Effectiveness of Wavelet Transform, Fractals, Lacunarity and Spatial Indices in Content-Based Image Retrieval

Sivagurunathan Chinniah

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Geography Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/3921
EFFECTIVENESS OF WAVELET TRANSFORM, FRACTALS, LACUNARITY AND SPATIAL INDICES IN CONTENT-BASED IMAGE RETRIEVAL

by

Sivagurunathan Chinniah

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master in Arts
Department of Geography

Western Michigan University
Kalamazoo, Michigan
April 2005
ACKNOWLEDGMENTS

Standing at the near end of my GIS journey, I would like to look back and thank all the terrific people who walked hand in hand with me, helping and guiding along the way, to overcome obstacles and cheer challenges. Besides being rough and stormy now and then, these past two years have been really wonderful.

Firstly, my heartfelt thanks to Dr Charles (Jay) Emerson, the man of many ideas, for constantly encouraging and supporting me in every step of the way, bringing out creativity in me by being fantastically flexible.

Secondly, I would like to take this opportunity to convey my countless thanks to Dr David Dickason for creating a world-class environment for conducing learning and research, and always lending out a helping hand for those in need. He is great indeed.

Also, my many thanks goes to Dr James Biles for providing reference materials and reviewing this thesis.

A special thanks to NASA for funding this research through the Intelligent Systems program at NASA Ames Research Center. The title of the grant was: "Fractal and Geostatistical Metadata for Monitoring Global Change Using Remote Sensing Imagery NCC-2-1246."
Acknowledgments—continued

Last but not least, my loving thanks to my fabulous family (wife Nalina, daughter Neevetha Malar and son Shuvethan) for backing me up and putting up with me. What will I do without them?

Sivagurunathan Chinniah
The thesis demonstrates the feasibility of applying wavelet transforms, fractals, and spatial indices in content-based image retrieval (CBIR) of remotely sensed imagery. It consists of three major parts: the design and development of thick-client CBIR software, population of two samples of SPOT™ and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) image databases with metadata describing their visual properties, and, exploration of the effectiveness of region quadtree query and retrieval by content using various combinations of indices. 272 SPOT™ and 507 Landsat 7 ETM+ 512 x 512 pixel panchromatic images of Atlanta, Georgia were subsetted from two scenes. Moran’s I index of spatial autocorrelation, two sets of fractal dimensions (based on the triangular prism and box counting methods), lacunarity, 5-bin spectral histogram and wavelet energy signatures (12 indices at 3 levels) were computed at each step of the six level region quadtree and the results were stored in a relational database. To query the populated database, region quadtree was used to select features of interest within a randomly selected image. Associated indices of the selected quadrants of the query region were computed and compared to the metadata in the database. Individual applications of each index on the whole image was found to reveal a promising pattern but picking the right indices and combining them with equal weights seem to pose a challenge especially when smaller quads are used.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ...................................................................................... ii

LIST OF TABLES ................................................................................................ vi

LIST OF FIGURES ........................................................................................... vii

CHAPTER

1. INTRODUCTION ................................................................................. 1

2. BACKGROUND .................................................................................... 3

3. DATA AND METHODOLOGY ........................................................... 5
   3.1. Data .......................................................................................... 5
   3.2 Region Quadtree ........................................................................ 6
   3.3 Metadata .................................................................................. 13
      3.3.1. Daubechies’ Wavelet Transform ....................................... 13
      3.3.2. Spatial Autocorrelation and Moran’s I......................... 22
      3.3.3. Fractal Dimension .......................................................... 25
      3.3.4. Lacunarity ...................................................................... 27
   3.4. Content-based Image Retrieval Software Overview ............ 29
   3.5. Methodology .......................................................................... 32

4. RESULTS .......................................................................................... 35
   4.1. Retrieval Results from SPOT Database ................................... 35
   4.2. Retrieval Results from Landsat Database .......................... 41
      4.2.1. Wavelet Energy Signature ............................................. 41
      4.2.2. Moran’s I Results .......................................................... 51
Table of Contents—continued

4.2.3. Fractal Dimension Using Box Counting Method
Results.......................................................................................... 54

4.2.4. Lacunarity Results.............................................................. 57

4.2.5. Results – Combination of Various Indices. ....................... 60

5. CONCLUSION............................................................................................. 65

REFERENCES ................................................................................................... 66
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Quadtree Levels, Number of Records and the Quadrant Size.</td>
<td>9</td>
</tr>
<tr>
<td>2.</td>
<td>Wavelet Energy Signature of Figure 13.</td>
<td>22</td>
</tr>
<tr>
<td>3.</td>
<td>Fractal Dimensions (Box Counting Method) Statistics for Landsat Database.</td>
<td>26</td>
</tr>
<tr>
<td>4.</td>
<td>Lacunarity Statistics.</td>
<td>28</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Region Quadtree Decomposition of 512 x 512 Image ................................... 7
2. Hierarchical Representation of Quadtree ....................................................... 8
3. Selecting a River Bend Using Region Quadtree ........................................... 9
4. Third Level Region Quadtree Quadrant Labeling .......................................... 10
5. Approximate Outline of a Lake Using Region Quadtree. Selected Quadrants are Marked in Bold and Labelled from 1 to 7. ......................... 11
6. Logical Representation of the Selected Quadrants from Figure 5 ................. 12
7. Notation for Level 2 Standard Wavelet Decomposition. ................................. 15
8. 2D Forward Wavelet Transform Using Daubechies’ db4 .............................. 16
9. Sample 16x16 Image of a River Bank for Illustrating Daubechies’ db4 ........... 17
10. Actual 16 x 16 Grayscale Values of a River Bank in Figure 9(a) .................... 17
11. Daubechies’ db4 Smoothing and Non-Smoothing Filters. Smoothing and Differencing Filters are Also Known as Lowpass and Highpass Filters Respectively ........................................................................................................ 18
12. The Effect of Level 1 Wavelet Transform Decomposition on a Real 256x256 Image ........................................................................................................ 20
13. Computed Daubechies’ db4 Wavelets Coefficients. Figure 7(b) is the Result of the Stretching These Coefficients and Display it as an Image ........................................................................................................ 20
14. Moran’s / Indices on Simulated Images ......................................................... 23
15. Moran’s / Indices Computed in the Marked Region on Actual Images That Closely Matches the Theoretical Values of Figure 16 from the Landsat Database .................................................................................................... 24
16. Pixel Dump of Marked Region of Figure 15(a) ............................................. 24
17. Pixel Dump of Marked Region of Figure 15(b) ............................................. 25
List of Figures—continued

18. The Fractal Dimension of the Two Squared Region is 2.41 but Their Lacunarity is Very Different. ................................................................. 27

19. Snapshot of the Spatial Index Table Containing 371280 Records at 6 Different Levels of the Quadtree Representation for the 272-Segmented Imagery. .......................................................................................................................... 29

20. Snapshot of the Spatial Index Table Containing 690690 Records at 6 Different Levels of the Quadtree Representation for the 507-Segmented Imagery Consisting of Wavelets Energy Signature........................................ 30

21. CBIR Architecture. .................................................................................. 31

22. Search Parameters Dialog Box. ................................................................ 33

23. Similarity Search Based on Moran’s I and Fractal Dimension Using Box Counting Method on a Road Feature Without Translation. ................................................................. 37

24. Similarity Search Based on Moran’s I and Fractal Dimension Using Box Counting Method on a Road Feature with Translation. ................................................................. 38

25. Similarity Search Based on Mean and a Level 3 Selected Quadrant. .......... 39

26. Similarity Search Based on Moran’s I and Fractal Dimension Using Triangular Prism Method on a Portion of a River Bend. ................................................................. 40

27. Wavelet Energy Signature of LL, LH, HL and HH at Different Levels. ...... 42

28. Wavelet Energy (LL1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order. .................................................................................. 43

29. Wavelet Energy (LL1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order. .................................................................................. 44

30. Wavelet Energy (LH1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order. .................................................................................. 45

31. Wavelet Energy (LH1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order. .................................................................................. 46

32. Wavelet Energy (HL1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order. .................................................................................. 47
List of Figures—continued

33. Wavelet Energy (HL1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order. ................................................................. 48
34. Wavelet Energy (HH1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order. ................................................................. 49
35. Wavelet Energy (HH1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order. ................................................................. 50
36. Moran’s I Indices on Real Images at Level 0........................................................................ 51
37. Moran’s I at Level 0 to 3.................................................................................................. 52
38. Level 1 Moran’s I – Top Six Listing from Landsat Database in Ascending Order. .................................................................................................................................. 53
39. Level 1 Moran’s I – Bottom Six Listing from Landsat Database in Ascending Order. .................................................................................................................................. 54
40. Fractal Dimension at Level 0 to 4 .................................................................................. 55
41. Level 1 Fractal Dimension d (Box Counting Method) – Top Six Listing from Landsat Database in Ascending Order. .................................................................................. 55
42. Level 1 Fractal Dimension d (Box Counting Method) – Bottom Six Listing from Landsat Database in Ascending Order. .................................................................................. 56
43. Lacunarity at Level 0 to 4. ............................................................................................. 57
44. Top Three Biggest Lacunarity in Landsat Database. .................................................. 58
45. Lacunarity – Top Six Listing from Landsat Database in Ascending Order. ................................................................................................. 58
46. Lacunarity – Bottom Six Listing from Landsat Database. .......................................... 59
47. Box Counting, Lacunarity and Moran’s I Without Translation. .................................... 60
48. Wavelet Coefficient HH1, HH2, HH3 Without Translation. .......................................... 61
49. Lacunarity with Translation. ......................................................................................... 62
50. Lacunarity with Translation. ......................................................................................... 63
List of Figures—continued

51. Top 11 Retrieval Based on LH1 Applied on All the 4 Quadrants... ............... 64
Satellites capture huge numbers of images, demanding a system for efficient and effective image retrieval. Even though these images may be taken for different purposes, the inherent properties within the images may contain a wealth of useful information, which cannot be accessed if the images are not properly organized and stored for future retrieval. Basically, there are two ways to achieve this goal: text-based and content-based image retrievals. The first method, which uses text or keywords to describe and retrieve images, has been an active area of research since 1980s (Tamura et. al, 1984). One good example of such a system is the Google Internet search engine. After achieving great success in text retrieval, Google recently expanded its tool to retrieve images found on the Internet. On first impression, it appears to work well but careful visual inspection of the retrieved images will clearly reveal that the retrieval success solely depends on the validity of the text or keyword description associated with the image. The accuracy of such system is highly dependent on how well one perceives and interprets the image and indexes a precise set of descriptive keywords. Since it is not possible to specify all aspects of an image, a user who uses entirely different keywords in the search parameters will be unable to access it. This problem is exacerbated when the image contain several distinct objects. Furthermore, the process of identifying precise keywords associated with an image is laborious and time consuming. These issues prompt a need for alternative method and this is what content-based image retrieval is all about.

Content-based image retrieval (CBIR) is the querying of an image database to
find similar images based on one or a combination of visual characteristics such as texture, color or shapes of objects. It is not based on text or keyword description but rather on visual image attributes. Query by example and query by sketch are two common ways of retrieving the image of interest from an image database (Equitz et. al, 1994, Scassellati et. al, 1994). In this thesis, the former method was used. Usually, this method requires a tedious technique of image classification and segmentation, but in this research this procedure was avoided by employing the region quadtree (Chapter 3.2).

The rest of the thesis is organized as follows: Chapter 2 gives the background information on content-based image retrieval and the reason for conducting the research. Chapter 3 begins by talking about the data and methodology used. It also discusses the usage of region quadtree and the wavelet transform multi resolution analytical technique. After that, an overview of the CBIR architecture is provided, and, the section ends with the explanation of Moran’s I, fractal dimension and lacunarity. Chapter 4 and 5 presents retrieval results, concluding remarks and recommendations for future work.
CHAPTER 2

BACKGROUND

Humans invented the computer to improve their communication system and in the process not only unwittingly made the world smaller but also introduced a new challenging problem. Since Internet technologies provided businesses and individuals with not only cheaper but more effective forms of communication, large amount of data were generated and stored. This prompted a need to organize this data, and many engines such as Google, Yahoo and MSN began to emerge, producing innovative solutions. In the beginning, the data was comprised of text but as technology began to mature sound, graphics, images and video were included. As the speed of transmission increased the use of non-text media increased in frequency and today it has become the norm for such media to be transmitted and stored.

For thousands of years, since the first writings, humans have learned to organize textual data and it was rather easy for developers of search engines to adapt some of these ideas to the modern mode of electronic communication. But organizing digital multimedia information poses a very difficult problem. Multimedia includes animation, sound, graphics, video and text but in this research the focus is on images, particularly remote sensing imagery.

Keyword search has proven to be inadequate due to retrieval mismatches. It is common for documenters and retrievers to have a different interpretation of the same image that causes this problem. So researchers began to look for innate properties such as color, texture, structure and spatial relationships within the image. In content-
based image retrieval user-defined features from an input image are used to query an image database as opposed to the traditional method of querying based on textual information.

A number of frameworks for CBIR have been proposed (Niblack et al. 1994, Gupta et al. 1997, Dowe 1993). There are many CBIR prototypes but they have addressed only low-level image properties and not specific characteristics of the earth's surface (Agouris et al. 1999). Low-level features include color, texture and statistical characteristics of the image. High-level features relate to specific visual inspection application such as identification of fingerprint and human faces, is based on image semantics. There are many CBIR systems today (Tieu and Viola, 2003) and some examples include QBIC (Niblack et al., 1992), Virage (Gupta et al., 1997), VisualSEEK (Smith & Chang, 1996a), and Photobook (Pentland, 1994) but most of the systems do not address the unique problems associated with remotely sensed earth images.

The marriage of information technology with geography gave birth to a new, dynamic and challenging field – Geographical Information Systems (GIS). Just like non-spatial data, acquiring and analyzing geographical data used to be a tedious and laborious task. Today, with the availability of complex GOS spatial analysis tools and the abundance of remotely sensed imagery, these issues have become much easier. However, GIS and remote sensing technologies inherited the same problem encountered in the world of computing resulting in the proliferation of spatial data including imagery. There is an urgent need to organize and store these images into an image database according to some unique characteristics to make it easy for a user to search and retrieve them using a set of input parameters. In this research, I describe a content-based image retrieval system specifically for remote sensing images.
CHAPTER 3

DATA AND METHODOLOGY

3.1. Data

The choice of imagery for this paper was based on ease of availability and convenience. I used panchromatic SPOT™ satellite imagery of the city of Atlanta, Georgia with a spatial resolution of 10 meter obtained in 1995. In this preliminary study, the original image was divided into 272 512x512 non-overlapping images. This size was convenient for the pilot-scale study.

To verify the results found in the preliminary study, a larger database consisting of 507 Landsat 7 ETM+ with a 512 x 512 resolution of the same area was derived from a scene (Path 019, Row 036) taken on 28 October, 1999. Again the choice of data was convenience and was randomly picked. Both these image set were cut from an entire Landsat and SPOT scene and converted into tiff format. In this thesis, I will refer to these two image sets as SPOT and Landsat database.

This study dealt with randomly picked real remote sensing imagery. Many other studies in content-based image retrieval use sample artificial images selected from popular frameworks such as the VisTex (1995) texture database or the Brodatz album (Smith and Burns, 1997) and sometimes intermixed with their own set of imagery. In such environment, it is easy to obtain a more accurate retrieval rate because the distributions of the pixel values of the images to be retrieved are distinct from the rest.
3.2 Region Quadtree

Organizing based on hierarchy is a common method practiced by mankind since the beginning of civilization. The feudal system, government, business organizations, books in the library are some of the many examples of the hierarchical system. Due to its efficiency, effectiveness and simplicity as a management and problem-solving tool, this method of organization has been widely adapted to fields ranging from politics to science in different variations. Today, with the advent of computers, several new fields have emerged: image processing, image compression, computer vision, pattern recognition and GIS. These fields essentially deal with 2D imagery that consists of pixels. Even a relatively small 512 by 512 pixel image (Figure 1) has a large number of pixels (512 x 512 = 262144 pixels). When there are huge amounts of such imagery in a database, individually accessing and manipulating each pixel could be a time consuming process even for the fastest computer.

Computers are becoming faster. At the same time image databases are increasing in size too. Increasing the speed of processing alone will not solve the problem; an efficient method of organization is needed to help computer programs manage pixels in a huge image database.

The adaptation of the ancient hierarchical technique to organize image data is called region quadtrees (Figure 1), have been an active area of research for the past 25 years (Samet, 1980, Bartolini et al., 1999, Chad et al., 2000). One of the main advantage of using a region quadtree is that it avoids the need for accurate image segmentation (extraction of the feature of interest), which is usually an important step in CBIR (Ma 1997, Chad et al., 1997). The problem with region quadtrees is that it can be difficult for users to decide on which region to select to be used as retrieval
criteria when none of selected quadrants represents the whole object. In this paper, I will discuss a simple method used to implement region quadtree-based CBIR.

Region quadtrees are used to represent subsets of a two-dimensional imagery (Samet, 1982). Quadtree segmentation begins with the whole image and splits it into 4 equal regions. It can be conceptualized in two ways: Figure 1(a) and Figure 2, are the same hierarchy, illustrated differently.

(a) Level 1: 4 Quadrants   (b) Level 2: 16 Quadrants   (c) Quadtree: Level 4

Figure 1. Region Quadtree Decomposition of 512 x 512 Image.

Each of these 4 sub-regions can be split again into a maximum of 16 smaller blocks (Fig 1b) and so on. The number of blocks N at each level i is given by:

\[ N = 4^i, \; i = 0, 1, 2, \ldots \]

A geographical feature can be outlined in a similar way using region quadtrees (Figure 3). It uses a region quadtree at level 3 and 4 to select a river bend highlighted in yellow. The edge of the river is not well enclosed within the region quadtree. Higher accuracy can be achieved by using higher quadtree levels (Figure 3b). The
Figure 2. Hierarchical Representation of Quadtree.

If level 4 is used only 256 records need to be created (Table 1) and for each of these records, statistical and spatial relationships between the 32 pixels can be computed and this information can be used in content-based image retrieval. Fractal dimension using the box counting and lacunarity algorithms used in this study is only defined for squared quadrants above a 16 x 16 resolution. Using higher levels to outline a region is not a good idea even though it segments the feature more precisely. The imprecision at low levels is a blessing in disguise, since the algorithm will be able to produce better distinguishing indices when a mixture of heterogeneous pixels is used. So this is a good justification to use a region quadtree instead of the more
precise pixel level segmentation of the region of interest. The advantage of a region quadtree is its simplicity and does not require the segmentation of the region of interest.

<table>
<thead>
<tr>
<th>Level</th>
<th>No of Records</th>
<th>Quad Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>512</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>256</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>128</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>256</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>1024</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>4096</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>16384</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>65536</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>262144</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. Quadtree Levels, Number of Records and the Quadrant Size.

(a) Using Level 4 Quadrants to Select a River Bend.  
(b) Using Level 5 Quadrants Produce a Finer Outline of the River Bend.

Figure 3. Selecting a River Bend Using Region Quadtree.
There are many ways to implement a region quadtree. In computer science, usually a tree data structure (Figure 5, 6) is used, but here a simple list of quadrant information was sufficient since our imagery size is small (512 x 51). For each selected quadrant, the upper-left corner coordinate together with the width and height (width = height for a squared quadrant) was used. Level information for each selected quadrant that defines a feature is computed as follows:

\[
\text{Level} = \frac{\text{Whole Image Size}}{\text{Width of Selected Quadrant}}
\]

Example: If Image Size = 512, Width = 64 then \(2^{\text{Level}} = 512/64 \Rightarrow \text{Level} = 3\)

![Figure 4. Third Level Region Quadtree Quadrant Labeling.](image)

A 512 x 512 image decomposed to the 3rd level consists of 64 sub-images of size 64 (Figure 4). The sub images are logically labeled from 0 to 63. Only the upper-left coordinate (ULX, ULY) and width and height information is stored. The Quadrant number is computed as follows:

Quadrant number = ULY/height\(\times 2^{\text{Level}} + \text{ULX/Width}.\)
When the user selects Quadrant 20, the upper-left coordinate (256, 128) and width = height = 64 is stored. The quadrant number is computed from

\[
\text{Quadrant number} = \frac{128}{64} \times 2^3 + \frac{256}{64} = 20
\]

Figure 5. Approximate Outline of a Lake Using Region Quadtree. Selected Quadrants are Marked in Bold and Labelled from 1 to 7.
Figure 6. Logical Representation of the Selected Quadrants from Figure 5.
3.3. Metadata

3.3.1. Daubechies’ Wavelet Transform

While region quadtrees are a technique for organizing data, wavelet transform is a powerful analytical tool. In statistics, summarizing raw data provides a lot more useful information than just looking at the original. In the same way, when data is transformed to another form by applying some mathematical functions, patterns that are deeply hidden within the original data are revealed. Wavelets are special functions that not only serve as an analytical tool for such purposes but also provide a way to look at data at different levels of detail and scale making it a very valuable analytical tool.

Wavelets are used in many applications such as speech compression, music processing, image and video compression. The Joint Photographic Expert Committee’s JPEG 2000 format, a new image coding system produced significant improvement over the older JPEG format by ingeniously adopting wavelet transforms in producing highly efficient compression techniques, is one of the many examples clearly demonstrating the power of this technology. Even the FBI uses it for the compression of its huge collection of fingerprint data (Brislawn et al., 1996).

The history of wavelets dates back to 1807 when Fourier informally stated that any periodic function (waveform) can be expressed as the sum of sines and cosines of different frequencies. This lead to the development of the Fourier Transform (Fourier, 1955). This can be viewed as a mathematical prism, breaking up a function into the frequencies that compose it, as a prism breaks up light into colors (Hubbard, 1998). In other words, the Fourier transform enhances the frequency information and hides the time information of an original signal, making it a powerful analytical tool and letting
the user to discover underlying patterns and information, easily from a different perspective.

Fourier transforms solved many problems in digital signal processing and worked best when the original signals have periodic properties but they had shortcomings dealing with abrupt changes and discontinuity. They analyze the entire signal and loses time and location information. The need to preserve time and frequency localization gave birth to the wavelets transform which can be viewed as an improvement of Fourier transforms.

Wavelets are small or little ripples of finite duration, and are often irregular and asymmetric. A waveform is a mathematical function of a wavelet. A wavelet transform is the conversion of frequencies into a sum of sequentially ordered set of terms. When wavelet transformation is used on images, it is known as a discrete wavelet transform (DWT).

Multiresolution theory (Mallat 1989) is a framework for wavelet transform that laid a foundation for filtering techniques to be developed on images allowing image data to be viewed from different scale and detail. The process can be done by simply averaging groups of pixels together and keeping their differences. Ingrid Daubechies used this framework to further simplify wavelets through iterative methods and developed even simpler and shorter filters (Daubechies, 1988).

There are many forms of Daubechies wavelets based on different filter lengths. These Daubechies family wavelets are written as dbN, where N is filter length. db2 is a Haar transform and is based on a simple square waveform with 2 filters. It is the simplest to understand and easy to compute. In this research, a more complex waveform called the Daubechies wavelets db4 was used and since it has 4 filters (Figure 9), it has the advantage of picking up more missing details than the
Haar wavelet. The process of applying a forward 2D wavelet transform on images using db4 filters is illustrated in Figure 9,10,11,12.

Figure 7. Notation for Level 2 Standard Wavelet Decomposition.

Basically, an image of size $N \times N$ is decomposed into 4 sub images of size $N/2 \times N/2$. This sub images are represented as $LL_1$, $LH_1$, $HL_1$ and $HH_1$ as shown in figure 7. $LL_1$ is further decomposed into $LL_2$, $LH_2$, $HL_2$ and $HH_2$. $L$ represents low frequency information, $H$ denotes high frequency information and the number specifies the level of decomposition. For example, $HL_2$ is the convolution of high pass filter in the horizontal direction and low pass filter in the vertical direction at
level 2. This process can be repeated again in the same manner up to any level and is known as pyramidal or multi resolution decomposition. It is common to use the LL quadrant for further decomposition and this process is known as the standard forward wavelet transform. If HL, LH or HH is used, it is a called the non-standard decomposition.

Figure 8. 2D Forward Wavelet Transform Using Daubechies' db4.

LL is also known as an approximate image or average image. It basically takes the original image to the next scale. Since HL, LH and HH contain details of how to reconstruct LL to the original, it is called the detail images. LH and HL are referred to as middle frequencies. HL1 and LH1 quadrants contain horizontal and vertical features of the original image respectively, while HH1 faintly reveals diagonal features (Figure 7).
Figure 9. Sample 16x16 Image of a River Bank for Illustrating Daubechies’ db4.

Figure 10. Actual 16 x 16 Grayscale Values of a River Bank in Figure 9(a).
\[ f_1 = \frac{(1 + \sqrt{3})}{4\sqrt{2}} \]
\[ f_2 = \frac{(3 + \sqrt{3})}{4\sqrt{2}} \]
\[ f_3 = \frac{(3 - \sqrt{3})}{4\sqrt{2}} \]
\[ f_4 = \frac{(1 - \sqrt{3})}{4\sqrt{2}} \]

where \( \sum_{i=1}^{4} f_i = 0 \) and \( f_1 \cdot f_3 + f_2 \cdot f_4 = 0 \)

**Averaging (Smoothing Filters)**

**Differencing (Non-Smoothing Filters)**

\[ g_1 = f_4 \]
\[ g_2 = -f_3 \]
\[ g_3 = f_2 \]
\[ g_4 = -f_1 \]

**Figure 11. Daubechies’ db4 Smoothing and Non-Smoothing Filters.** Smoothing and Differencing Filters are Also Known as Lowpass and Highpass Filters Respectively.

### 3.3.1.1 Description of Daubechies’ Wavelet Transform on a Sample 16 x 16 Image.

Step 1: Let the first row of the 16-pixel width image be \( a_1, a_2, \ldots, a_{16} \) (Figure 10).

Step 2: Convolve averaging filters on \( a_1, a_2, \ldots, a_{16} \) to bring out approximate (average) coefficients \( b_1, b_2, \ldots, b_8 \).

**Averaging Coefficients**

\[ b_1 = a_1 f_1 + a_2 f_2 + a_3 f_3 + a_4 f_4 \]
\[ b_2 = a_3 f_1 + a_4 f_2 + a_5 f_3 + a_6 f_4 \]
\[ b_3 = a_5 f_1 + a_6 f_2 + a_7 f_3 + a_8 f_4 \]

\[ \cdots \]
\[ b_7 = a_{13} f_1 + a_{14} f_2 + a_{15} f_3 + a_{16} f_4 \]
\[ b_8 = a_{15} f_1 + a_{16} f_2 + a_1 f_3 + a_2 f_4 \]
Step 3: Apply differencing filters on $a_1, a_2, \ldots, a_{16}$ to obtain a set of differencing (detail) coefficients $b_9, b_{10}, \ldots, b_{16}$.

**Differencing Coefficients**

\[
\begin{align*}
b_9 &= a_1g_1 + a_2g_2 + a_3g_3 + a_4g_4 \\
b_{10} &= a_3g_1 + a_4g_2 + a_5g_3 + a_6g_4 \\
b_{11} &= a_5g_1 + a_6g_2 + a_7g_3 + a_8g_4 \\
\vdots \\
b_{15} &= a_{13}g_1 + a_{14}g_2 + a_{15}g_3 + a_{16}g_4 \\
b_{16} &= a_{15}g_1 + a_{16}g_2 + a_1g_3 + a_2g_4
\end{align*}
\]

Step 4: Replace $a_1, a_2, \ldots, a_{16}$ with pixels from the next row (Figure 10).

Step 5: Repeat step 2 to 4 for all rows.

If this process is applied to all the rows in a 256 x 256 image, the output will look like Figure 12b. The display on the left is the approximate image of the original and it appears squeezed horizontally due to effect of the averaging filters. The display on the right is the detail image is a result of a differencing filter. The detail image appears rather fuzzy due to difficulty in displaying negative floating points number (a product of applying the differencing filter). However, when linear contrast stretching is applied the detail image emphasizes the vertical features (Figure 12c).

Step 6: Go through step 1 to 4 but this time for all the columns in figure 10.

The final output will be figure 9b and the actual wavelet coefficients are shown if figure 11. If similar process was applied on a 256 x 256 image the result would be figure 12d.
Figure 12. The Effect of Level 1 Wavelet Transform Decomposition on a Real 256x256 Image.

Figure 13. Computed Daubechies' db4 Wavelets Coefficients. Figure 7(b) is the Result of the Stretching These Coefficients and Display it as an Image.
Texture classification accuracy using wavelet transform coefficients was shown to improve if middle frequencies (LH or HL) were used in the decomposition as opposed to the standard decomposition method where normally lower frequencies are used (Zhu, 1998). Zhu (1998) calls his method the best wavelet decomposition (BWD). The BWD was not used in the current research due to time constraints and the differences between BWD and standard wavelet transform is not very significant.

3.3.1.2 Wavelet Energy Signature

Energy can be defined as the capacity to do work. There are many forms of energy: kinetic energy, potential energy, heat energy, thermal energy, chemical energy, electrical energy, sound energy, nuclear energy and electromagnetic energy. In electrical engineering, the energy $E$ of a signal $f$ is defined by

$$E = f_1^2 + f_2^2 + ... + f_N^2.$$

In the same way the energy of a sub image $E$ is defined as

$$E = \frac{1}{N} \sum_{i=1}^{N} C_i(x,y)^2$$

where $N$ is the total number of wavelet coefficient and $C(x,y)$ is the wavelet coefficient at location $(x,y)$.

Wavelet energy signatures have proven to be very powerful for texture characterization (Zhao, 2001, Smith, J.R. et. al, 1994). In the case of the sample image in figure 9, 98% of the energy is concentrated on the LL1 sub image (Table 2). Even though LH1 and HL1 only make up 1.5% of the total energy, this small amount of energy is good indicator of sharp intensity changes in the image which corresponds to the boundary between the river and land.
Table 2. Wavelet Energy Signature of Figure 13.

<table>
<thead>
<tr>
<th>Sub image</th>
<th>Energy</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL1</td>
<td>4972487</td>
<td>98.3%</td>
</tr>
<tr>
<td>LH1</td>
<td>32319</td>
<td>0.6%</td>
</tr>
<tr>
<td>HL1</td>
<td>43532</td>
<td>0.9%</td>
</tr>
<tr>
<td>HH1</td>
<td>9911</td>
<td>0.2%</td>
</tr>
</tbody>
</table>

When an image is decomposed using 2D wavelet transform, it provides four different ways to view and analyze the image data. If it is taken to the next level and another set of four is available for analysis. At each level of decomposition the image data can be analyzed in a different light. That is why it is called the multi-resolution analysis tool.

3.3.2 Spatial Autocorrelation and Moran's $I$

Everything is related to everything else, but near things are more related than distant things (Tobler, 1969). This statement is also true for a remotely sensed image where adjacent pixels may be related or dependent. This means that these images contain spatial structures that are important characteristics for classifying the imagery and may also assist in image retrieval. Dependency in observation is a violation in measure of independence that is essential in non-social statistical analysis. Even Karl Pearson and William Gosset, founding fathers of statistics, had discussed these problems and pointed out that the validity of such a measure can only be applied on variables connected to time but not with space (Student 1914). However, P.A.P. Moran (1950) produced an entirely new class of global descriptive statistics known as Moran's $I$, which computes the degree of
dependency. The indices measured the strength of dependent observations indicating whether they are checkered (Moran’s $I = -1$), clustered (Moran’s $I = 1$), or randomly distributed (Moran’s $I = 0$).

Figure 14 shows simulated images that demonstrate the theoretical Moran’s $I$. Regions from the image database that are the closest to these theoretical values are illustrated in figure 15. It was hard to find a checker board image in the database and the lowest $I$ value was found to be -0.2698, which is far from -1. Even though the region in figure 15(a) appears dark and a Moran’s $I$ of one would be expected but the dump of the raw pixel values in figure 16 feebly exhibit a faint checkerboard properties. This is the closest value of a checkered image that can be found in the database. However, figure 15(b) and 15(c) closely resembles the theoretical values of figure 14(b) and 14(c) respectively. The cross-product statistical form of Moran’s $I$ [1] that was used for computing global spatial autocorrelation is:
\[ I = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w(i,j)(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \]

Where, \( \bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \) and \( S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} w(i,j) \)

(a) \( I = -0.2698 \)  
Best Checkered Image

(b) \( I = 0.0197 \)  
Best Random Image

(c) \( I = 0.9763 \)  
Best Clumped Image

Figure 15. Moran's I Indices Computed in the Marked Region on Actual Images That Closely Matches the Theoretical Values of Figure 16 from the Landsat Database.

| 74 | 74 | 65 | 70 | 65 | 65 | 61 | 74 | 61 | 65 | 70 | 65 | 61 | 61 | 74 | 57 |
| 57 | 74 | 61 | 61 | 57 | 65 | 57 | 57 | 70 | 53 | 65 | 61 | 74 | 53 | 65 | 65 |
| 70 | 53 | 74 | 57 | 74 | 65 | 61 | 61 | 61 | 53 | 74 | 61 | 70 | 57 | 65 | 53 |
| 65 | 57 | 65 | 61 | 65 | 70 | 61 | 61 | 61 | 74 | 57 | 78 | 65 | 70 | 65 | 70 |
| 74 | 65 | 65 | 57 | 65 | 65 | 65 | 65 | 83 | 53 | 74 | 61 | 65 | 74 | 70 | 48 |
| 53 | 74 | 57 | 61 | 65 | 65 | 65 | 65 | 61 | 65 | 53 | 65 | 53 | 74 | 57 | 74 |
| 74 | 65 | 61 | 61 | 65 | 65 | 61 | 78 | 65 | 65 | 70 | 70 | 70 | 70 | 65 | 57 |
| 53 | 57 | 65 | 70 | 65 | 61 | 70 | 61 | 65 | 53 | 65 | 61 | 70 | 74 | 65 | 65 |
| 70 | 65 | 65 | 70 | 70 | 74 | 65 | 65 | 61 | 57 | 83 | 65 | 78 | 57 | 70 | 61 |
| 65 | 65 | 61 | 61 | 61 | 65 | 65 | 57 | 70 | 61 | 65 | 57 | 57 | 70 | 61 | 78 |
| 74 | 65 | 70 | 65 | 70 | 61 | 65 | 65 | 61 | 65 | 65 | 70 | 61 | 70 | 65 | 65 |
| 61 | 65 | 78 | 65 | 70 | 61 | 65 | 61 | 74 | 61 | 65 | 65 | 65 | 61 | 70 | 70 |
| 70 | 65 | 65 | 61 | 83 | 57 | 61 | 70 | 57 | 65 | 57 | 70 | 65 | 65 | 65 | 65 |
| 57 | 70 | 53 | 57 | 57 | 61 | 61 | 57 | 65 | 65 | 65 | 70 | 61 | 70 | 53 | 70 |
| 57 | 74 | 65 | 61 | 65 | 65 | 61 | 74 | 57 | 70 | 74 | 61 | 70 | 70 | 61 | 74 |
| 61 | 61 | 61 | 78 | 48 | 70 | 61 | 65 | 65 | 57 | 74 | 53 | 61 | 61 | 70 | 65 |

Figure 16. Pixel Dump of Marked Region of Figure 15(a).
3.3.3 Fractal Dimension

Fractal dimensions also play an important role in texture analysis (Zhu et al., 1998; Chaudhuri et al., 1995; Lam et al. 2002). The term fractal came from the Latin word frangere, which means to break or fragment (Mandelbrot 1983). A fractal is a shape or pattern that is self-similar when it is enlarged. Fractal dimension $d$ can be computed by taking the limit of the quotient of the log change in fractal or object size and the log change in measurement scale, as the measurement scale approaches zero.
\[ \lim_{r \to 0} \frac{\log(N_r)}{\log(1/r)} = d \] [2]

Fractal dimension is a real number that expresses the complexity of a spatial pattern and usually ranges between zero and three. A point pattern is represented by a number between zero to one, a curve between one and two and a surface between two and three. For a surface, value two indicates a smooth surface and a number that approaches three indicates a rough surface. The range of \( d \) depends on the size of the region: the larger the area, the smaller the interval size (Table 3). There are many ways to compute fractal dimension and in this research, triangular prism (Clarke, 1986) and box counting method (Chaudhuri et al, 1995) was used.

<table>
<thead>
<tr>
<th>Level</th>
<th>Size</th>
<th>( d ) (Minimum)</th>
<th>( d ) (Maximum)</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>512</td>
<td>2.5067</td>
<td>2.6841</td>
<td>2.6172</td>
<td>0.0287</td>
</tr>
<tr>
<td>2</td>
<td>256</td>
<td>2.4500</td>
<td>2.7321</td>
<td>2.6366</td>
<td>0.0362</td>
</tr>
<tr>
<td>3</td>
<td>128</td>
<td>2.2952</td>
<td>2.7251</td>
<td>2.5865</td>
<td>0.0419</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>1.9804</td>
<td>2.6224</td>
<td>2.4494</td>
<td>0.0464</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>1.6018</td>
<td>2.9260</td>
<td>2.4544</td>
<td>0.0670</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 3. Fractal Dimensions (Box Counting Method) Statistics for Landsat Database.

In a study to characterize remotely sensed images it was found that fractal dimension was helpful for characterization, analysis, and classification of remotely sensed images (Zhao 2001). In still another study, it was found that adding a fractal dimension layer in multispectral classification of image texture increased accuracy through the addition of pixels in some land cover classes and a decrease of them in
others (Emerson, et al, 1999). Fractal dimension and standard deviation could be collectively used as part of the metadata because they form a broad impression of an image even before classification (Lam et al., 2002).

3.3.4 Lacunarity

A single fractal dimension does not provide a unique textural description, as it may only be applicable on a limited scale (Parrinello and Vaughan, 2002). Using fractal dimension alone will not be very useful since a single fractal dimension may show strikingly different texture (Figure 18). Another complementary measure called lacunarity was introduced by Mandelbrot (1982) to describe different textures represented by the same fractal dimension.

![Lacunarity: 2.94](image1) ![Lacunarity: 1.34](image2)

Figure 18. The Fractal Dimension of the Two Squared Regions is 2.41 but Their Lacunarity is Very Different.

Lacunarity was initially introduced as a way to differentiate fractals and textures which had the same fractal dimension but a different visual appearance. It is a scale-dependent measure of spatial heterogeneity or texture of a landscape. Methods for calculating lacunarity were first developed in general terms by Mandelbrot (1982) and a number of algorithms for calculating lacunarity have been proposed, developed
and revised (Allain and Cloitre, 1991; Plotnick et al., 1993, Dong, 2000). In this research, a straightforward and computationally efficient gliding box counting algorithm method proposed by Dong (2000) was used. Lacunarity $\Lambda$ was estimated from:

$$\Lambda(r) = \frac{\sum M^2 Q(M,r)}{\left[ \sum M Q(M,r) \right]^2}$$

where $Q(M,r)$ is the probability function for the number of boxes of size $r$ and mass, $M$.

<table>
<thead>
<tr>
<th>Level</th>
<th>Size</th>
<th>Lacunarity (Minimum)</th>
<th>Lacunarity (Maximum)</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>512</td>
<td>1.0185</td>
<td>1.4315</td>
<td>1.0609</td>
<td>0.0357</td>
</tr>
<tr>
<td>1</td>
<td>256</td>
<td>1.0195</td>
<td>2.0752</td>
<td>1.0641</td>
<td>0.0518</td>
</tr>
<tr>
<td>2</td>
<td>128</td>
<td>1.0246</td>
<td>3.0139</td>
<td>1.0724</td>
<td>0.0790</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>1.0432</td>
<td>625.0000</td>
<td>1.1142</td>
<td>3.4857</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>1.1211</td>
<td>4645.7851</td>
<td>1.2129</td>
<td>13.2371</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Lacunarity Statistics.

Lacunarity is basically a measure of holes or gaps of a fractal. A dense (with few holes) fractal is characterized by a small lacunarity values and higher lacunarity values suggest increased gappiness associated with heterogeneity of pixel values.
3.4. Content-based Image Retrieval Software Overview

The Image Characterization and Modeling System (ICAMS) was originally a C++ application that provided specialized spatial analytical functions for visualizing and interpreting remote sensing imagery (Emerson et al., 1999). The program was completely revised and rewritten in Java and can run as a stand alone application or as a Java Web Start application. ICAMS-Java works in Windows, Linux, Macintosh and UNIX environments.

The application was also extended to include wavelet transforms, Getis’ G index of spatial autocorrelation (Anselin, 1995) and region quadtree-based CBIR capabilities. The CBIR system architecture is outlined in Figure 21. A batch version of ICAMS was also created and was used to populate the two image databases. Classes in ICAMS were called to compute and store the various indices as metadata in the image database for all the 272 SPOT and 507 Landsat test images. These computed indices were stored into the appropriate fields in an Access database (Figure 19).

![Image of Microsoft Access table]

Figure 19. Snapshot of the Spatial Index Table Containing 371280 Records at 6 Different Levels of the Quadtree Representation for the 272-Segmented Imagery.
Figure 20. Snapshot of the Spatial Index Table Containing 690690 Records at 6 Different Levels of the Quadtree Representation for the 507-Segmented Imagery Consisting of Wavelets Energy Signature.
CBIR System Architecture Overview

**Database Population Phase**

- Select Landsat 7 ETM+ and SPOT scene

- Image Preparation 507 and 209 512 x 512 subset Landsat 7 and SPOT scene to TIFF format

- Apply region quadtree decomposition up to 6 levels on each image

- Compute Moran’s I, Fractal Dimensions, Spectral Histogram, Wavelet coefficients for 3 levels of decomposition for each quadrant

- Populate Database with indices and pointers to image locations

**Image Retrieval Phase**

- SQL Query Construction

- Query Database and select required levels and quadrants.

- Assign weights to indices to allow the use of combined indices

- Use least square classifier

- Rank and mark matched region

**User Interface - Input**

- Select Query Image

- Use Region Quadtree to select feature of interest

- Select query parameters

**User Interface - Output**

- Ranked query results marked region for display

**Figure 21. CBIR Architecture.**
3.5. Methodology

Each of the 272 SPOT and 507 ETM+ tiff images was segmented into non-overlapping regions using six levels of a region quadtree. Lacunarity and fractal dimension using box counting were only computed up to fifth level because they require at least a segment size of 32x32 pixels. Each image produced a total of 1,325 set of indices. Hence the total number of records in the table was $1365 \times 272 = 371,280$ and $1365 \times 506 = 690 \times 690$ for the SPOT and Landsat database respectively. Each record is linked to an image table, which has a pointer to the actual image location in the hard disk. A batch version of ICAMS was written in Java to do this task.

A user-friendly interface was implemented to query the database to retrieve stored information and images in an appropriate form. Java Swing, a set of classes supplied in Java for the development of Graphical User Interface (GUI) was used to develop the interface extending ICAMS functionalities to query the image database for content-based image retrieval. Each computed index was analyzed to determine the underlying distribution and look for outliers. The results were used to decide on the type of parameters to be used in the matching and ranking algorithm.

Typically, the user selects a geographical feature of interest such as a river bend using a region quadtree structure (Figure 3a, 3b). The software then searches the database for similar images with or without translation. Textures indices associated with the selected feature will be computed and a dialog box will display the results of the query indices (Figure 22).
The user may inspect the feature indices in the combo box and select a combination of indices to be used in the similarity search for a river bend in the image database. The program will then proceed to retrieve imagery based on the given criterion and rank the images in the database. The ranking score is based on least sum of squared differences (LSSD) or least square classifier. The ranked images are displayed in the combo box ready to be displayed for inspection.

Combination of two distinct sets of indices in the ranking process can amplify the influence of the indices with the larger range of values. For example, Moran’s $I$
range of values between -1 and +1 and the mean value of the pixels ranges between 0 and 255. Since the sum of squared differences of Moran's I will be much smaller, the sum of squared differences of the mean will become dominant in the ranking process. In order to assign equal weighting to each of the search indices, the range of values of all the indices was linearly adjusted to be between 0 and 1. This was accomplished by dividing each index by its range (maximum – minimum).

Translation was implemented by enclosing all selected quadrants into a smallest bounded rectangle. The upper left coordinate of the bounding rectangle was determined and used as a translation vector to move the rectangle to location (0,0). The smallest quadrant size in the bounded rectangle is then taken as the step size to traverse the entire the region of an image.

Scaling and rotation of 90, 180 and 270 degree could be implemented in the same way as translation but was not used in this study due to time constraint.
CHAPTER 4

RESULTS

Retrieval experiments were conducted on SPOT and Landsat image database and the results are presented below. In all the retrieval results, the top left image is the query image. The retrieved images are ranked based on the least sum of square deviation and displayed in ascending order. In some queries, especially those pertaining to the whole image, the top six and the bottom six are presented to show the contrast.

4.1 Retrieval Results from SPOT Database

Figure 23 shows the retrieval of a road feature selected by using a non-translated region quadtree. The two index criteria included were Moran’s $I$ and fractal dimension using the box counting method. Both the indices were given equal weights. The first one is obviously the image itself but the third and the fifth images have a road running along the quadtree outline of the query image. This seems to indicate that image retrieval based on texture indices is not impossible.

However, when translation was included, the results appear to be rather disappointing (Figure 24). Only the first two images matched. Nonetheless, the outline of the location on the images where the best matches occurred seems to cross road(s) at some point. This needs to be examined more carefully before a conclusion can be drawn.

Figure 25 shows the top 5 matching images ranked based on a single quadrant at level 4 shown in bold outline, which is a part of the river. The retrieved images are
based solely on the mean. No spatial indices were used. Exactly one level 4 quadrant is selected in the query image. The first two retrieved images correctly displayed imagery with river content at that location. Since the query image was one of the 272 images, the top image ranked is the query image itself. The SSD of such an image is always 0. This shows that simple descriptive statistics can be useful in retrieval of remote sensing imagery.

When Moran’s $I$ and fractal dimension based on triangular prism was used, to retrieve images for the river bend, the results appear to be disappointing (Figure 26). None of the images matched except the first one. Even so, the unmatched images do reveal a patch, which mimics a river at the marked location. This needs to be examined more carefully. It can be inferred that fractal dimension using box counting seems to give more accurate matches than using the triangular prism method.
Figure 23. Similarity Search Based on Moran’s $I$ and Fractal Dimension Using Box Counting Method on a Road Feature Without Translation.
Figure 24. Similarity Search Based on Moran’s I and Fractal Dimension Using Box Counting Method on a Road Feature with Translation.
Three sets of similarity search results are presented.

Figure 25. Similarity Search Based on Mean and a Level 3 Selected Quadrant.
Figure 26. Similarity Search Based on Moran’s $I$ and Fractal Dimension Using Triangular Prism Method on a Portion of a River Bend.
4.2 Retrieval Results from Landsat Database

After getting mixed results from the SPOT database, the Landsat database was populated and the retrieval accuracy analyzed. Moran's I, fractal dimension, lacunarity and wavelet energy signature was applied independently to each of the whole image. The top six and bottom six in ascending order are presented. Then, the usage of the region quadtree technique, together with equal weight combination of various indices in image retrieval is demonstrated below.

4.2.1 Wavelet Energy Signature

One of the factors that contributes to the effectiveness of image retrieval is the choice of appropriate levels. Wavelets energy signatures from the Landsat database were very sensitive to the level used in the search query. As the level increases, the range of wavelet energy increase exponentially for LH1, HL1 and HH1. This increases the complexity in the retrieval accuracy when a region is selected with a mixture of quadrants from different levels and used with combination of different indices. Even though every index was linearly adjusted to be between 0 and 1, the distribution wavelet energy signature tended to be skewed as revealed in Figure 27 which shows the huge variation of the maximum from the average.
In this research, wavelet signatures were computed to 3 levels of decomposition for each region. The sorted energy signature (sub image LH1) for the whole image for the Landsat database seems to reveal a pattern. The top six images with the lowest energy signature (sub image LH1) contain a river or a lake (Figure 30). All six show strong simple forms as well. The bottom six images are generally urban areas (Figure 31). A similar property is shown for energy signatures for sub image HL1 (Figure 32, 33) and there is slight variation in this trend for HH1 (Figure 34, 35). However, in the case of the energy signature (LL1 sub image), the situation
appears to be reversed (Figure 28, 29) with the bottom six displaying a mixture of mountains and rivers (Figure 29).

Figure 28. Wavelet Energy (LL1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order.
Figure 29. Wavelet Energy (LL1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order.
Figure 30. Wavelet Energy (LH1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order.
Figure 31. Wavelet Energy (LH1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order.
Figure 32. Wavelet Energy (HL1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order.
Figure 33. Wavelet Energy (HL1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order.
Figure 34. Wavelet Energy (HH1 Sub Image) – Top Six Listing from Landsat Database in Ascending Order.
Figure 35. Wavelet Energy (HH1 Sub Image) – Bottom Six Listing from Landsat Database in Ascending Order.
4.2.2 Moran's $I$ Results

It is possible to classify images according to the type of land cover as demonstrated in figure 38 and 39 when Moran's $I$ is computed over the entire image. Moran's $I$ for the 512 x 512 image ranges from 0.7693 to 0.9436. The lower values are associated with urban areas and higher values are linked with water or river features (Figure 36(a) and 36(b)). This result is consistent with the results found for the wavelet energy signature found for sub image LL1. Further analysis on other levels needs to be conducted to determine the minimum and maximum size where this trend occurs.

<table>
<thead>
<tr>
<th>Type</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>$0.7963$</td>
</tr>
<tr>
<td>Mountain</td>
<td>$0.8642$</td>
</tr>
<tr>
<td>Farm land</td>
<td>$0.8700$</td>
</tr>
<tr>
<td>Lake</td>
<td>$0.9436$</td>
</tr>
</tbody>
</table>

Figure 36. Moran's $I$ Indices on Real Images at Level 0.
Moran's $I$ does not seem to vary too much as the level increases (Figure 38). However, due to algorithm constraints, Moran’s $I$ is not reliable beyond level 3 and cannot be computed for levels higher than 4.

Figure 37. Moran’s $I$ at Level 0 to 3.
Figure 38. Level 1 Moran’s $I$ – Top Six Listing from Landsat Database in Ascending Order.
4.2.3 Fractal Dimension Using Box Counting Method Results

Similar to the energy signature for sub image HL1, LH1 and HH1, the lower values of fractal dimension at level 1 are associated with lake or river (Figure 41) and higher values predominantly consist of urban settings (Figure 42).

Similar to Moran’s $I$, fractal dimension does not seem to vary much as the level increases but fractal dimension using box counting method can be computed for a squared image of size 32 pixels. When the level increases (or region size decreases), there is a wider range of values for fractal dimension.
Figure 40. Fractal Dimension at Level 0 to 4.

Figure 41. Level 1 Fractal Dimension \( d \) (Box Counting Method) – Top Six Listing from Landsat Database in Ascending Order.
Figure 42. Level 1 Fractal Dimension $d$ (Box Counting Method) – Bottom Six Listing from Landsat Database in Ascending Order.
4.2.4 Lacunarity Results

The lacunarity range for the whole image (512 x 512) in the study of the Landsat database varied from only 1.02 to 1.43 but at level 4 (32 x 32), it had a very wide range of 1.12 to 4645.79. When the size of the region decreases (or level increases), the maximum lacunarity increases drastically and in isolated cases can go be up to 4645.79 (Figure 44). In small regions, this is possible because the black regions can be interpreted as gaps and sufficient pixels with 0 values can increase to big numbers. The highest lacunarity (Figure 44(a)) seems to appear in the areas containing water and the lowest values appear in urban areas. Figure 45 and 46 shows that range of values needs to be adjusted statistically when lacunarity is combined with other indices.
Figure 44. Top Three Biggest Lacunarity in Landsat Database.

Figure 45. Lacunarity – Top Six Listing from Landsat Database in Ascending Order.
Figure 46. Lacunarity – Bottom Six Listing from Landsat Database.
4.2.5 Results – Combination of Various Indices.

Figure 47. Box Counting, Lacunarity and Moran’s $I$ Without Translation.
Figure 48. Wavelet Coefficient HH1, HH2, HH3 Without Translation.
Figure 49. Lacunarity with Translation.
Figure 50. Lacunarity with Translation.
Figure 51. Top 11 Retrieval Based on LH1 Applied on All the 4 Quadrants.
A need for an effective and efficient content-based image retrieval system leads to the proliferation of remotely sensed imagery. Preliminary results using textural indices such as Moran’s $I$, wavelet energy signatures lacunarity, and fractal dimensions show some promise but much work is required before it can be successfully applied in CBIR. Instead of relying on spatial indices, inclusion of other low-level features such as spectral histograms and other per-pixel measures may improve the overall accuracy of image retrieval system. An effective algorithm to combine and weight the indices must be implemented by taking into consideration factors such as the way images are classified by each of these indices with different region sizes before assigning weights. The remote sensing image retrieval framework developed in this research must also be integrated with text-based system to increase its effectiveness rather than relying solely on image content. The region quadtree was found to be an elegant means of avoiding laborious, time-consuming pixel level segmentation. The use of the region quadtree structure together with Moran’s $I$ and fractal dimensions in CBIR showed mixed results and require more rigorous research and analysis. The results presented must be compared with research obtained using pixel level segmentation to evaluate the real benefits and true potential of the hierarchical tool.
REFERENCES


Samet, H., 1990, The Design and Analysis of Spatial Data Structures, Addison-Wesley, Reading, MA.


Student, W., 1914, The elimination of spurious correlation due to position in time or space, Biometrika, 10: 179-180.

VisTex, 1995, VisTex texture database, Maintained by the Vision and Modeling group at the MIT Media Laboratory
http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html


Zhao, W., 2001, Multiscale analysis for characterization of remotely sensed images, Ph.D. Dissertation, Louisiana State University, December.