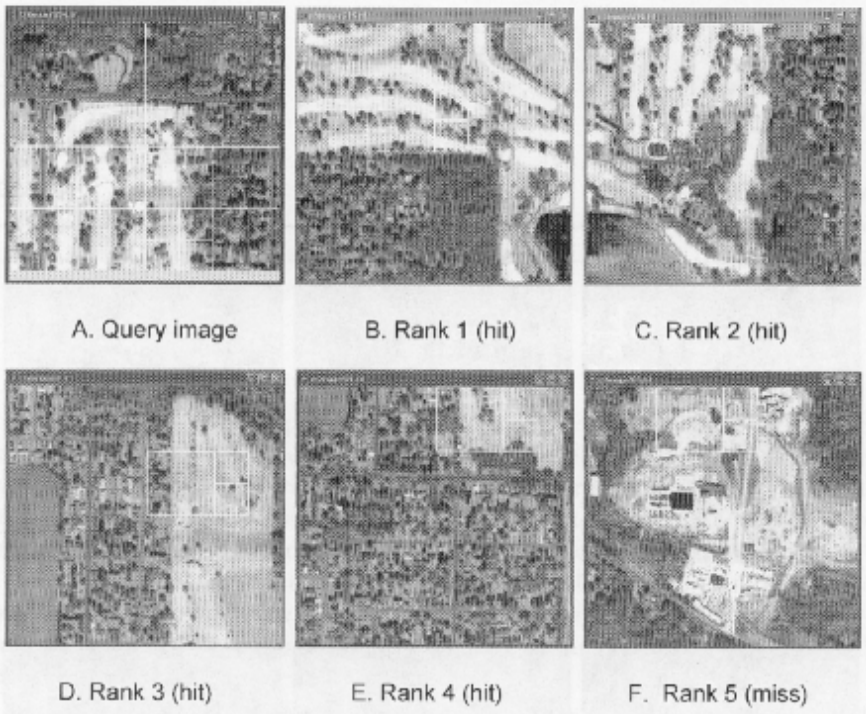

Fig. 12. Ranked results from five-bin histogram analysis of a golf course.

In the IKONOS source image, trailer parks were a unique, easily identifiable land cover, characterized by closely interspersed rectangular structures, usually located at an angle. Ten of the 380 tiles contained a trailer park. Possibly due to the oblique orientation of most of the trailer parks in the IKONOS image, the diagonal high-pass wavelet energy signatures yielded the best results, followed by box counting fractal dimension. Other measures such as triangular prism fractal dimension and Moran's *I* yielded poor results and are not summarized. Figure 13A shows the feature definition, with the top five ranked tiles (Figures 13B–13F) that were obtained using the high-pass wavelet energy signatures. Although it was possible to retrieve all 10 of the tiles in the database that depicted large trailer parks, other urban features, such as the

Table 7. Results of Golf Course Image Retrieval

Grayscale 5-bin Hist.		Grayscale Mean/St. Dev		Fractal dim. Box counting		Fractal dim. Tri. Prism		Spatial autoc. Moran's *I*		Wavelet High pass energy	
Image	Rank	Image	Rank	Image	Rank	Image	Rank	Image	Rank	Image	Rank
295	1	128	1	147	5	128	4	31	1	31	1
296	2	295	2	295	7	147	6	275	2	295	2
31	3	127	3	108	9	275	14	128	4	128	4
51	4	296	10	31	11	31	22	295	6	296	7
128	6	275	11	128	14	127	25	296	7	108	10
108	7	108	12	275	35	296	69	127	8	127	14
127	14	31	16	296	62	295	77	108	10	51	15
275	24	51	17	127	79	108	110	147	31	275	28
147	39	147	35	51	30	51	201	51	122	147	47
Mean	11.11		11.89		28.00		58.67		21.22		14.22

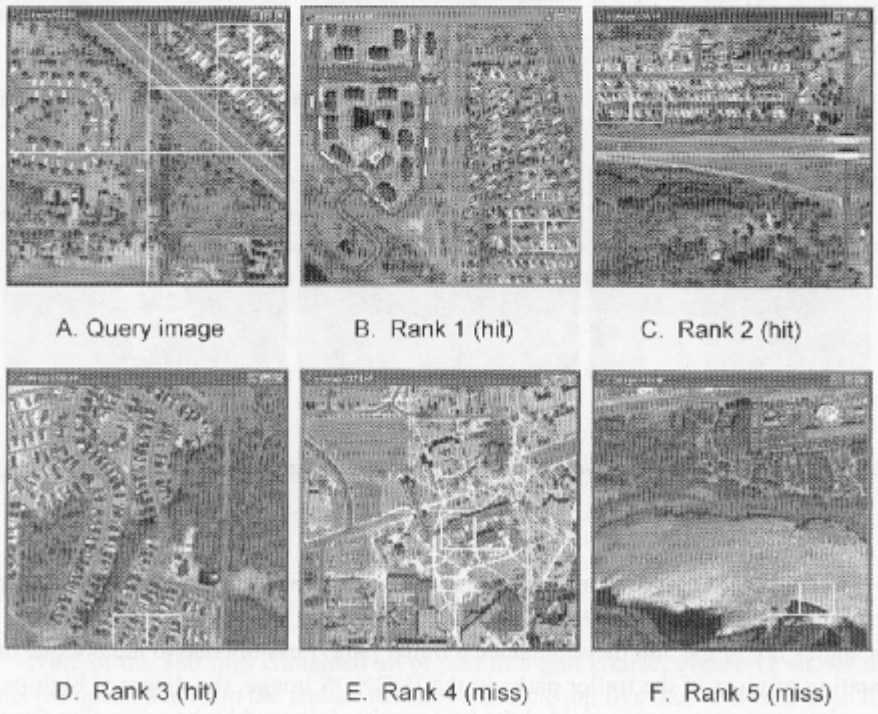

Fig. 13. Ranked results from wavelet analysis of a trailer park.

college campus tile (Fig. 13E) and rough textures such as the sun glint on the lake (Figure 13F), interfered with the positive results. Table 8 shows the results of the trailer park image retrieval process.

Table 8. Results of Trailer Park Image Retrieval

Grayscale 5-bin. Hist		Fractal dim. Box counting		Lacunarity		Wavelet High-pass energy	
Image	Rank	Image	Rank	Image	Rank	Image	Rank
14	1	365	4	364	1	183	1
183	3	183	5	365	9	167	8
36	10	211	11	212	10	14	9
211	13	184	17	184	31	36	11
265	19	36	18	167	33	211	13
229	23	265	22	211	56	365	16
184	39	14	27	203	77	184	18
365	75	167	30	36	97	210	24
167	78	210	42	210	110	212	43
210	111	366	62	366	123	229	75
Mean	37.2		23.8		54.7		21.8

CONCLUSIONS

The quadtree structure provides an efficient means of indexing and searching a database of low-level image characteristics. Some of the indices that have been examined in this work, most particularly fractal dimension and lacunarity, are greatly affected by the statistical support (size) of the selected quads, and although this encourages the use of larger quads with more stable indices, it simultaneously leads to a more approximate definition of the complex shape of a defined feature. Selecting only smaller quads to define a feature greatly increases search times (particularly if the translation capability is enabled), because a much larger number of records in the index database must be examined.

We have found that the hierarchical quadtree structure provides a good basis for a user to isolate the feature of interest from the background if the user selects homogeneous regions that characterize the feature itself. However, it is often useful to select surrounding regions that are not part of the defined feature, if the figure/ground contrast is important to determining the success or failure of a match. Size and shape characteristics may or may not be important to a feature definition, and the degree of success at retrieving matching features depends in part on the user's skill at identifying the salient characteristics of the feature of interest.

Although the basic framework for performing content-based image retrieval is presently encoded into ICAMS-Java, there still remains much to accomplish before the tool can be considered completely successful. Only single-band images can be analyzed at this point, and this limits the accuracy of the retrieval. Adding multispectral capability should significantly improve performance, because many of the key identifying characteristics of features are related to color (Yao and Chen, 2003), as represented by the information contained in multispectral bands.

The region quadtree approach provides a framework for evaluating the performance of potentially any combination of textural or spectral measures for content-based image retrieval. Pre-processing a large image scene or a collection of scenes and storing the spatial and spectral indices in a database can potentially extend the definition of metadata beyond the usual descriptions of acquisition date, sensor identification, and lineage to a richer content-based description that can facilitate access to imagery that depicts specific features and conditions on the Earth's surface that may be of interest to researchers engaged in earth science, resource development, planning, and security investigations.

ACKNOWLEDGMENTS

This research is supported by a research grant from NASA Intelligent Systems (research grant NAS-2-37143).

REFERENCES

Allain, C. and M. Cloitre, 1991, "Characterizing the Lacunarity of Random and Deterministic Fractal Sets," *Physical Review A,* 44:3552–3558.

Baatz, M. and A. Schape, 2000, "Multiresolution Segmentation—an Optimization Approach for High Quality Multi-scale Image Segmentation," in *Angwandte Geographische Informationsverarbeitung XII*, Strobl, J., Blaschke, T., and G. Griesebner (Eds.), Heidelberg, Germany: Wichmann-Verlag, 12–23.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., and P. Zeil, 2000, "Object-Oriented Image Processing in an Integrated GIS/Remote Sensing Environment and Perspectives for environmental applications," in *Environmental Information for Planning, Politics and the Public*, vol. 2, Cremers, A. and K. Greve (Eds.), Marburg, Germany: Metropolis Verlag, 555–570.

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

Cliff, A. and J. Ord, 1973, *Spatial Autocorrelation*, London, UK: Pion Limited.

Datcu, M., Daschiel, H., Pelizzari, A., Quartulli, M., Galoppo, A., Colapicchioni, A., Pastori, M., Seidel, K., Marchetti, P., and S. D'Elia, 2003, "Information Mining in Remote Sensing Image Archives: System Concepts," *IEEE Transactions on Geoscience and Remote Sensing*, 41(12):2923–2936.

Daubechies, I., 1988, "Orthonormal Bases of Compactly Supported Wavelets," *Communications in Pure and Applied Mathematics*, 41:909–996.

Daubechies, I., 1991, *Ten Lectures on Wavelets,* Philadelphia, PA: SIAM, CBMS-NSF Series in Applied Mathematics.

Dong, P., 2000, "Test of a New Lacunarity Estimation Method for Image Texture Analysis," *International Journal of Remote Sensing*, 21(17):3369–3373.

Getis, A. and J. Ord, 1992, "The Analysis of Spatial Association by the Use of Distance Statistics," *Geographical Analysis*, 24:189–206.

Jaggi, S., Quattrochi, D., and N. Lam, 1993, "Implementation and Operation of Three Fractal Measurement Algorithms for Analysis of Remote Sensing Data," *Computers and Geosciences*, 19:745–767.

Lam, N., Qiu, H., Quattrochi, D., and C. Emerson, 2002, "An Evaluation of Fractal Methods for Characterizing Image Complexity," *Cartography and Geographic Information Science*, 29:25–35.

Li, J. and R. Narayanan, 2004, "Integrated Spectral and Spatial Information Mining in Remote Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 42(3):673–685.

Liebovitch, L. and T. Toth, 1989, "A Fast Algorithm to Determine Fractal Dimensions by Box Counting," *Physics Letters A*, 141(8,9):386–390.

Mallat, S. G., 1989, "A Theory for Multi-resolution signal decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11: 674–693.

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

Manjunath, B. and Y. Ma, 1996, "Texture Features for Browsing and Retrieval of Image Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842.

Ohm, J. R., Bunjamin, F., Liebsch, W., Makai, B., Muller, K., Smolic, A., and D. Zier, 2000, "A Set of Visual Feature Descriptors and Their Combination in a Low-Level Description Scheme," *Signal Processing: Image Communication*, 16:157–179.

Paquet, E., Rioux, M., Murching, A., Naveen, T., and A. Tabatabai, 2000, "Description of Shape Information for 2-D and 3-D Objects," *Signal Processing: Image Communication*, 16:103–122.

Plotnick, R., Gardner, R., Hargrove, W., Prestegaard, K., and M. Perlmutter, 1996, "Lacunarity Analysis: A General Technique for the Analysis of Spatial Patterns," *Physical Review E*, 53(5):5461–5468.

Quattrochi, D., Lam, N., Qiu, H., and W. Zhao, 1997, "Image Characterization and Modeling Systems (ICAMS): A Geographic Information System for the Characterization and Modeling of Multiscale Remote Sensing Data," in *Scaling in Remote Sensing and GIS*, Quattrochi, D. and M. Goodchild (Eds.), Boca Raton, FL: CRC/Lewis Publishers, 295–307.

Samet, H., 1984, "The Quadtree and Related Hierarchical Data Structures," *Computing Surveys*, 16(2):187–260.

Sarkar, N. and B. Chaudhuri, "An Efficient Approach to Estimate Fractal Dimension of Textural Images," *Pattern Recognition*, 25(9):1035–1041.

Smeulders, A., Worring, M., Santini, S., Gupta, A., and R. Jain, 2000, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380.

Yao, C. and S. Chen, 2003, "Retrieval of Translated, Rotated and Scaled Color Textures," *Pattern Recognition*, 36:913–929.