Graduate Theses, Dissertations, and Problem Reports

1998

Scale and texture in digital image classification

Christopher John Sebastian Ferro
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation

SCALE AND TEXTURE
IN DIGITAL IMAGE CLASSIFICATION

Christopher John Sebastian Ferro

Thesis Submitted to
the Eberly College of Arts and Sciences at West Virginia University
in partial fulfillment of the requirements for the degree of

Master of Arts
in
Geography

Timothy Warner, Chair
Gregory Elmes
M. Duane Nellis

September 25, 1998
Morgantown, WV

Keywords: remote sensing, scale and texture, edge effect

ABSTRACT

SCALE AND TEXTURE IN DIGITAL IMAGE CLASSIFICATION

Christopher John Sebastian Ferro

(ABSTRACT)

This thesis is a theoretical and empirical study on textural properties of digital images. Spatial information exists at a hierarchy of scales and texture is a consequence of the objects in that hierarchy. Within-class texture results from the spatial arrangement of objects at the next finer level in the hierarchy than the informational class. Between-class texture results from spectral differences between adjacent classes and is most obvious near class edges, especially for smooth classes. In rough classes between-class variance may not differ much from within-class variance. Errors in classifications using texture, therefore, are most likely associated with class edges; however, investigators often avoid edges in evaluating texture or classification.

The window sizes needed to produce a stable texture measure are often large. Experiments with ADAR 1-meter data suggest that windows of 50 to 300 meters are necessary. Small windows are required to minimize edge effects. This is inherently contradictory as windows used to produce stable texture measures also produce a large edge effect.

Experiments with simulated data showed that separability of classes increased when texture was used in addition to spectral information. Separability of texture also improved with larger scale windows. This improvement was over-estimated when pixels were chosen away from class edges. The ADAR data showed that separability of the interiors of classes improved with the addition of texture, but for the class as a whole, the class separability actually fell. Maximum Likelihood classification of the ADAR data demonstrated the effect of edges and multiple scales in reducing the accuracy of classification incorporating texture.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

Remote sensing incorporates both the processes by which information is gathered about an object through reflected or emitted electromagnetic radiation and the acquisition and analysis of images that are derived from that information. Remotely sensed images can be analog (as in a photograph) or digital (as in electro-optical or digital camera data). This thesis is an investigation of the textural properties of digital images and how those properties affect spectral classification.

A. Types of Information

There are four primary types of information in digital images: radiometric, temporal, spectral, and spatial. Radiometric properties are partly determined by the number of gray levels, whereas temporal properties are related to the timing of the imagery (Lilesand and Kiefer, 1994). Spectral information is derived from measurements at specific wavelengths of the light reflected from objects on the ground back to the sensor. The width of the wavelength bands, the number of those bands, and their extent is known as *spectral resolution*. Spectral resolution of current satellite and aircraft-borne sensor data has a broad range. Low resolution *Satellite Pour l'Observation de la Terre* (SPOT) panchromatic data is acquired in one broad band extending from 0.51 μm to 0.73 μm (Begni, 1982). Medium spectral resolution Landsat Thematic Mapper (TM) data is acquired in seven bands from 0.45 μm through 2.35μm in the optical wavelengths and 10.4 μm through 12.5μm in the thermal band (Markham and Barker, 1985). Some aircraft-borne hyperspectral sensors, such as NASA's Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) and the Hyperspectral Digital Image Collection Experiment (HYDICE), collect data in hundreds of narrow bands and at very high spectral resolution (Vane, *et al.*, 1993; Basedow, *et al.*, 1994).

Spatial Resolution can be described by the smallest angular or linear separation between two objects of a specific contrast ratio that can be resolved, or imaged, by the sensor. For non-photography based imaging systems, a common measure is the linear distance of the ground projected *instantaneous field of view* (IFOV). For example, the French SPOT sensor has an IFOV of 20×20 meters with the multi-spectral scanner and 10×10 meters with the panchromatic scanner. The IFOV of the Landsat Thematic Mapper scanner is approximately 30×30 meters for the optical bands and 120×120 meters for the thermal band (Kramer, 1996). As a general rule, for an object to be detected by the imaging system, it needs to be twice the

length of the IFOV of the sensor and have contrasting spectral or spatial properties when compared to the background. An important property of any image is the spatial structure, or arrangement, of bright and dark tones. Spatial information can be expressed in terms of spatial autocorrelation, a measure of the similarity between a pixel and its neighbors. The degree of spatial autocorrelation is determined by the size and shape of the objects in the scene and the spatial resolution of the image (Woodcock, Strahler and Jupp, 1988a; 1988b). Without some kind of spatial structure, digital images would appear to be nothing more than random collections of pixels, or *noise.*

B. Image Texture

Texture is another expression of the local spatial structure in digital images. Wang and He (1990) have defined texture as the tonal or gray level variation of an image. Hsu (1978) defined it as the spatial distribution of tones of the pixels in remotely sensed images. Haralick (1979) described texture as a scale dependent phenomenon that arises from the spatial interrelationship between *tonal primitives* that comprise the scene. The term tone describes a specific brightness level along a continuum of gray levels from black to white. Tonal primitives are regions that may comprise a pixel, or group of adjacent pixels that have similar tonal properties (Haralick 1979). Texture is most often measured from the local spectral variation within a window of interest, such as a 3×3 or 5×5 pixel matrix, or window. This window is a grid of weighted values that is "passed" across the image to assign the central pixel a value based on the spectral variability of its neighbors. Central to this thesis is the idea that although the spatial resolution of the sensor is important in determining image properties (Woodcock and Strahler, 1987), it is the size of the window over which texture is calculated that determines the usefulness of texture measures. A further complication is that texture may exist at different scales in an image, with the finest scale depending on the size of an object in relation to the spatial resolution of the sensor instrument.

C. Texture versus Pattern

Texture is often confused with or used synonymously with *pattern*. For this thesis the term pattern is used to describe some spatial regularity within an area of interest. By contrast, texture is a measure of spatial variability, random or non-random, within an area of interest. Many computerized texture measures can not distinguish patterns from other textures since the technique is aspatial at the scale of measurement. The figure below (*Figure 1*) shows how some measures of texture can not distinguish pattern. The grid on the left contains twenty-five

numbers with the values 5, 10 and 20 in a specific pattern. The grid on the right contains the same numbers in random sequence.
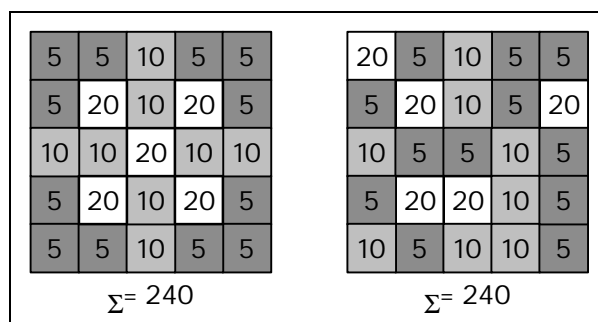
Figure 1: *Different patterns with the same first order statistical value.*

If the numbers are simply added in each grid, the sum for both grids is 240. Any further processing of this number is independent of the order in which the values occur in the matrix. This is true of texture measures using simple, or first order, statistics (*i.e., mean, variance etc.*). Texture measures that use higher order statistics, such as the Gray-Level Co-occurrence Matrix, are explicitly spatial and the ordering of number is significant. Texture measures will be discussed further in the next chapter. For this thesis first order statistics were employed in texture calculation.

D. The Edge Effect

One significant artifact in texture analysis is the so-called "edge effect," a domination of the texture image by bright pixels along the interfaces of spectrally different regions. Edges are bright because spectral differences between classes are usually greater than the differences within classes. Since the texture algorithm is simply assigning a value to a pixel based on variability of the values of its neighbors, texture estimates for locations near the edges of spectral classes will often therefore be greater than those within the interior of classes. For a texture image to be useful in standard spectral classification, each class should have a relatively distinct and homogeneous calculated texture value. Adaptive filters have been proposed to overcome this problem, but the results have been mixed. For example, the adaptive filter of Ryherd and Woodcock (1996) tends to produce a rather blocky output.

E. The Issue of Scale

The scale dependent aspects of texture have largely been ignored by previous studies, which have generally focused instead on the algorithms of texture calculation (Haralick, *et al.*,

1973; Weszka, *et al.*, 1976; Haralick, 1979; Connors and Harlow, 1980; Dutra and Mascarenhas, 1984; Marceau, *et al.*, 1990; Gong, *et al.*, 1992). This study will focus on scale through the development of a theoretical framework to facilitate understanding the relationship of scale and texture and an empirical study of texture in simulated and real image data. The theoretical framework will develop terms and concepts related to remote sensing models, scale and texture. The framework will also improve understanding about the nature of texture and draws on the terminology and ideas of Strahler *et al.* (1986) and Woodcock and Strahler (1987). The empirical study will then develop of a method for determining the appropriate scale (i.e., window size) for texture analysis using image plots of local variability of variance against window size.

II. *LITERATURE REVIEW*

Early texture analysis carried out in the 1970s with black and white (panchromatic) and multispectral images yielded promising results. Haralick, *et al.* (1973) achieved a classification accuracy of 74 to 77 percent using only spectral information. Identification accuracies of 83.5 percent were obtained with a combination of spectral and textural features. Hsu (1978) found that with data derived from texture alone, hit-rate analysis for test sites yielded accuracies of 85 to 90 percent. Numerous experiments in texture feature extraction and analysis from the 1970s to the present have included performance evaluations of, and comparisons between, various texture measures (Weszka, *et al.*, 1976; Conners and Harlow, 1980; Dutra and Mascarenhas, 1984; Latty, *et al.*, 1985; Marceau *et al.*, 1990; Gong, *et al.*, 1992).   An example of a more recent study is that by Agbu and Nizeyimana (1991), in which it was suggested that texture features had a potential for assisting with the initial stages of detailed soil survey programs and land-use planning. The use of texture derived from radar images has shown to be particularly beneficial in the discrimination of sea ice types (Barber and LeDrew, 1991). Ryherd and Woodcock (1996) used texture and spectral data in image segmentation of simulated forests, as well as real data including mixed urban and suburban land cover types.

A. Measures of Texture

As discussed in the last paragraph, the evaluation and development of new approaches to calculating texture has been a particular focus in the remote sensing literature. One class of image characteristics used for determining texture is first-order statistics of local areas. These include *mean, entropy,* and *variance*, shown in *Table 1* (Jensen, 1996):

Table 1*: First-order statistics as texture measures***.**

Texture Measure	Equation
Mean	$AVE = \dfrac{1}{W} \displaystyle\sum_{i=0}^{quant_k} i \times f_i$
Entropy	$ENT = \displaystyle\sum_{i=0}^{quant_k} \dfrac{f_i}{W} \ln \dfrac{f_i}{W}$
Variance	$VAR = \dfrac{1}{W} \displaystyle\sum_{i=0}^{quant_k} (i - AVE)^2 \times f_i$
Where $\quad f_i$ =frequency of gray level i occurring in a pixel window $\quad quant_k$ =quantization level of band k (e.g., 2^8=0 to 255) $\quad W$ =total number of pixels in a window.	

The variance filter provides a measure of local homogeneity in an image and can also be regarded as a non-linear non-directional edge detector (Wilson, 1997), because it emphasizes sudden changes in surface images brightness without any directional bias. For this reason it is useful for capturing within-class variability by detecting edges of scene objects.

Other ways of calculating texture include second-order statistical measures such as the gray-level co-occurrence matrix (GLCM). Haralick and Shanmugam (1974) define 32 GLCM texture measures. Weszka, *et al.* (1976) and also Marceau, *et al.* (1990) have evaluated the GLCM. Weszka, *et al.* (1976) found that the GLCM and other measures of texture using second order statistics resulted in poorer classification results compared to texture calculated from simple statistics. Fourier transform features, an approach borrowed from geophysics, had the lowest performance.

A method employed by Wang and He (1990) entails calculating texture units as elements of a *texture spectrum*. Jensen (1996) explains this method as "...computed texture based on an analysis of the eight possible clockwise ways of ordering the 3×3 matrix of pixel values". This is a set of nine elements representing the brightness value of the central pixel and the intensity of the neighboring pixels. The result is a *texture unit*, containing eight elements. There are 6561 possible texture units, with higher numbers representing "coarser" textures. The frequency of occurrence for each texture unit in an image or sub-image is the texture spectrum. Although this method could theoretically be extended to larger window sizes, the possible permutations of the texture spectrum would be impractically large.

Cushnie (1987) took a different approach to texture by considering intraclass variability as "scene noise." A two step experiment was employed by first simulating satellite data at progressively coarser spatial resolutions and comparing classification accuracies. Secondly, an averaging filter was used before classification reducing intraclass variability, enhancing classification accuracy. The effect of averaging on pixels near class edges was not evaluated.

B. Scale and Texture

According to Marceau *et al.* (1990), ninety percent of the variability in classification accuracy using texture is accounted for by the size of the window used to calculate texture while the particular texture algorithm determines only ten percent. This is a particularly significant finding because the majority of texture studies focus on algorithm development and evaluation without addressing the effect of window size.

Hodgson (1998) presents a good overview of window size as it relates to texture. Hodgson's (1998) research focused on cognitive (i.e., visual) classification, not automated classification. In his work, classification accuracy was related to window size in ground area units. It was found that a minimum size window was needed for accurate cognitive classification, and that the window size needed increased as spatial resolution increased. Once the accuracy reached a maximum, no further increase or decrease in accuracy occurred. Hay, Niemann and Goodenough (1997) also provide an excellent overview of scale and the methods for estimating image properties at new scales.

The subject of window size as it relates to texture analysis for computerized image classification has been mentioned by Hsu (1978), Chavez and Bauer (1982), Dutra *et al.* (1984), Marceau *et al.* (1990), and Briggs and Nellis (1991). Texture and scale were investigated in a systematic manner by Franklin and McDermid (1993) using image semivariograms calculated over forest stands to determine appropriate window sizes for texture calculation. Franklin, *et al.* (1996) also developed a technique for determining the optimal window size for digital image processing based on semivariograms. Chavez and Bauer (1982) developed technique for automatic window size selection for edge enhancement using the horizontal first difference, or derivative, of an image. This approach was solely concerned with a two-pixel difference as a measure of the scale rather than the influence of all pixels in a neighborhood from which texture might be calculated. Curran (1988) has demonstrated the use of semivariograms in determining image scale. Hsu (1978) and Dutra *et al.* (1984) suggest that small window sizes are a better choice due to the contaminating effect of edge pixels in classification. Dutra, *et al.* (1984) also suggest that smaller window sizes preserve "microtextures". Marceau, *et al.* (1990), on the other hand, found that maximum classification accuracy was achieved with window sizes that were class specific. Relatively large windows (17×17 and 25×25) worked well for most cover types. Marceau, *et al.* (1994) used graphs of local variance plotted against spatial resolution to derive optimal spatial resolutions for forested environments. Profiles of spectral separability were also graphed to determine the optimal spatial resolution for separating forest classes. Nellis and Briggs (1989) applied texture analysis to various scales of remotely sensed data and related them to tallgrass prairie landscape units. In addition to spatial scale, the sizes of the features in an image were found to influence texture measures. Larger window sizes were suggested for homogeneous landscapes.

Woodcock and Strahler (1987) provide a framework that graphically relates local variance to pixel size (spatial resolution). Their use of plots (see *Figure 2*) is conceptually similar to this project's second objective of developing a method to determine the appropriate scale for texture analysis using image plots of texture variability against window size.
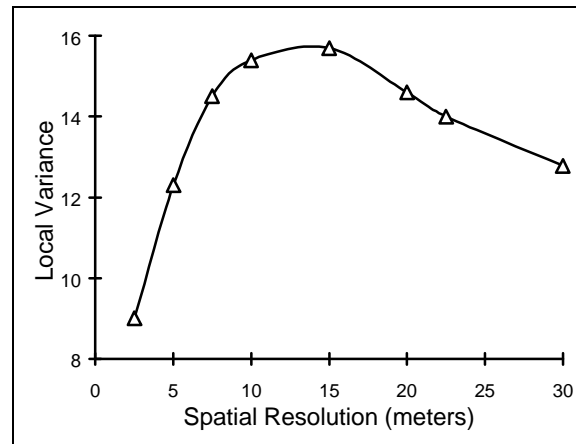
Figure 2: *Local variance plotted against spatial resolution (pixel size).*

Woodcock and Strahler (1987) related the shape of these plots to the scale of elements in the image. The peak of the local variance was found to occur when the pixel size reached approximately one-third to three-quarters the size of the objects, depending on object spacing, in the scene for both the simulated and observed images.

This thesis extends Woodcock and Strahler's work in two important ways. Firstly, their work dealt with a fixed window size of 3×3 pixels with varying spatial resolution. Secondly, Woodcock and Strahler's (1987) graphs were based on images which were considered a single class. Thus between class variability and its effect on texture was not considered. This project will therefore investigate the effects of varying both image resolution and window size. Furthermore, the effect of class edges will be a special focus of this work.

C. Models

A model that helps provide a framework for the analysis of remotely sensed images, and particularly texture, is the *discrete model* (Strahler, *et al.*, 1986). In this model, the scene is conceived as a background with discrete objects existing on it. These ground objects are abstracted as *elements* that have uniform *parameters* or *properties*. Elements may be unique, but often belong to *classes*. Elements of a single class are assumed to have the same set of parameters or properties. The background, in this case, is often regarded as spatially continuous with uniform properties or parameters, and is usually partially obscured by the

elements in the scene (Strahler, *et al.*, 1987). Strahler, *et al.* (1987) define two resolution models: the H-resolution model and the L-resolution model. When the spatial arrangement of ground features can be detected directly by the sensor as individual elements, the scene is said to conform to the H-resolution model. When the sensor is unable to distinguish individual elements, because the IFOV is larger than the elements, the scene is said to have an L-resolution.

III. THEORETICAL FRAMEWORK

In order to perform an evaluation of image texture for use in digital image classification, it is important to have an understanding of what controls texture in an image. The fact that there is no standard terminology for the concepts and processes involved in texture analysis adds to the difficulty of doing so. This chapter, therefore, is an attempt to provide greater clarity regarding terminology and the concepts of scale and texture which should, in turn, promote better use of texture measures in digital image processing.

A. Scenes and Images

The terms scene and image are often used interchangeably though the two are fundamentally distinct (*Figure 3*). The real scene is defined by the physical objects of which it is comprised. The image is an abstraction of the scene with scene objects manifested as image elements. These elements are analogous to the tonal primitives defined by Haralick (1979). The relationship between scene and image is two-way; the scene is defined by the extent of the image and the image is made of abstractions of scene objects.

One of the most important controls on the relationship between the scene and the image is the *scale of capture*, as defined by the sensor's effective resolution, normally approximated by the IFOV. It is this scale that determines whether a scene object is resolved in the image as a discernable element. The interface of the scale of capture and the scene results in the digital image (*Figure 3*). The sensor predetermines this scale, the equivalent of the spatial resolution of the image, and the investigator generally has no control over it.
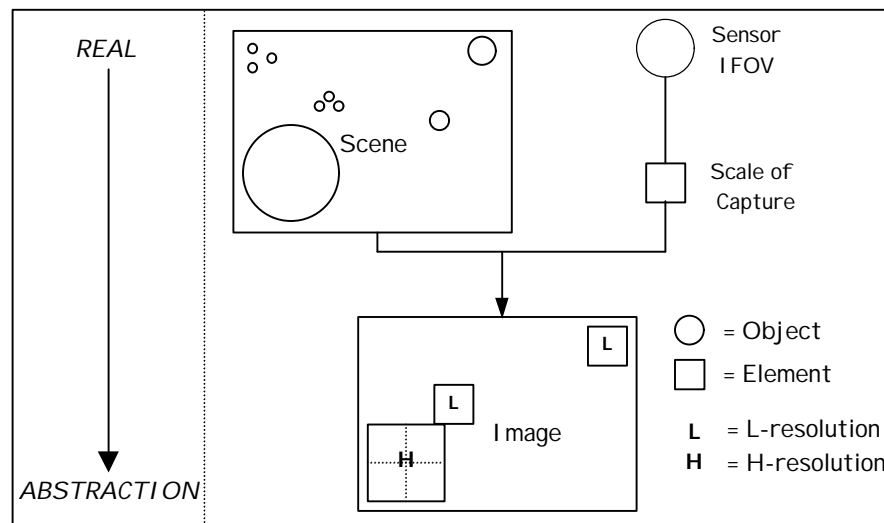
Figure 3: *Abstraction from real objects to image elements.*

Real scenes have objects that exist at various scales and consequently the image will have both *H-resolution elements* and *L-resolution elements*. L-resolution elements are often referred to as "mixed pixels." While Woodcock and Strahler (1987) state that H- and L-resolutions are image properties, it is the elements themselves that represent H- through L-resolution manifestations of real scene objects. Therefore, in this thesis, it is the image elements that will be referred to as having H- through L-resolution characteristics and it is expected that all images have both H- and L-resolution elements. Furthermore, in a single image the H- through L-resolution elements may vary in size along a continuum; however, generally image elements will cluster at specific scales. For example, real objects, such as *leaves*, are H-resolution elements at centimeter scales of capture, and give rise to an image that exhibits local spatial autocorrelation. Spatial autocorrelation is high with H-resolution elements because the scene objects are resolved and thus like pixel values are clustered. The objects become L-resolution elements at meter scales of capture where images *tend* to show spatial autocorrelation due to the spatial groupings at the next hierarchical scale; in this case trees.

Informational classes, determined by the investigator, exist at one of the various hierarchical scales in an image and thus there is an optimal scale of capture. For example, if 'Residential' is an informational class, scales of capture that resolve both 'Rooftops' and 'Lawn' as spectral classes will provide unwanted spectral variability. An optimal scale of capture (ignoring image collection costs) for identification and mapping the boundaries of an object is one that has an IFOV *smaller* than the informational class, but *larger* than the size of objects at the next finest hierarchical scale. This idea of optimal scale of capture differs from the well-known concept of efficient sampling, where the aim is to avoid spatial autocorrelation in samples.

B. Textures

Image textures result from an interface between the scale of capture and the scale of image elements. Because there is a hierarchy of scales of image elements, therefore there must exist a hierarchy of texture scales. Investigators are generally interested only with those texture scales within the informational class of interest.

For example, if an image includes houses captured at an H-resolution scale and "Residential" is the class of interest, the texture will be relatively coarse for that class. However, if "Rooftop" is the class of interest, the texture will be relatively smooth for that class.

While textures exist in the abstract, it is only by applying textures to images containing informational classes determined by the investigator that texture can be made useful for image classification. Texture is present if the *class statistics* are influenced by varying proportions of image elements at the next finer scale in the hierarchy. In conventional aspatial classification, using standard classifiers such as Maximum Likelihood, texture that is not smooth degrades classification accuracy by increasing the class variance.

The intersection of image texture and predetermined informational classes results in both a *between-class texture* and a *within-class texture.* The latter arises from the finer scale elements in relation to the class of interest. In order for image texture measures to be useful in spectral classification the within-class texture must be of a constant value and distinct from the texture measures of the other classes.   This is obtained through window sizes that are sufficiently large to encompass texture elements; however, between-class texture should not overwhelm the image due to excessive window sizes.

One possible approach to drawing on texture to *improve* classification is to use *local variance* over an appropriate window size. The window size must be large enough to span the texture elements and should be *at least* twice the size of the spacing between them. Window sizes larger than this are desirable since object sizes and spacing are not typically uniform.

The need for a large window size results in a trade-off between large window sizes that give stable texture measures and the increasing proportion of between-class variance texture pixels such large windows produce. *Figure 4* shows the relationship between window size, the percentage of pixels not influenced by between-class variance and the size of the class polygon. As window size increases the proportion of within-class texture pixels to total pixels generally decreases. For small polygons the rate of decline is most rapid. With a window size of 5×5 pixels a 10×10-pixel polygon has only 35% of its pixels unaffected by between-class variance. These unaffected pixels are referred to as *interior pixels*. Relatively large class polygons can have a sizable number of pixels that incorporate within-class variance with large window sizes.   A 600×600-pixel polygon, for example, has its percentage of with-class variance pixels reduced by five percent at a window size of 31×31 pixels.

Proportion of Within-Class Texture Pixels to Total Pixels
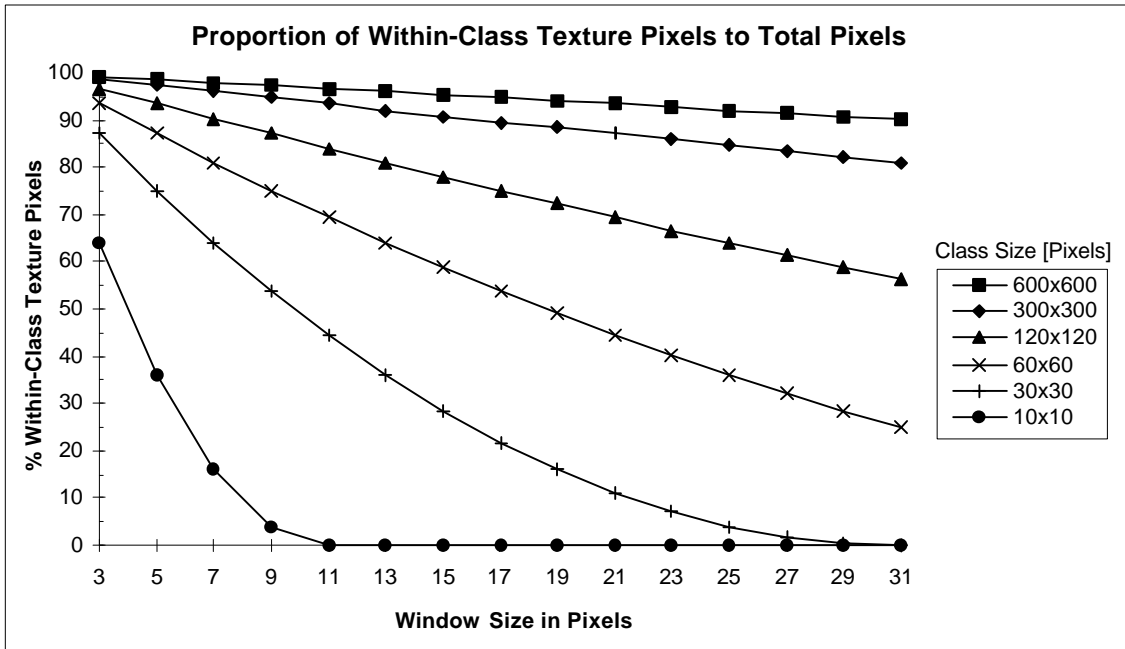
Figure 4: *Proportion of Within-Class Texture Pixels to Total Pixels.*

The effect of high between-class variance is illustrated in *Figures 5a and 5b,* which show a cross-section of a hypothetical image. In these figures, Class I is a relatively dark class. It has a smooth texture in that its pixels' values are relatively uniform. Class II is a relatively bright class, with highly variable pixel values and thus a coarse texture (*Figure 5a*).
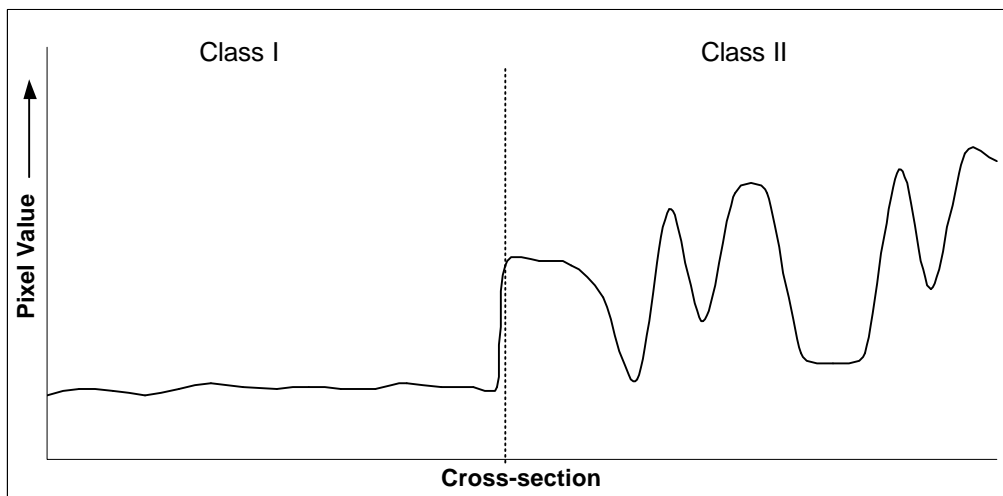
Figure 5a: *Cross-section pixel values at a class interface.*
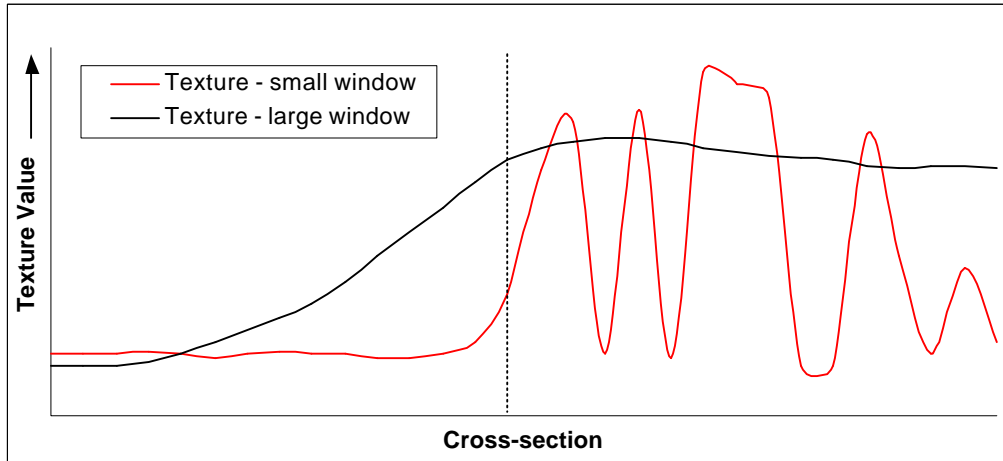
Figure 5b: *Cross section of texture values at a class interface.*

The curves on the graph in *Figure 5b* indicate a trade-off between different scales of window sizes used in calculating texture along the cross-section. With textures calculated over a small window the texture values of Class I remain low until right near the edge of Class II. This small window, however, produces a very unstable measure of within-class variability of Class II. Thus the texture measure will provide poor discrimination of Class II in that it has some texture values similar to those of Class I. With large window sizes the effect of the between-class variance is significant much further away. Many pixels in Class I are classified as the rough class, since some texture values are high inside the Class I region. The effect of the edges is not important for Class II, which has a relatively homogeneous calculated texture value.

By contrast the large window size gives a distinctive higher texture value for Class II. This comes at the cost of a broad "edge effect" for Class I, which would tend to cause those pixels to be mis-classified as Class II. It is interesting to note that there is no equivalent edge effect for Class II, because the within-class variability of this class is similar to that of the between-class variability of the two classes. In summary, when texture is used in a standard automated classification, the scale of the window size will determine the relative location of mis-classified pixels. Specifically, coarse classes have errors distributed throughout the class, but these will progressively be eliminated with larger window sizes. Little edge effect will be noticed. Smooth classes will give high accuracies for relatively small window sizes; however, even at the small window size there will be an edge effect, and thus errors will be most common on the edge of the class. As the window size increases, this edge effect will increase commeasurately.

There are two important implications from this discussion. Firstly, it is standard practice, though generally unstated, to avoid edges in selecting both training and testing pixels. Edges of classes may be poorly defined and pixels near edges are often contaminated by other classes; however, evaluations of classification using texture will significantly underestimate the errors, particularly for relatively smooth classes. Secondly, this theoretical framework leads to relatively precise definitions for the extremes of "smooth" and "coarse" textures. "Smooth" is where the within-class variability is much lower than the between-class variability. By contrast, "coarse" is where the within-class variability is similar to the between class variability.

IV.   *RESEARCH OBJECTIVES*

The previous chapter provided an extended discussion of how image texture arises from the spatial distribution of scene objects. It was shown that due to the hierarchy of scales in imagery, there are two types of texture – between-class and within-class. Consequently, the edges of classes are of particular importance for evaluating texture in digital image classification. This theoretical discussion provides the framework for the remaining part of this work, which consists of an empirical analysis of simulated and real image data. Simulated data was used because it consists of known and specified spectral and spatial parameters. The real data consists of ADAR aerial imagery of Morgantown, WV. Actual image data is often much more variable and of course the investigator does not have 100 percent *a priori* knowledge of the scene. The objective of the empirical analysis was to investigate the appropriate window size for texture analysis. Following the theoretical discussion of Chapter III, it was expected that the results would be class-dependent, and influenced by the spatial dimensions of class polygons. In order to facilitate multi-scale comparisons, window sizes were to be defined by real-world units (meters) rather than pixels. Although the main objective was to describe the relationships found, a subsidiary aim was to quantify the scale relationships, for example, through spectral class discrimination measures.

V. *METHODS*

The objective of investigating appropriate window sizes for texture analysis involves experiments on both simulated and real data sets with a wide range of window sizes. The variability of the resulting texture was plotted against window size. The texture measure was a variance filter algorithm on a pixel matrix at ten window sizes ranging from 3×3 pixels to 21×21 pixels. The variance measure is defined in *Table 1* and compared to other measures in Chapter II, Section A.

Two sets of data were used for this study: a simulated scene and a scene of Morgantown, WV. Both simulated and real data were resampled to give a total of six spatial resolutions

A. Simulated Data

The simulated data consisted of a scene model with two sub-images pasted onto a larger background image (*Figure 6*). This image was constructed with raster GIS software and given the attributes shown in *Table 2*. The texture of each sub-image was initially defined by random pixel generation with specified means and standard deviations. The classes and background were then re-classed so that each consisted of groups of three distinct digital number values corresponding to three distinct objects within each class.

Table 2: *Characteristics of the simulated data set.*

Class	Size of Objects Within the Class	Mean Pixel Value	Standard Deviation	Narrowest Dimension Of Class
Background	1 meter	158.9	4.7	600 meters
Class A	5 meter	162.4	28.3	600 meters
Class B	10 meter	106.4	5.3	600 meters
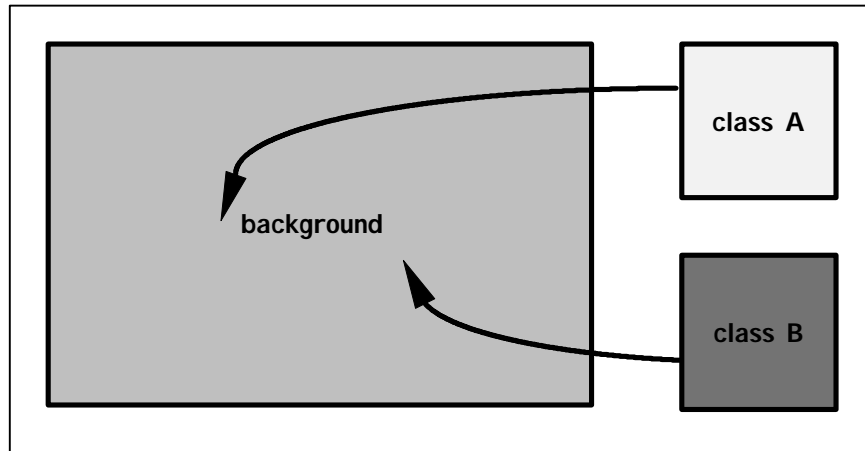
Figure 6: *The sub-images are pasted onto the background.*

The simulated scene was assumed to represent an image with a 1-meter pixel size. The scene model was then used to create a set of five additional simulated images by aggregation of pixels to coarser resolutions of 2, 5, 10, 20 and 30 meters as shown in *Figure 7* (Hay, *et al.*1997). A small random noise element with a mean of 15 and a standard deviation of 1 was added to each image.
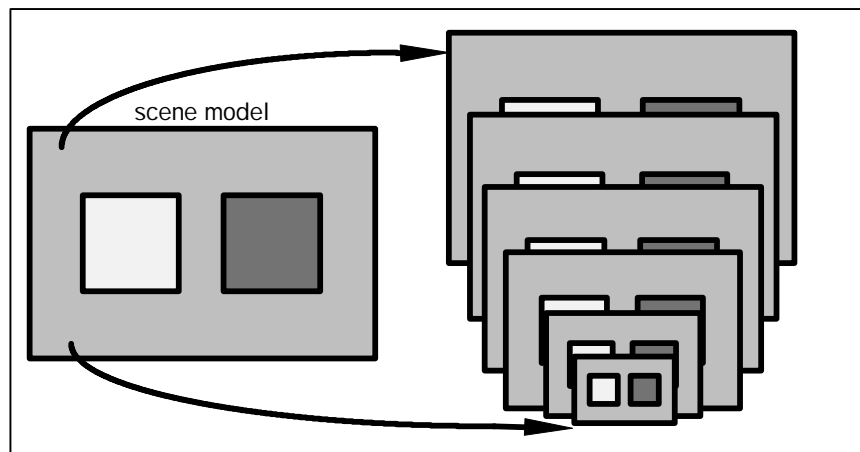
Figure 7: *The scene model is resampled.*

Since the object sizes of the background and the two classes were 1, 5 and 10 meters, respectively, this produced simulated images with elements ranging from H- through L-resolution. An example of the simulated 10-meter data is shown in *Figure 8*. Class A is the on the left and Class B is on the right.
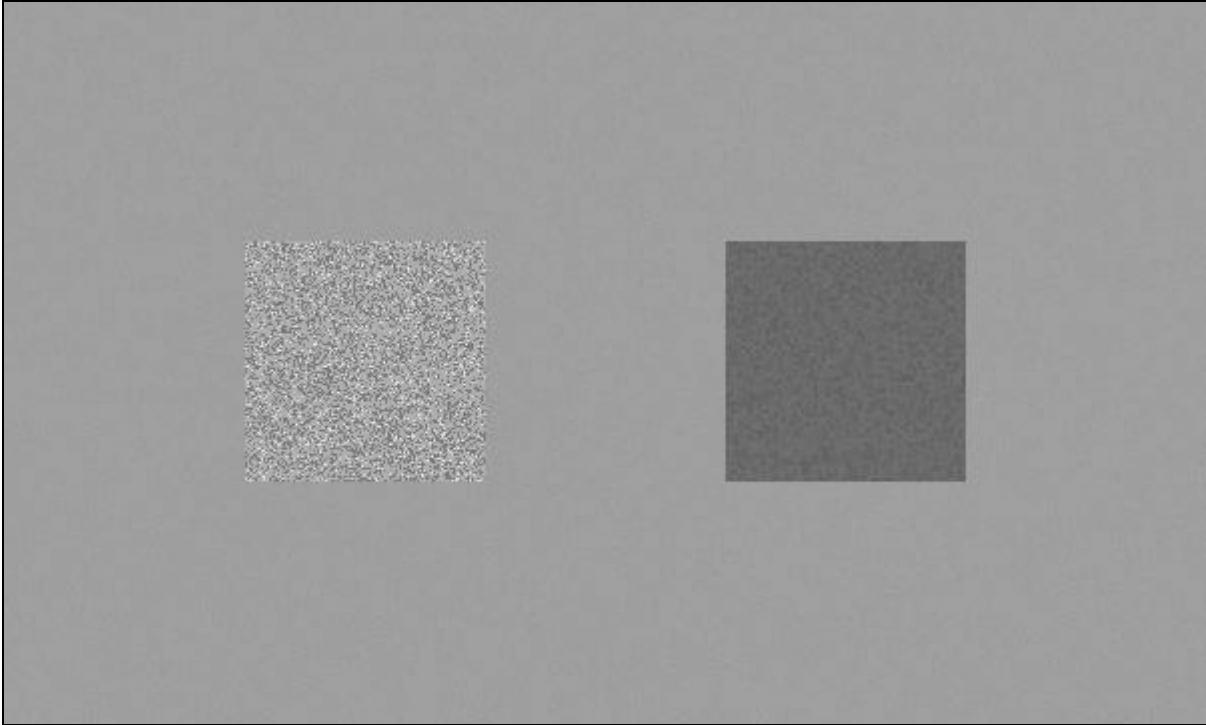
Figure 8: *Simulated 10-meter image.*

Texture analysis was performed on each re-sampled image at ten window sizes. The window sizes used in this study were 3×3 to 21×21 pixels. The entire process is summarized in *Figure 9.*
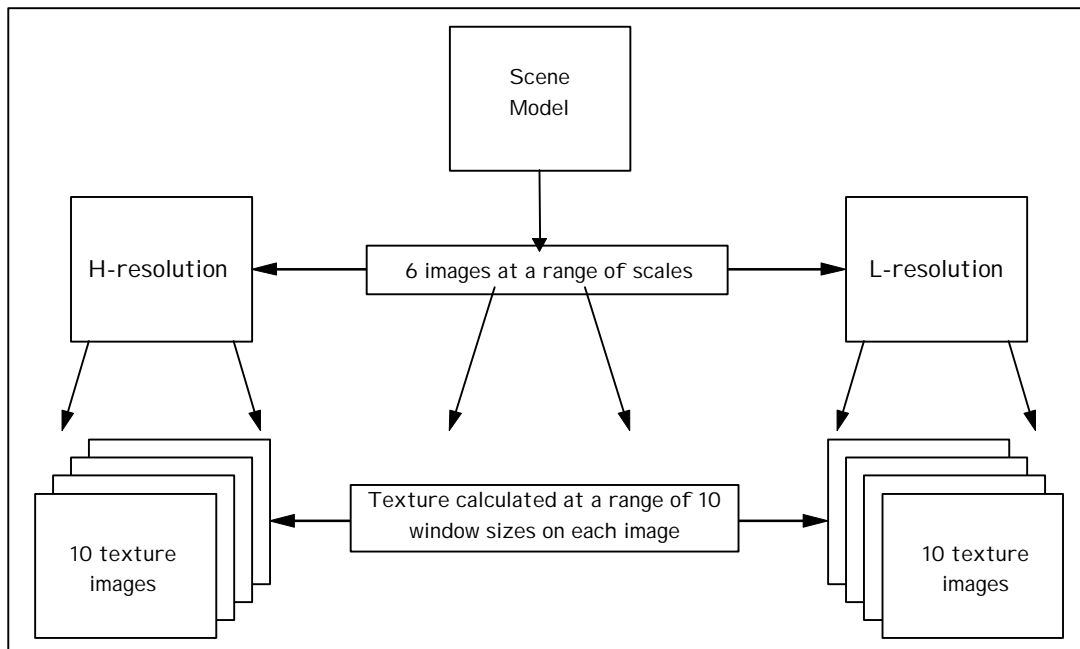
Figure 9: *Simulated data experiment.*

The variability of the scene texture and class texture was plotted against window size for each re-sampled image on one graph. Furthermore, to illustrate the effects of pixels that incorporate between-class variance, variability of these pixels were compared to pixels that were not effected by the between-class variance. To facilitate the application of these experiments to actual image data, separation between classes was expressed by a general separability index. The Jefferies-Matusita (J-M) Distance is one such index that takes into account both the measure of distance between class means and the differences between the covariance matrices of the classes (Richards, 1993; Warner and Shank, 1997). For each texture image the J-M Distance was calculated and plotted for selected classes.

B. Real Data

To demonstrate the practical value of these concepts the approach was tested on actual remotely sensed image data. This second data set consists of 1-meter ADAR data of the Morgantown, WV area.

The ADAR 5500 digital camera system was used to collect the data used for this thesis. ADAR images are captured directly with four digital cameras (Stow, *et al.*, 1996). The charge coupled device (CCD) on each camera produces an image of 1500 by 1000 pixels. ADAR image data can be acquired from any fixed wing aircraft with a conventional aerial camera mount.

The ADAR data were captured on four dates in late March and early April 1997. The sensor characteristics are listed in *Table 3*. Flight altitude was approximately 8,600 feet (2,620 meters) above mean seal level. This altitude resulted in a spatial resolution of approximately 1-meter. Each frame covered a rectangular area of approximately 1.5 by 1.0 kilometers. Image brightness is controlled in flight by adjusting the shutter speed of each camera (Stow *et al.*, 1996).

Table 3: *The ADAR sensor characteristics.*

Band	Color	Bandwidth (μm)	Aperture	Exposure (sec.)
1	Blue	0.45 – 0.54	f 2.8	1/500
2	Green	0.52 – 0.60	f 2.8	1/800
3	Red	0.61 – 0.69	f 4.0	1/640
4	Near Infrared	0.78 – 1.00	f 5.6	1/2500

A GPS receiver simultaneously captured aircraft latitude, longitude and elevation data with each image frame. The coordinates were projected to UTM Zone 17 using NAD27 datum. Each waveband image was automatically registered to create a four band digital frame (Stow, *et al.*, 1996). Multiple frames were digitally mosaicked to create a coverage equivalent to the United States Geologic Survey (USGS) 7½-minute Morgantown North topographic quadrangle. *Figure 10* is a false-color composite of the ADAR data for the Morgantown, WV area.
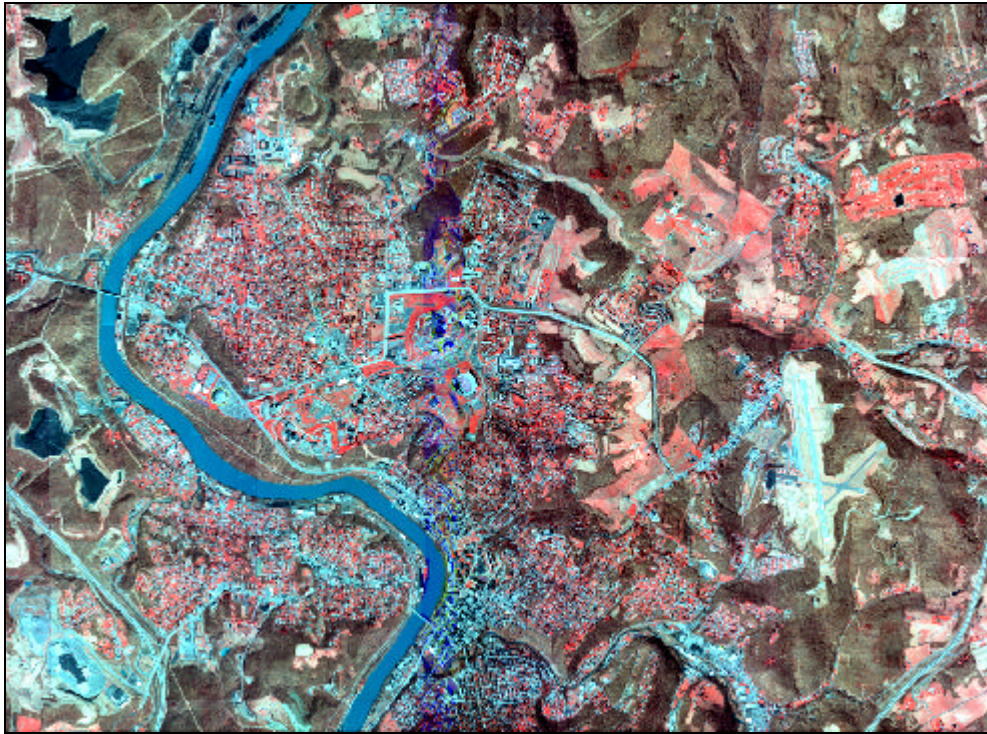
Figure 10: *False-color composite of the Morgantown, WV area.*

Errors in band registration and sensor performance necessitated the exclusion of band 1 (blue) from the analysis without significantly deleterious effects to the analysis. This band is susceptible to atmospheric haze and is often highly correlated to the two other visible bands (Stow, *et al.,* 1996). It therefore was not essential for the analysis. Edgematching errors also existed, but they did not pose a significant problem for the image processing undertaken in this project. Illumination differences between panels in the mosaic were also present due to atmospheric differences and time of image acquisition. The images for the mosaic were taken over a period of time in which trees were beginning to show buds and agricultural fields were being modified. There were subtle texture differences between panels in the mosaic, which had budding trees and those that did not. The texture differences did not significantly affect the classification or separability analyses.

For texture calculation band 2 (green) was chosen because of its relatively superior spectral response and contrast properties. Also, differences in shadows, because of time of day and the illumination differences between panels in the mosaic are minimized at shorter wavelengths. Pilot tests of texture feature extraction and visual analysis showed that band 2 was least affected by the illumination differences between panels. The data was then resampled to five spatial resolutions ranging from 2 meters to 30 meters. This yielded a total of six images at 1, 2, 5, 10, 20, and 30-meter resolutions. The vicinity of Morgantown includes land cover classes that exhibit a variety of textures and scales:

- Water
- Agriculture/Grass
- Barren/Bare Soil
- Forest

- Low Density Residential
- High Density Residential
- Commercial/Industrial
- Surface Mining

Texture was calculated on each image at ten window sizes ranging from 3×3 to 21×21 pixels. The variability of image texture was plotted against window size for all images on one graph.

In addition to the calculation of J-M distances for evaluating class separability these data combinations were classified using Maximum Likelihood. After image classification an error analysis was performed on each classification to test for accuracy against ground reference data. Training data were extracted from both the interior of classes, away from the edges, and entire class polygons that therefore included the pixels adjacent to class boundaries.

C. Ground Reference Data

Ground reference data included a National High Altitude Photography (NHAP) program color infrared photograph enlarged to a scale of 1:14,500, *a priori* knowledge of local landscape and land cover features.

VI.   EXPERIMENTAL RESULTS AND DISCUSSION

A. SIMULATED DATA

1.   Texture Variability

Many studies that have evaluated the usefulness of texture for image classification have done so on an image basis. *Figure 11* demonstrates the change in texture variability for the entire simulated image. The window size in pixels has been converted to meters so that window sizes can be compared between image resolutions on the same graph. *Figure 11* shows that as data is aggregated to a coarser spatial resolution the texture variability generally decreases for a window of a specified size in real world units (meters).   Exceptions to this are the 5-meter resolution data and, to a certain extent, the 30-meter resolution data.   The anomalous data for the 5-meter resolution data may result from the fact that this window size corresponds with that of class elements in the simulated scene model.
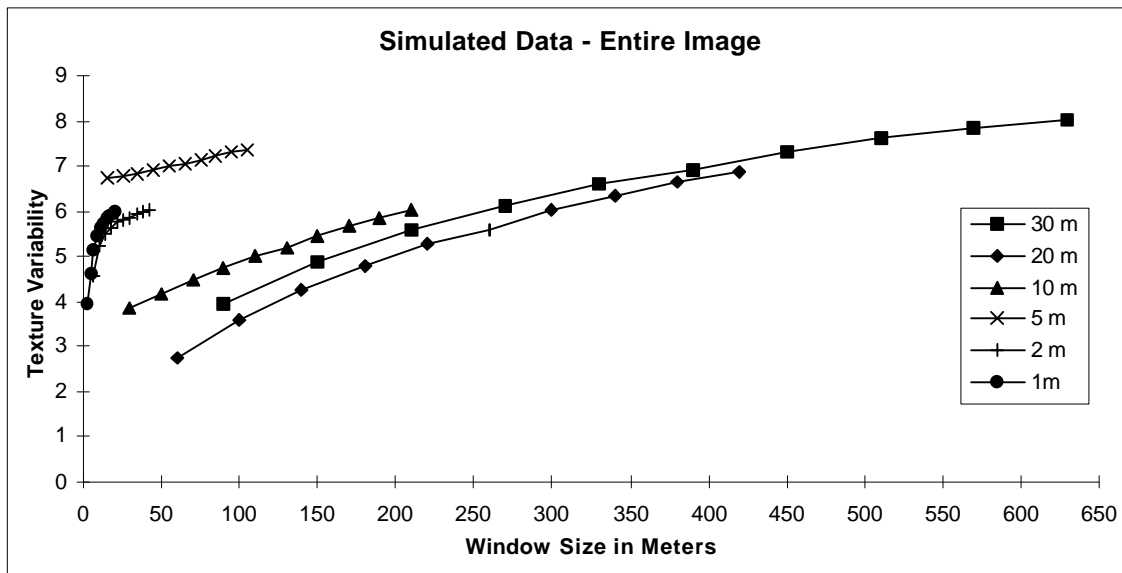
Figure 11: *Texture variability for the entire simulated scene.*

Texture variability tends to increase with an increase in window size because a greater proportion of the texture calculation window includes between-class texture. The effects of the between-class texture can be seen in the images of *Figure 12*, which shows examples of texture images of the simulated data. The image on the right is 15-meter texture data derived from 5-meter data with a 3×3 pixel window while the image on the left is 105-meter texture derived from 5-meter data with a 21×21pixel window. The edge effect can clearly be seen in the right-hand image, particularly for Class B.
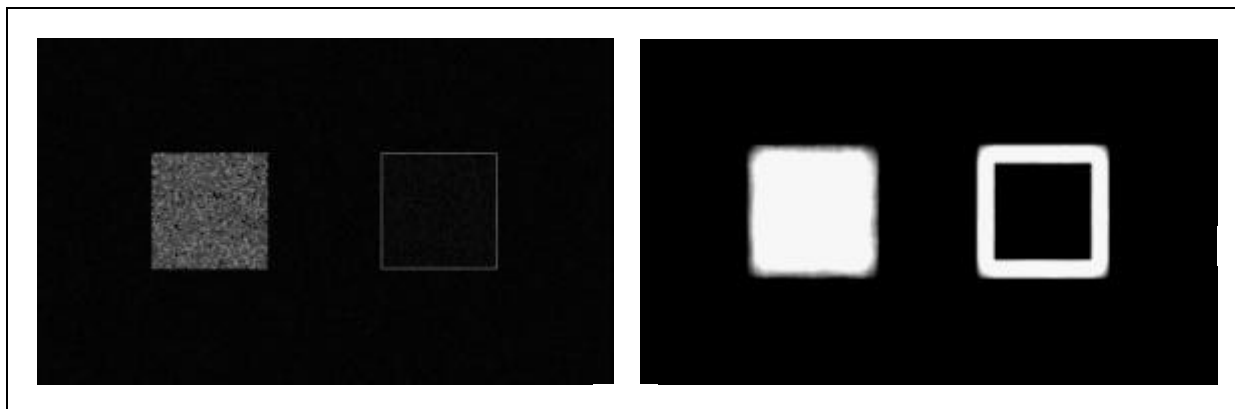
Figure 12: *Sample texture images.*

The graph in *Figure 11* might be interpreted to imply that the smallest possible window size is preferable because it produces the lowest variability in texture. However, since the texture image is not uniform across the scene, or even within a single class, plots of texture variability against window size for each class individually provide more useful information than averages for the entire scene.

The graph for Class A (*Figure 13*) shows a pattern very different from that of the entire scene. Variability declines rapidly as window size increases, leveling off in the 50 to 150 meter range. Again, the variability for the 5-meter resolution data is higher than the other resolutions.
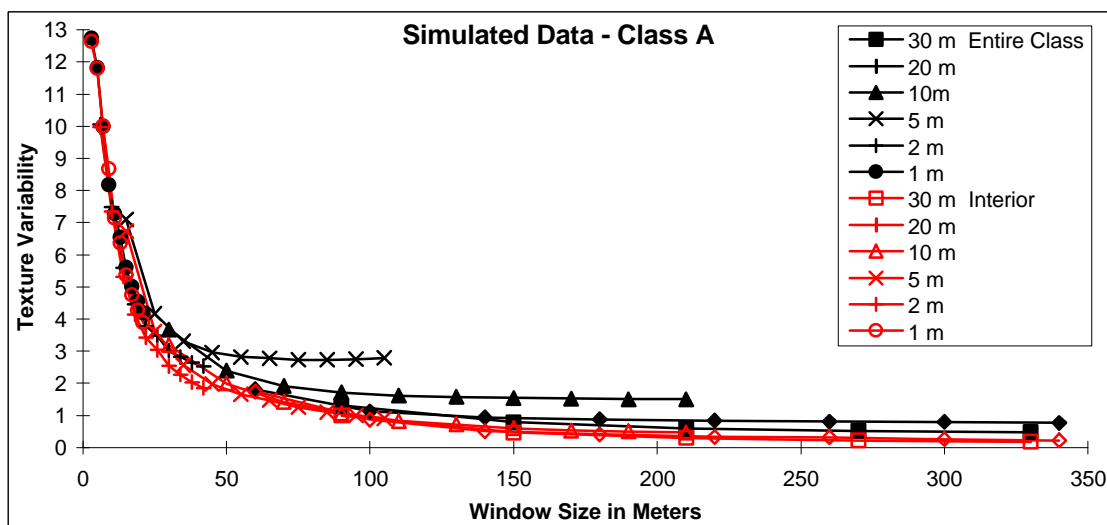
Figure 13: *Texture variability of Class A.*

Class A is spectrally similar to the background, thus the between class-variance is similar to the within-class variance. This is shown in the graph by the similarity in the curves for texture variability of the interior of the class and texture variability for the entire class, which

includes between-class variance. The influence of the pixels in the background in creating an edge effect is negligible in this case. The initial high texture variability at relatively small window sizes occurs when the elements in the class are transitioning from H-resolution to L-resolution. When the curves reach the texture variability minimum the elements are at L-resolution. The class remains at H-resolution throughout the graph.

Compared to Class A, Class B has larger within-class objects (10 meters versus 5 meters), but a much smaller spectral variability (standard deviation of 5.3 versus 28.3). Furthermore, Class B has a mean DN value that contrasts strongly with that of the background. Consequently, the texture variability for the interior of Class B is very different from the variability for the entire class. In the graph of Class B (*Figure 14*) variability of the texture for the entire class (including edges) increases due to the high proportion of pixels of between-class texture dominated pixels as window size increases. At the largest window sizes there is a slight decline as texture becomes increasingly dominated by between-class texture. Without the effect of the high-variance between-class texture pixels, the texture variability for pixels from the class interior is very low. The curve shows a decline as window size increases and begins to flatten out after the 50-meter scale. Variability continues to slowly decline but does not reach a minimum even at the largest scale that was evaluated (350 meters). The extremely low variability is only achieved at relatively large window sizes where a significant percentage of the class polygon is influenced by between-class texture.
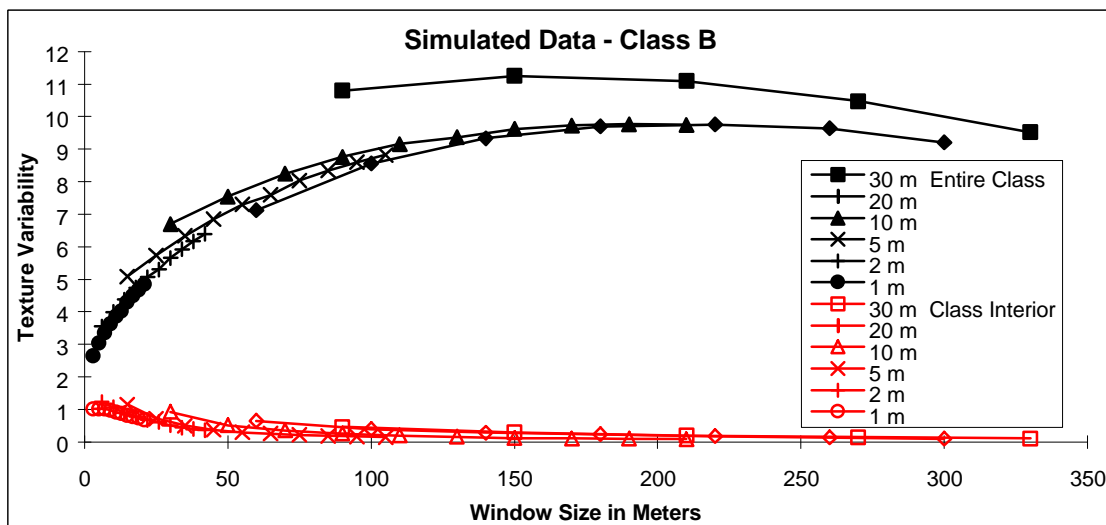
Figure 14: *Texture variability of Class B.*

The results on this graph are of vital importance to investigators who wish to use texture-derived data in classification schemes. The effect of window size is class dependent.

The value of the texture variability appears to remain fairly constant from image resolution to resolution; thus the choice of resolution does not affect the value at which the texture variability minimum occurs. While an evaluation of pixels in the interior of the class might indicate a useful scale for calculating texture, the effect of the surrounding classes is large. Such a significant portion of the class polygon is affected by between-class variance that the incorporation of texture data might not produce fruitful results in classification. This also suggests that evaluations of classifier accuracy should always include data from class edges, otherwise an overly optimistic result may be obtained.

2. Class Separability

The results to this point illustrate the trade-offs associated with window size. However, in order to quantify how these effects influence classification, a separability index is needed which can relate to classification accuracy. The Jefferies-Matusita distance was used to calculate the general separability of classes with combination of spectral and textural data. The range of J-M distances is from 0, which indicates no separability, to 1.414, which indicates complete separability (Richards, 1993). The J-M distances were plotted for simulated 10-meter data and the texture calculated on it at ten window sizes (*Figures 15* and *16*). Ten-meter texture was chosen because, as is seen in *Figures 13* and *14*, this provides a good range in spatial properties for these images.

Figure 15 shows that texture in addition to the single simulated 10-meter band greatly enhances the separability of the two classes. Without the confusing effect of between-class variance pixels, the two classes are completely separable with the addition of texture at all scales. While the inclusion of 'edge effected' pixels results in a less significant increase in separability, there is nonetheless a marked improvement over the spectral data alone. Larger window sizes result in an improved accuracy over the range of 30 to 110 meters. Accuracy did not change significantly for window sizes greater than 100 meters.
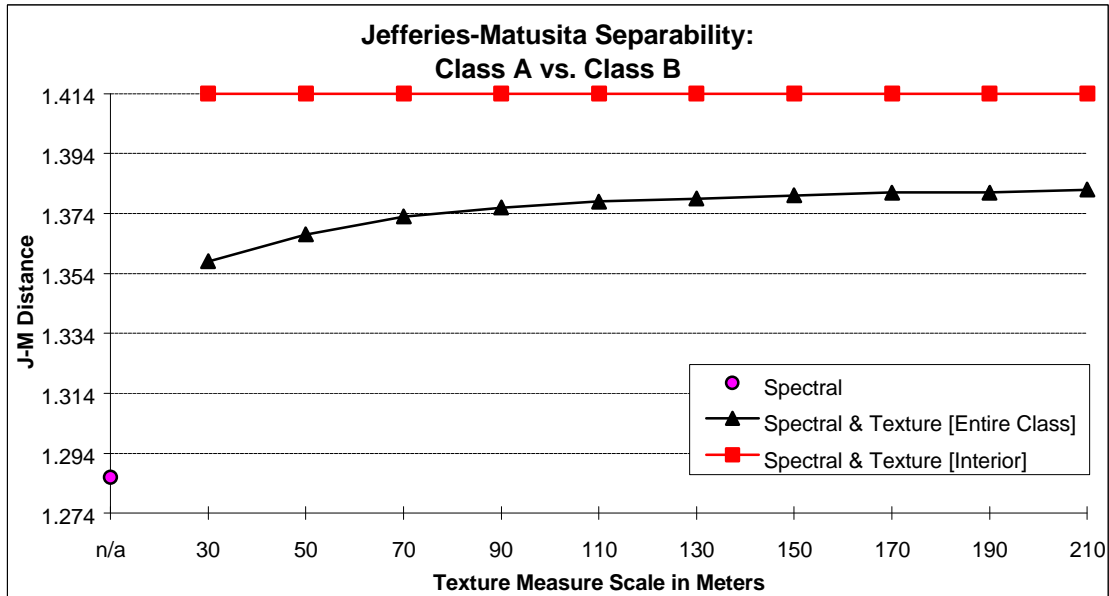
Figure 15: *Jefferies-Matusita Separability for Class A vs. Class B.*

The texture variability minimum for Class A, as seen in *Figure 13* occurs at approximately 50 to 200 meters in scale. The texture variability for Class B, as shown in *Figure 14*, drops sharply until about the 50-meter scale and beings to flatten out after 150 meters. However, window sizes in the range of 150 meters for these images are not practical since there are so few interior pixels. Furthermore, it is interesting that these window sizes are 5 to 40 times the size objects in the classes. These figures are large because the texture window must span the spacing of elements to capture texture effectively. The elements are of irregular spacing and randomly distributed within the classes; thus a large window is necessary to cover a large enough sample of elements.

The values shown on the graph in *Figure 16* suggest an optimal texture scale of approximately 50 meters in combination with the 10-meter simulated data in terms of class separability. When the separability of Class A versus the background is plotted, texture again provides a valuable discriminant, increasing separability from 0.880 to 1.250. As window size increases further, separability falls steadily. If only interior pixels are evaluated, Class A and the background are completely separable over the scale range of 50 to 170 meters. Whether the entire class or only interior pixels are included in the analysis, the addition of textural data greatly enhances the separability of the classes. The J-M distances between the background and Class A show an increase in separability from texture scales of 30 to 50 meters.
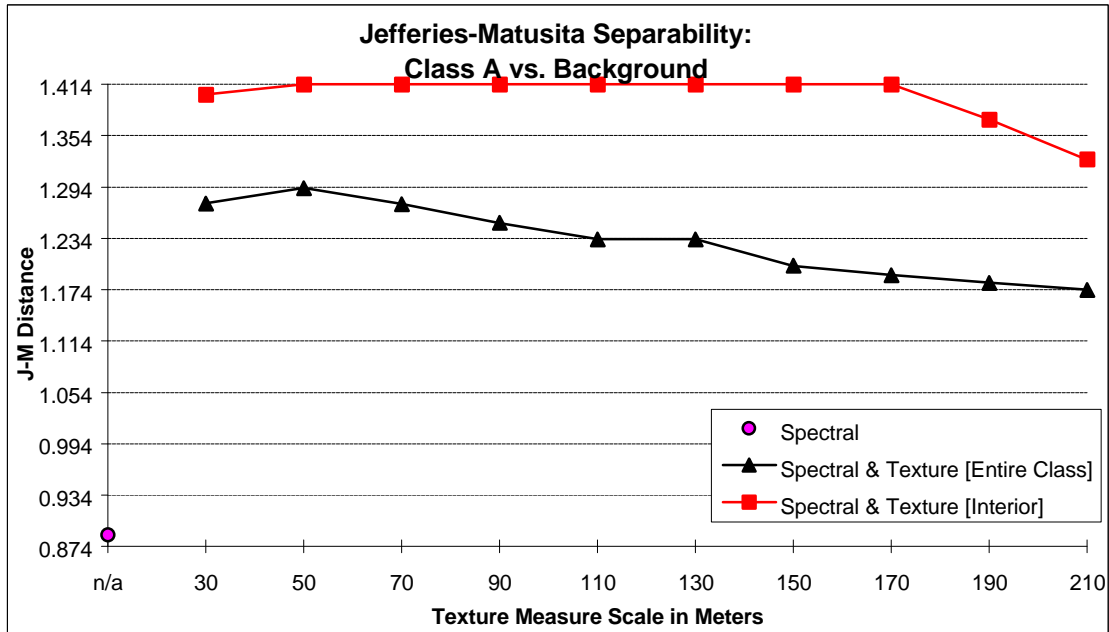
Figure 16: *Jefferies-Matusita Separability for Class A vs. background.*

The J-M distance between Class B and the background was 1.414 in all instances, meaning that with or without the inclusion of the textural data they are completely separable.

The analysis so far has focused on simulated data, with simple, known spatial parameters. In the next section the evaluations of scale and texture are applied to real data to evaluate how well these methods and results can be applied to real data.

B.   REAL DATA

1.   Texture Variability

The graphs of texture variability against window size for the Morgantown, WV data generally follow the pattern shown by the simulated data graphs (*Figures 13* and *14*). For the Central Business District (CBD), which is an example of the Commercial/Industrial class, shown in *Figure 17*, texture variability is very similar for the entire class and the interior of the class for all window sizes (*Figure 18*). The differences in collection dates for the mosaic panels had little effect on the Commercial/Industrial classes because the objects found in those classes are generally unaffected by seasonal variation.
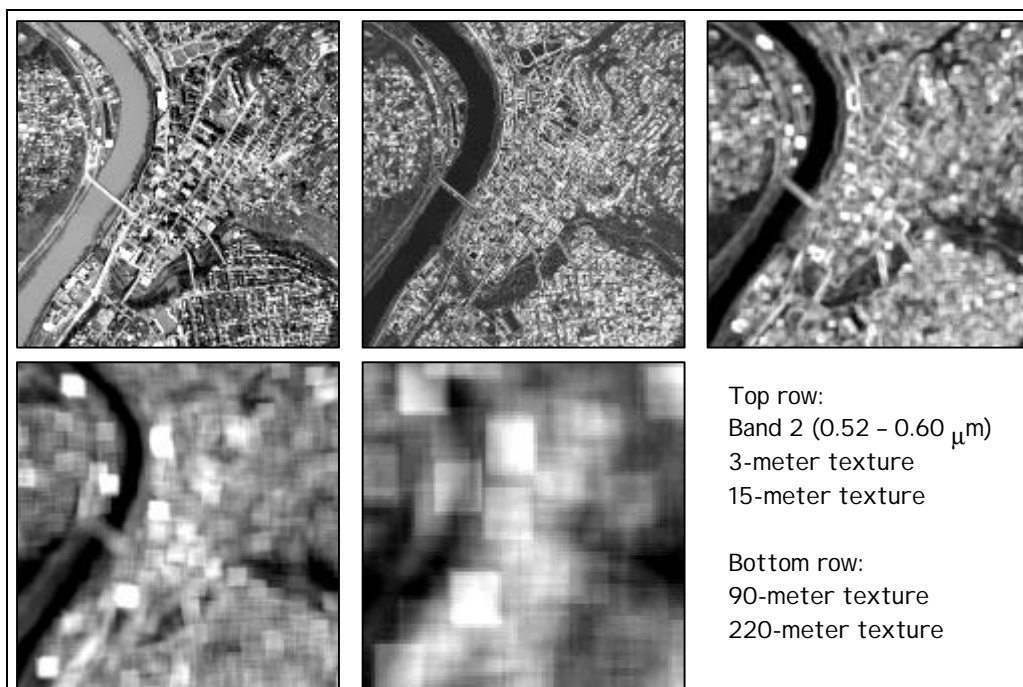
Figure 17: *Sub-images of the CBD.*
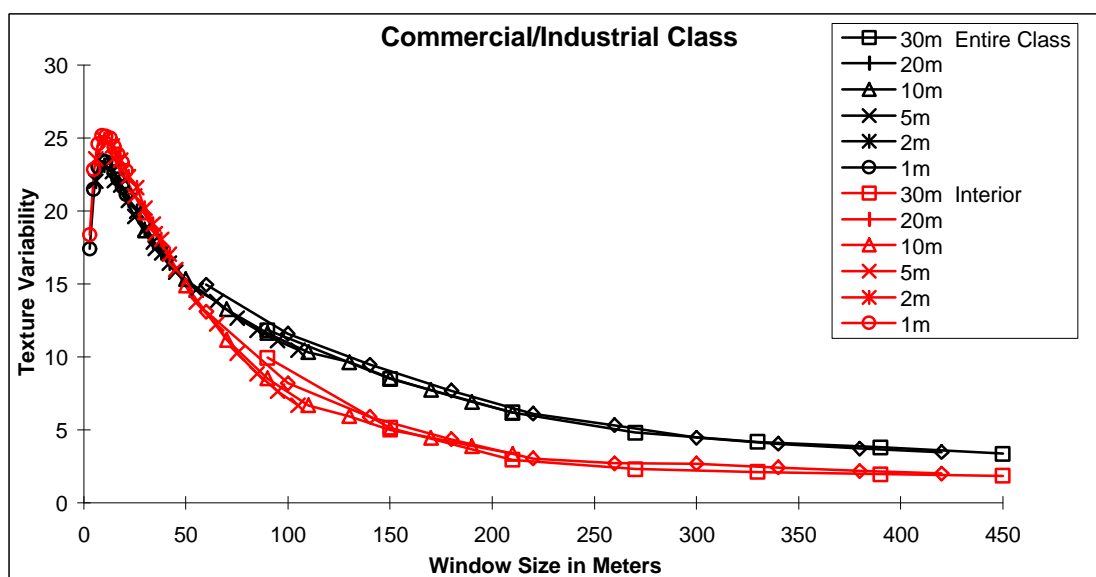
Figure 18: *Texture variability for Commercial/Industrial class (CBD), Morgantown, WV.*

The results of *Figure 18* matched well with the trends found with Class A in the simulated data experiment (*Figure 13*), except with the CBD there is a pronounced increase in variability of texture with increasing window size over the smallest window sizes. Spectral differences between the (CBD) and the surrounding Residential class are similar to the spectral differences found within the CBD; thus, the graphs for the entire class and the interior plot closely together. The effect of between-class variance texture pixels is therefore small.

Texture variability seems to reach a minimum level near the 400-meter scale. The relatively low texture variability minimum at the 3 meter scale occurs because the texture measure widow is falling within the elements of the next finest scale in the hierarchy (in this case, the individual buildings and urban structures in the CBD). The elements within the CBD are therefore also at H-resolution at this texture scale and scale of capture. This is important because although the average texture variability is quite low, the texture data gathered at that scale might not be useful for classification of elements at the next level of the hierarchy. The peak in texture variability occurs at the 15-meter scale. This is due to the texture window consistently falling on the edges of objects in the next finer class in the hierarchy, in this case buildings, roads, and other urban structures. This is similar to Woodcock and Strahler's (1987) finding that when pixel size reached one-third to three-quarters the size of the scene objects, a peak in local variance occurred.

Texture analysis on the Agricultural/Grass (*Figure 19*) class yielded some important and interesting results (*Figure 20*). The texture variability of the interior pixels is much different from the variability of the entire class (including edges). The peak for texture variability occurs at approximately 15 meters for the interior pixels and at approximately 45 meters for the entire class.

Top row:
Band 2 (0.52 – 0.60 μm)
3-meter texture
15-meter texture

Bottom row:
45-meter texture
90-meter texture
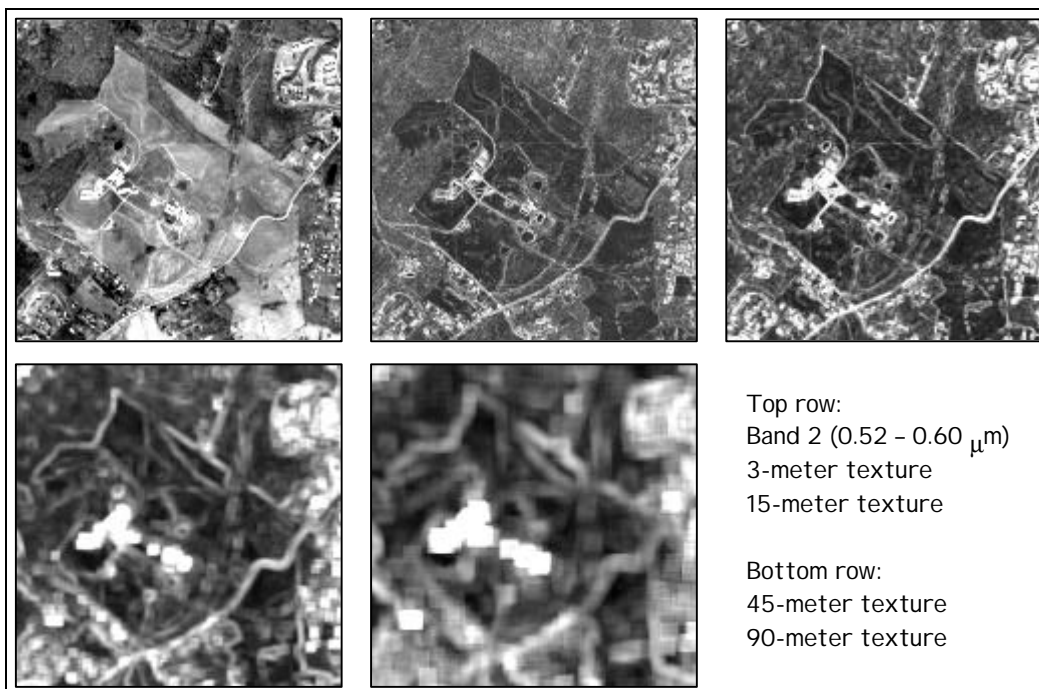
Figure 19: *Sub-images of the Agricultural/Grass class.*
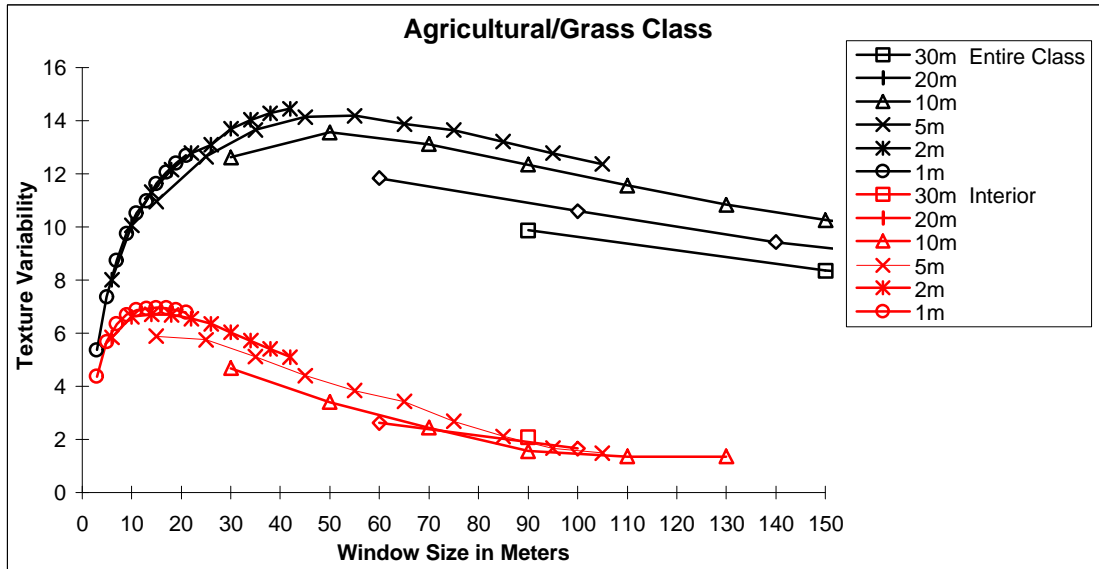
Figure 20: *Texture variability for agricultural/grass class, Morgantown, WV.*

As with the CBD, at the 3-meter texture measure scale a level of low texture variability occurs. Thus, at this scale, the texture elements for this class are also H-resolution. This scale includes local variation of the land cover, such as soil and topographic variations, and accounts for the low variability observed. For the interior pixels the elements at the finer scale cause the graph to peak at approximately 15 meters, as the elements are transitioning from H-resolution to L-resolution. The peak at 45 meters for the entire class is probably related to the changing properties of interior and edge pixels. It seems that for this particular agriculture/grass class polygon, though the elements are small, surprisingly large window sizes are needed to capture the appropriate texture scale. The texture variability minimum for the interior pixels occurs near the 90-meter scale.

2. Class Separability

J-M separability graphs for the real data differ from those of the simulated data in that with the inclusion of the 'edge-effected' pixels, separability was less than that of the spectral data alone. Forest and Commercial/Industrial classes are completely separable without the inclusion of texture data (*Figure 21*) due to their highly different spectral characteristics. The inclusion of texture results in only a very small degradation in separability in this instance.
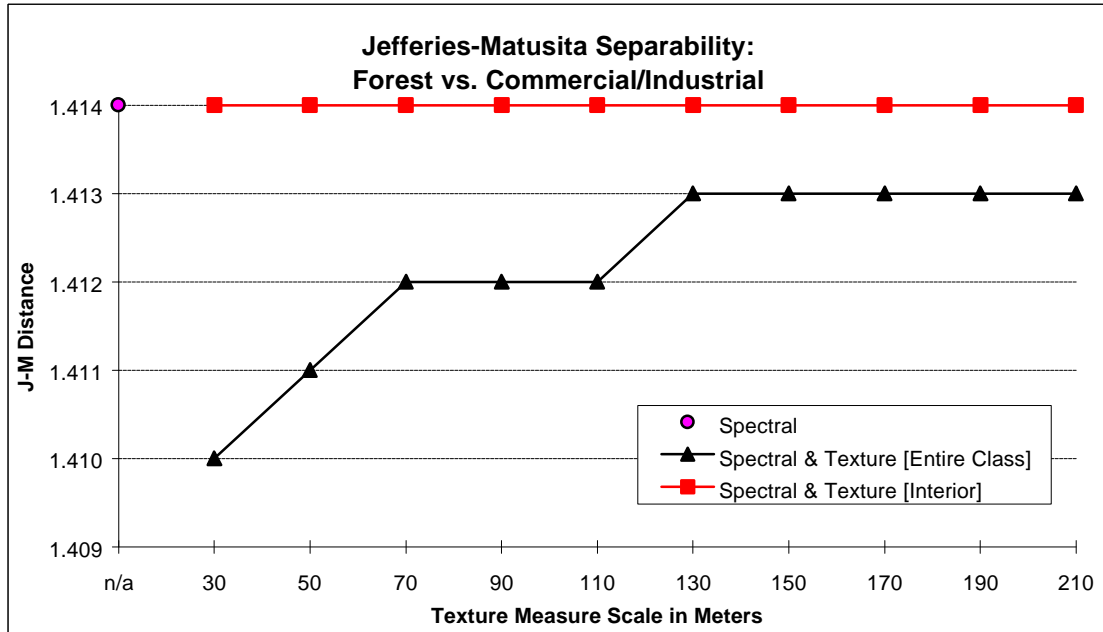
Figure 21: *Jefferies-Matusita separability of Forest vs. Commercial/Industrial classes.*

Potentially confusing spectral classes such as "Residential" and "Commercial/Industrial" (*Figure 22*) which both include roads, driveways, parking areas, buildings, and vegetation, are generally found to be more separable with the inclusion of texture when evaluated based on pixels from the interior of classes only. In the analysis, increasing the size of the window produced a near linear increase in accuracy over the 30-meter to 210-meter range studied for class interiors and entire classes. However, even at the largest window sizes, the accuracy of the entire class only improves to approximately the same level as can be obtained from the spectral data alone. When pixels that incorporate between-class texture are included separability is relatively low.
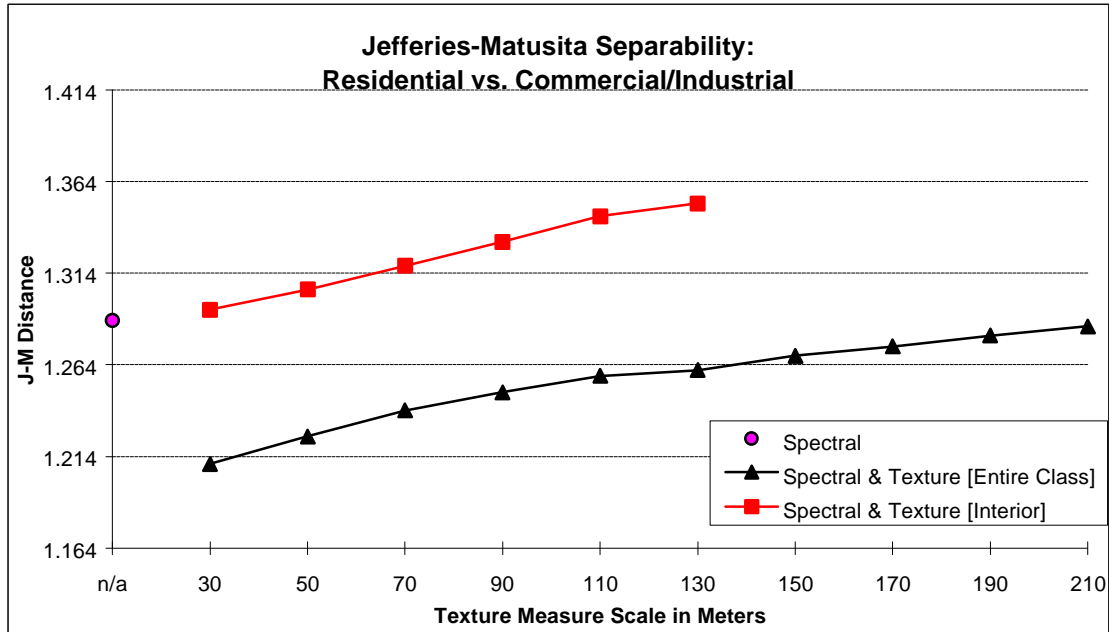
Figure 22: *Jefferies-Matusita separability of Residential vs. Commercial/Industrial.*

An even stronger example of the potentially deleterious effects of texture is illustrated by the 'Residential' and 'Agricultural/Grass' class separability graph (*Figure 23*). Without textural data, a J-M distance of 1.288 is obtained; however, the excluding the problematic between-class variance pixels on the edges of classes, separability increases immediately to 1.360 and continues to improve to total separability of 1.414. These classes are in fact less separable because of the between-class variance pixels in the textural data. Thus for separability evaluated on the entire class, the J-M distances never exceed 1.114, even at the largest scale.
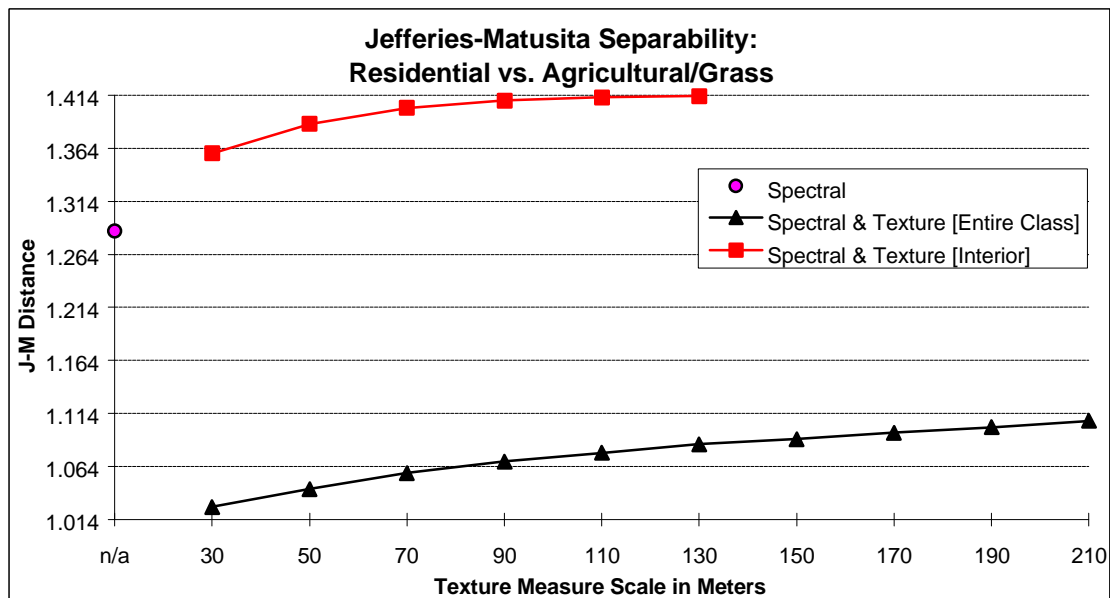
Figure 23: *Jefferies-Matusita separability for Residential vs. Agriculture/Grass.*

3. Classification Results

Classification was performed to provide a more comprehensive evaluation of the effect of scale on texture for this data. Six combinations of spectral and textural data were used. Pixels were assigned to one of six classes using a maximum likelihood classifier (*Table 4*).

Table 4: *Classes and their representative land cover types.*

Class	Range of cover types
Water (WTR)	Rivers, ponds
Forest (FOR)	Deciduous forest, evergreen forest
Agricultural/Grass (AGR)	Crops, pasture, fields
Residential (RES)	Homes, lawns, trees, roads, trailer parks
Commercial/Industrial (COM)	Businesses, factories, roads, parking lots
Surface Mining Related (MIN)	Strip mines, coal piles/storage

The test reference data was compared to the classified images on a pixel-by-pixel basis. Agreement and disagreement were summarized in the cells of error matrices (Jensen, 1996). An example of an error matrix for test pixels is shown in *Table 5*. Overall accuracy is computed by dividing the number of correct pixels by the total number of pixels in the error matrix. Classification accuracy was evaluated using the Kappa statistic. The Kappa statistic is a measure of the difference between the observed accuracy and the random possibility of

chance agreement between the reference data and the classification (Lillesand and Kiefer, 1994). Conceptually, Kappa can be defined as:

$$\hat{k} = \frac{observed\ accuracy\ -\ chance\ agreement}{1\ -\ chance\ agreement}$$

Table 5: *Example of an error matrix.*

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	19553	0	0	0	0	0	19553	0.0000
FOR	0	87905	518	3765	165	0	92353	0.0482
AGR	0	1647	26655	13148	68	6	41524	0.3581
RES	707	7487	6025	56622	10199	99	81139	0.3022
COM	63	698	32	15893	66884	465	84035	0.2041
MIN	12	99	0	178	2331	16647	19267	0.1360
Total	20335	97836	33230	89606	79647	17217	337871	
Omission	0.0385	0.1015	0.1979	0.3681	0.1602	0.0331		0.1883

Image classification of various spectral and textural combinations were evaluated and graphed. The classification was performed on combinations of 2, 5, and 10-meter spectral data, because they fall within the range of resolutions used by current remote sensing platforms, and a range of texture scales from 10 meters through 220 meters. These were also the scales used in the spectral separability evaluations.

For the overall accuracy, it was found that if the accuracy is evaluated with pixels from the edge of classes that are influenced by between-class variance, there is a general degradation of classification accuracy (*Figure 24*).
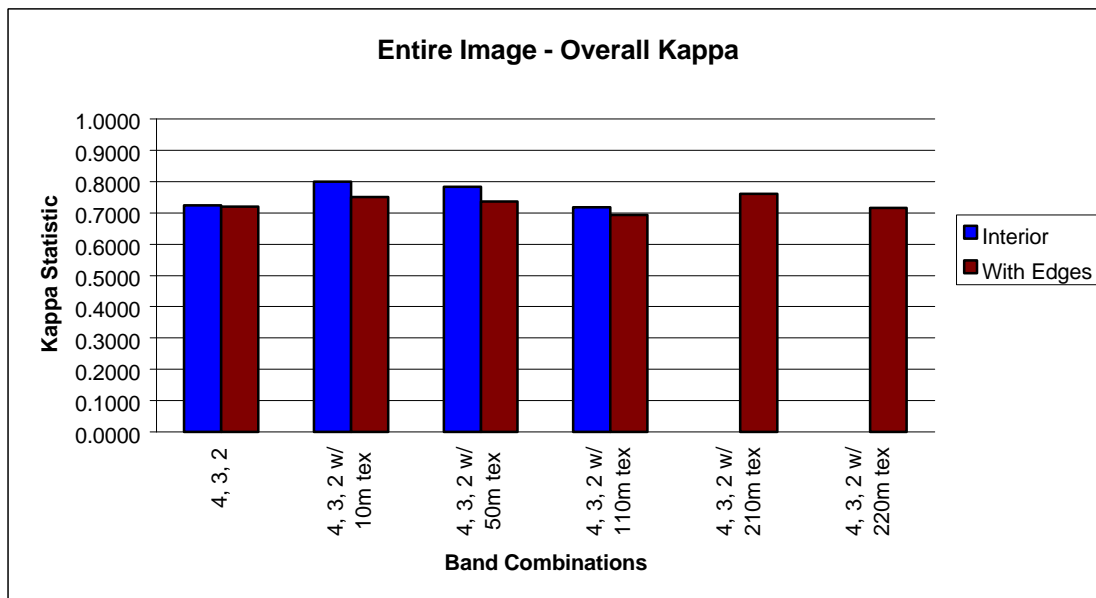
Figure 24: *Comparison of Kappa statistics on an image basis.*

Kappa increases significantly (at the 99% confidence level) with the addition of texture information. In general band combinations which included texture resulted in higher classification accuracies than those which did not. The exception to this is band combination 4, 3, 2 and 110-meter texture. The slight reduction in classification accuracy at the 110-meter scale can be accounted for by the multiple texture scales present in the image. Any class that has a low accuracy at the 110-meter scale influences the overall accuracy at that scale. As the texture scale changes, the effects of the variability of the six classes have different effects on the classification results. The greatest level of accuracy including pixels near the edges of classes was found with bands 4, 3 and 2 with 210-meter texture data. Classes that achieved their highest accuracy at the 210-meter scale (Residential and Commercial/Industrial) comprise a significant proportion of the total pixels in the image. However, an increase in just ten meters in window size resulted in a decrease in Kappa to 0.7895 for class interiors and 0.7151 for interior and edge pixels. It is interesting to note that this scale is twice that of the scale at which an accuracy minimum was found (110 meters). This decrease in accuracy at 220 meters suggests that the scale at which texture is calculated should be chosen very carefully, as even with large windows a small change in window size can result in large differences in classification accuracy.

Though useful for understanding how texture scale relates to the error assessment of an entire image, a comparison of selected classes at various texture scales should provide added insight when determining optimal texture scales for image classification. Computing the

accuracy of individual classes is more complex than that of an entire image and involves two types of accuracy: *producer's accuracy* and *user's accuracy*. When the total number of correctly classified pixels in a class is divided by the total number of pixels that should have been classified in that class, it is known as producer's accuracy (Jensen, 1996). This term is given because the producer of the classification is interested in how well a certain area can be classified. The producer's accuracy is related to omission error by the equation 1 - producer's accuracy. The *omission error* indicates how often pixels were left out of the class of interest. If the total number of correctly classified pixels in a class is divided by the total number of pixels that were actually classified in that class, (both incorrectly and correctly), the result is a measure of *user's accuracy*. User's accuracy is also related to *commission error* by the equation 1 - user's accuracy. Commission error is an indication of how often pixels were assigned to a class other than the correct one, as determined by the reference data (Jensen, 1996).

4. User's Accuracy

As suggested by the texture variability plots, classes that have a between-class variance higher than the within-class variance, and thus a relatively smooth within-class texture, should show degradation in classification accuracy when the test pixels include areas near the edges of the class. The class interiors, consisting of pixels unaffected by between-class variance shrink as the window size increases. The rate at which this occurs is a function of the dimensions of the class polygon (*Figure 4*). For the Morgantown ADAR data at window sizes greater than 110 meters there were an insufficient number of interior pixels to evaluate classification accuracy for pixels unaffected by between-class variance because the texture calculation window approached the size of class the polygons. *Figure 25* is a graph of user's accuracy for the Agricultural/Grass class for both interior pixels and pixels that include the edges of the test class. User's accuracy increases from the 10-meter to the 50-meter texture scales and then decreases at 110 meters. For the interior of the class this is somewhat surprising given the texture variability minimum at 90 meters in *Figure 18*. The accuracy then increases at 210 meters before declining again at 220 meters. The scale of maximum classification accuracy is 210 meters. This scale is large, considering the relative smoothness of this land cover type and the size of the class dimensions. At these large scales most of the measured texture arises from the between-class variance. The pattern of peaks and troughs in

this graph might be an indication of the complexity of multiple scales in real land cover classes and the changing effects of within- and between-class variance.
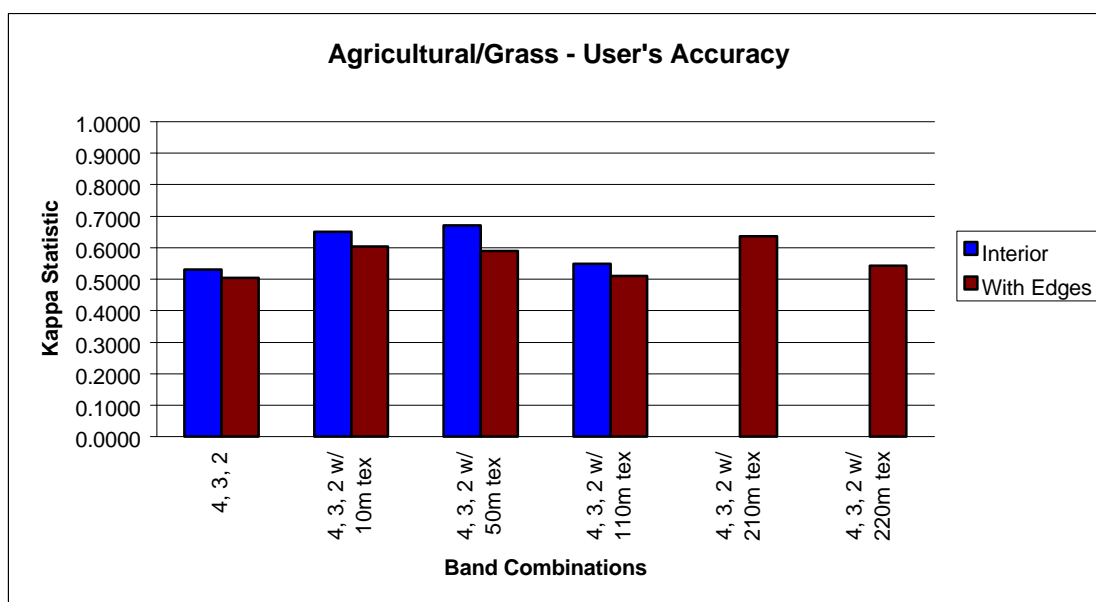
Figure 25: *User's accuracy of Agriculture/Grass class.*

Classes that have a relatively rough texture with high within-class variability that is similar to the between class variability, for example the Commercial/Industrial class, show little change in Kappa with the inclusion of pixels near the edges of the test polygon (*Figure 26*). At window sizes greater than 110 meters there is an insufficient number of pixels that are unaffected by between-class variance to evaluate the class interiors. Classification accuracy generally increases as window size increases across the entire graph, though the increased accuracy from the inclusion of texture is small. The 220-meter texture scale is the maximum scale used for evaluating classification accuracy; thus it is unknown if accuracy would improve further with the addition of texture at larger scales. All band combinations that include texture resulted in higher Kappa statistics than those with spectral data alone. In fact for very large window sizes (210 and 220-meter scales) the between-class texture seems to be more valuable than within-class texture.
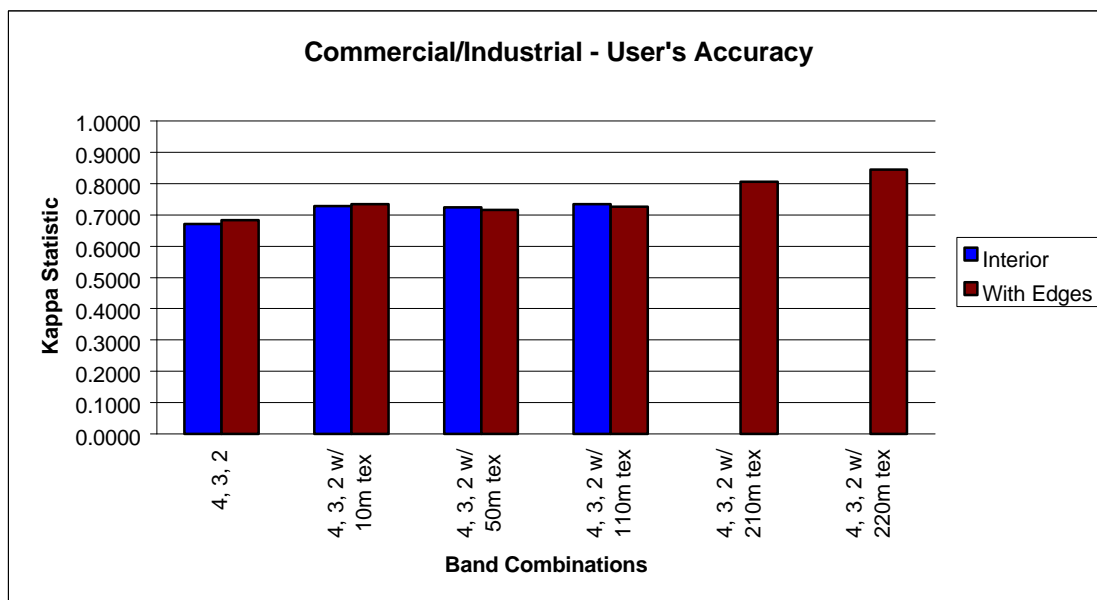
Figure 26: *User's accuracy of the Commercial/Industrial class.*

5. Producer's Accuracy

The producer's accuracy graphs more clearly illustrate the effect of edge pixels on the Kappa statistic because it provides a measure of the pixels missed due to the edge effect. As with the graphs of user's accuracy, there were insufficient pixels unaffected by between-class variance to evaluate class interiors at window sizes greater than 110 meters.

For the Agricultural/Grass class (*Figure 27*) the inclusion of texture data improved classification results at the 10, 50 and 110-meter scales for the interior of classes. A sharp reduction in Kappa can be seen at the 210 and 220 meter scales. The effect of edge pixels is clearly evident in that when accuracy is evaluated on the entire class, larger window sizes generally cause a decline in accuracy. The between-class texture pixel values overlap with the values of other classes, confusing the classifier, and are not classified as Agriculture/Grass.
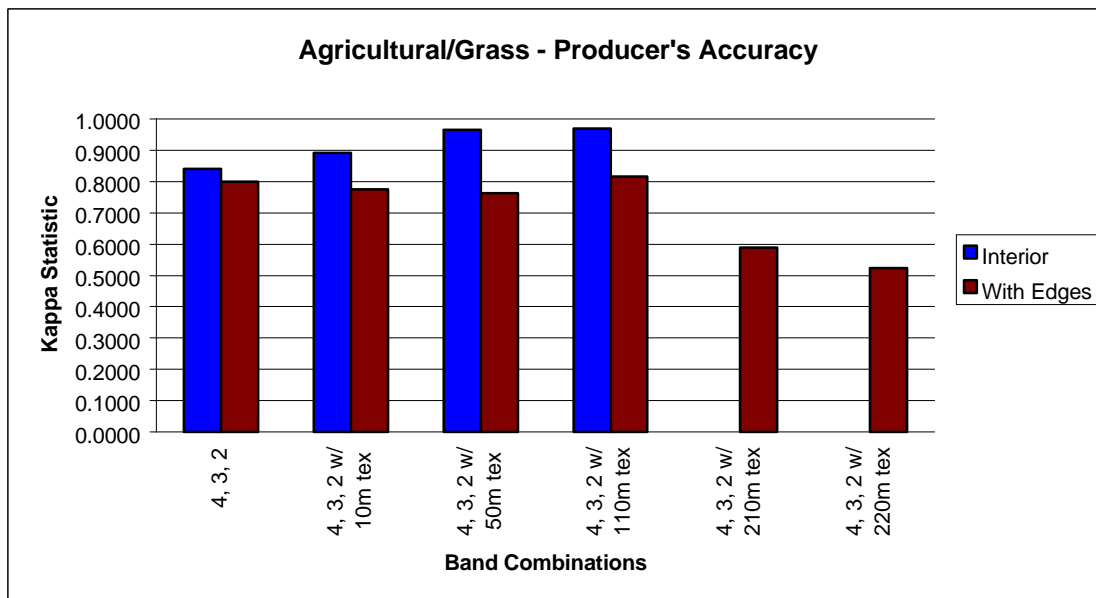
Figure 27: *Producer's accuracy for the Agricultural/Grass class.*

As the window size becomes increasingly larger than 110 meters and is dominated by between-class variance, the influence of the between-class texture confused the classification and resulted in Kappa statistics below .6000.

The graph for the Commercial/Industrial class is more complex (*Figure 28*). As was found for the user's accuracy (*Figure 26*), the inclusion of edge pixels has no significant effect on the Kappa statistic for any band combination or textural scale. Kappa increases with the addition of texture at the 10-meter scale then drops before reaching a minimum at the 110-meter scale before increasing again at 210 meters. At 220 meters accuracy was found to be lower.
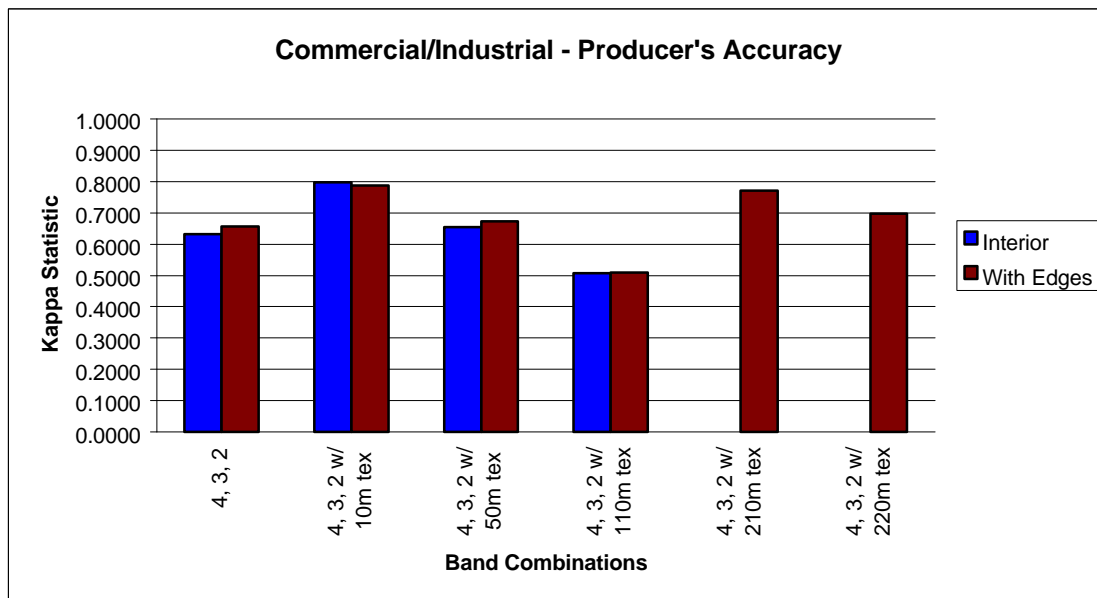
Figure 28: *Producer's accuracy of the Commercial/Industrial class.*

It is surprising that the interior pixels show a decline in accuracy with larger scales of texture, because *Figure 18* suggests that texture variability for the CBD is at its greatest at the 10-meter scale. The multiple-scale objects found in the Commercial/Industrial class might explain the apparent trough in Kappa at the 110-meter textural scale. Buildings and other structures are of a variety of sizes and shapes, creating a complex arrangement of image elements.

VII.   CONCLUSIONS

The theoretical framework presented here provides a basis for discussion and explanation of the concepts and relationships involved in image texture. Through these concepts it is possible to clarify the role of scale in digital image classification and evaluate the usefulness of texture. In conceptualizing scene objects as H- and L-resolution elements representing those objects, the multi-scale aspects of images can be more adequately investigated and analyzed. Because the use of image texture for classification depends on the informational classes chosen by the investigator, special emphasis must be given to hierarchies of scales and how the image elements relate to the classes. The concepts in this framework were used in developing the simulated and real image data experiments.

The results of the simulated data experiment show that per class data is often more useful for texture investigation than averages for an entire scene. A clear relationship was seen between window size for texture calculation and class separability. Class A had a coarse texture with a high within-class variance compared to the background; thus, the variability of texture dropped to a minimum value for both interior pixels and the entire class polygon. This is contrasted by Class B which had a smooth texture with a high between-class variance. When pixels near the edges were included in the analysis the texture variability remained high and only dropped as the number of edge affected pixels dominated the class statistic.

While plots of texture variability and tests on simulated data show the promise of finding appropriate window size for texture feature extraction, the results of the real data experiment reveal the inherent complexities involved in interpreting and analyzing actual image data. Multiple peaks in classification accuracy help to illustrate this complexity. Although objects within classes exist at multiple scales, shapes and orientations it is possible to make some general statements about the role scale plays in image texture.

The window sizes necessary to achieve a stable, relatively homogeneous measure of texture are often quite large even for relatively "smooth" informational classes such as Agricultural/Grass. For "coarse" classes, such as Commercial/Industrial, the window size might well measure larger than the dimensions of the class polygon. While classification often results in a large increase in accuracy for pixels far from class edges, the accuracy tends to fall for pixels near the edges of classes due to between-class variance. Since a useful application of textural information requires the inclusion of these edge pixels, the large window sizes necessary for achieving the stable texture measures result in a large number of pixel values

being influenced by between-class variance. The investigator must weigh the potential usefulness of texture against the added confusion of the between-class variance pixels. The size of the image elements, the class polygon dimensions (*Figure 4*) and the spectral properties of the classes and elements adjacent to the class of interest must be considered.

Measures of producer's accuracy and user's accuracy showed the effect of high between-class variability in image classification. Large window sizes for some classes resulted in improved user's accuracy, since no pixels were omitted from the class. The same windows resulted in lower producer's accuracy since the effect of between-class variance on the edge pixels caused many of these pixels to be placed in an incorrect category due to their high calculated texture values. In this case the Agricultural/Grass class is analogous to Class I in *Figure 5a* and the Commercial/Industrial class is analogous to Class II.

This thesis is an important step in applying the concepts of scale and texture to digital image classification. Further analysis is needed to fully determine the impact of class edges at various texture scales on image classification.

VIII.   REFERENCES

Agbu, P. A. and Nizeyimana, E. (1991), Comparisons between spectral mapping units derived from SPOT image texture and field soil map units. *Photogramm. Eng. Remote Sens.* 54(4):397-405.

Atkinson, P. M. and Curran, P. J. (1997), Choosing an appropriate spatial resolution for remote sensing investigations. *Photogramm. Eng. Remote Sens.* 63(12):1345-1351.

Barber, D. G. and LeDrew, E. F. (1991), SAR sea ice discrimination using texture statistics: a multivariate approach. *Photogramm. Eng. Remote Sens.* 57(5):385-395.

Basedow, R. W., Carmer, D. C. and Anderson, M. E. (1994), HYDICE system, implementation and performance. *Proceedings of the SPIE, Orlando Florida, 17-18 April, 1994,* V2480:258-267.

Begni, G. (1982), Selecting the optimal spectral bands for the SPOT satellite. *Photogramm. Eng. Remote Sens.* 48(10):1613-1620.

Briggs, J. M. and Nellis, M. D. (1991), Seasonal variation of heterogeneity in the tallgrass prairie: a quantitative measure using remote sensing. *Photogramm. Eng. Remote Sens.* 57(4):407-411.

Chavez, P. Jr. and Bauer, B. (1982), An automatic optimum kernel-size selection technique for edge enhancement. *Remote Sens. Environ.* 12:23-38.

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

Curran, P. J. (1988), The semivariogram in remote sensing: an introduction. *Remote Sens. Environ.* 24:493-507.

Cushnie, J. L. (1987), The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies. *Int. J. Remote Sensing* 8(1):15-29.

Dutra, L. V. and Mascarenhas, N. D. A. (1984), Some experiments with spatial feature extraction methods in multispectral classification. *Int. J. Remote Sensing* 5(2):303-313.

Franklin, S. E. and McDermid, G. J. (1993), Empirical relations between digital SPOT HRV and CASI spectral response and lodgepole pine (*Pinas contorta*) forest stand parameters. *Int. J. Remote Sensing* 14(12):2331-2348.

Franklin, S. E., Wulder, M. A. and Lavingne, M. B. (1996), Automated derivation of geographic window sizes for use in remote sensing digital image texture analysis. *Computers & Geosciences* 22(6):665-673.

Gong, P., Marceau, D. J. and Howarth, P. J. (1992), A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sens. Environ.* 40:137-151.

Haralick, R. M. (1979), Statistical and structural approaches to texture. *Proceedings of the IEEE* 67(5):786-804.

Haralick, R. M. and Shanmugam, K. S. (1974), Combined spectral and spatial processing of ERTS imagery data. *Remote Sens. Environ.* 3:3-13.

Haralick, R. M., Shanmugam, K. S. and Dinstein, I. (1973), Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 3(6):610-621.

Hay, G. J., Niemann, K. O. and Goodenough, D. G. (1997), Spatial thresholds, image-objects, and upscaling: A multiscale evaluation. *Remote Sens. Environ.* 62:1-9.

Hodgson, M. E. (1998), What size window for image classification? A cognitive perspective. *Photogramm. Eng. Remote Sens.* 64(8):797-807.

Hsu, S. (1978), Texture-tone analysis for automated land-use mapping. *Photogramm. Eng. Remote Sens.* 44(11):1393-1404.

Kramer, H. J. (1996), *Observation of the Earth and its Environment. Survey of Missions and Sensors.* Springer Verlag, Berlin, 960 pp.

Jensen, J. R. (1996), *Introductory Digital Image Processing: A Remote Sensing Perspective,* Prentice Hall, Upper Saddle River, New Jersey, 316 pp.

Latty, R. S., Markham, N. B., Williams, D., Toll, D., and Irons, J. (1985), Performance comparisons between information extraction techniques using variable spatial resolution data. *Photogramm. Eng. Remote Sens.* 51(9):1459-1470.

Lillesand, T. M. and Kiefer, R. W. (1994), *Remote Sensing and Image Interpretation.* John Wiley & Sons, Inc., New York, 747 pp.

Marceau, D. J., Howarth, P. J., Dubois, J. M., and Grattton, D. J. (1990), Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing* 28(4):513-519.

Marceau, D. J., Gratton, D. J., Fournier, R. A., and Fortin, J. (1994), Remote sensing and the measurement of geographical entities in a forested environment, 2., The Optimal Spatial Resolution. *Remote Sens. Environ.* 49:105-117.

Markham, B. L. and Barker, J. L. (1985), Spectral characterization of the Landsat Thematic Mapper sensors. *Int. J. Remote Sens.* 6(5):697-716.

Nellis, M. D. and Briggs, J. M. (1989), The effect of spatial scale on Konza landscape classification using textural analysis. *Landscape Ecology* 2(2):93-100.

Richards, J. (1993), *Remote Sensing Image Data Analysis,* Springer Verlag, New York, 340 pp.

Ryherd, R. and Woodcock, C. (1996), Combining spectral and texture data in the segmentation of remotely sensed images. *Photogramm. Eng. Remote. Sens.* 62(2):181-194.

Stow, D., Hope, A., Nguyen, A. T., Phinn, S., and Benkelman, C. A. (1996), Monitoring detailed land surface changes using an airborne multispectral digital camera system. *IEEE Transactions on Geoscience and Remote Sensing* 34(5): 1191-1203.

Vane, G., Green, R. O., Chrien, T. G., Enmark, H. T., Hansen, E. G., and Porter, W. M. (1993). The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sens. Environ.* 44:127-143.

Wang, L. and He, D. C. (1990), A new statistical approach for texture analysis. *Photogramm. Eng. Remote Sens.* 56(1):61-66.

Warner, T. A. and Shank, M. C. (1997), Spatial autocorrelation analysis of hyperspectral imagery for feature selection. *Remote Sens. Environ.* 60:58-70.

Weszka, J. S., Dyer, C. R. and Rosenfeld, A. (1976), A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics* 6(4):269-285.

Wilson, P. A. (1997), Rule-based classification of water in Landsat MSS images using the variance filter. *Photogramm. Eng. Remote Sens.* 63(5):485-491.

Woodcock, C. E. and Ryherd, S. L. (1989), Generation of texture images using adaptive windows. *Technical Papers, ASPRS/ACSM Annual Convention,* 2-7 April, Baltimore, Maryland. 2:11-22.

Woodcock, C. E. and Strahler, A. H. (1987), The factor of scale in remote sensing. *Remote Sens. Environ.* 21:311-332.

Woodcock, C. E., Strahler, A. H. and Jupp, D. L. B. (1988a), The use of variograms in remote sensing, I: scene models and simulated images. *Remote Sens. Environ.* 25:323-348.

Woodcock, C. E., Strahler, A. H. and Jupp, D. L. B. (1988b), The use of variograms in remote sensing, II: real digital images. *Remote Sens. Environ.* 25:349-379.

IX.   APPENDIX A – ERROR MATRICIES

A. Spectral bands only

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	346	0	0	0	0	0	346	0.0000
FOR	0	1688	2	76	12	0	1778	0.0506
AGR	0	2	504	383	3	0	892	0.4350
RES	2	5	81	1494	202	0	1784	0.1626
COM	0	0	0	515	1715	0	2230	0.2309
MIN	0	0	0	7	394	392	793	0.5057
Total	348	1695	587	2475	2326	392	7823	
Omission	0.0057	0.0041	0.1414	0.3964	0.2627	0.0000		0.2153

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9940
FOR	0.9354	0.9947
AGR	0.5297	0.8404
RES	0.7622	0.4865
COM	0.6713	0.6326
MIN	0.4676	1.0000

Overall Kappa = 0.7234

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	798	0	0	0	0	0	798	0.0000
FOR	0	3612	18	165	21	0	3816	0.0535
AGR	0	120	1051	718	12	0	1901	0.4471
RES	16	148	249	2000	361	3	2777	0.2798
COM	0	40	1	694	2336	15	3086	0.2430
MIN	0	3	0	13	450	673	1139	0.4091
Total	814	3923	1319	3590	3180	691	13517	
Omission	0.0197	0.0793	0.2032	0.4429	0.2654	0.0260		0.2254

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9791
FOR	0.9247	0.8895
AGR	0.5045	0.7636
RES	0.6190	0.4426
COM	0.6822	0.6561
MIN	0.5688	0.9716

Overall Kappa = 0.7140

B. Spectral bands and 10 meter texture

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	8444	0	0	0	0	0	8444	0.0000
FOR	0	42007	69	1665	79	0	43820	0.4140
AGR	0	30	13279	6308	23	0	19640	0.3239
RES	166	291	1362	42072	5864	0	49755	0.1544
COM	0	0	0	11857	50324	0	62181	0.1907
MIN	0	0	0	66	2112	9768	11946	0.1823
Total	8610	42328	14710	61968	58402	9768	195786	
Omission	0.0193	0.0076	0.0973	0.3211	0.1383	0		0.1527

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9799
FOR	0.9472	0.9902
AGR	0.6498	0.8919
RES	0.7741	0.5695
COM	0.7283	0.7973
MIN	0.8081	1.0000

Overall Kappa = 0.8002

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	19553	0	0	0	0	0	19553	0.0000
FOR	0	87905	518	3765	165	0	92353	0.0482
AGR	0	1647	26655	13148	68	6	41524	0.3581
RES	707	7487	6025	56622	10199	99	81139	0.3022
COM	63	698	32	15893	66884	465	84035	0.2041
MIN	12	99	0	178	2331	16647	19267	0.136
Total	20335	97836	33230	89606	79647	17217	337871	
Omission	0.0385	0.1015	0.1979	0.3681	0.1602	0.0331		0.1883

Class	Commission	Omission
WTR	1.0000	0.9592
FOR	0.9322	0.8603
AGR	0.6029	0.7744
RES	0.5888	0.5156
COM	0.7330	0.7867
MIN	0.8567	0.9649

Overall Kappa = 0.7587

C. Spectral bands and 50 meter texture

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	347	0	0	0	0	0	347	0.0000
FOR	0	1691	4	76	13	0	1784	0.0521
AGR	0	4	569	236	10	0	819	0.3053
RES	1	0	14	1727	235	0	1977	0.1265
COM	0	0	0	430	1783	0	2213	0.1943
MIN	0	0	0	6	285	392	683	0.4261
Total	348	1695	587	2475	2326	392	7823	
Omission	0.0029	0.0024	0.0307	0.3022	0.2334	0.0000		0.1680

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9970
FOR	0.9335	0.9969
AGR	0.6700	0.9657
RES	0.8150	0.5956
COM	0.7235	0.6745
MIN	0.5515	1.0000

Overall Kappa = 0.7825

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	799	0	0	0	0	0	799	0.0000
FOR	0	3590	33	179	23	0	3825	0.0614
AGR	0	48	1045	555	12	0	1660	0.3705
RES	15	229	240	2256	445	0	3185	0.2917
COM	0	52	1	583	2375	24	3035	0.2175
MIN	0	4	0	17	325	667	1013	0.3416
Total	814	3923	1319	3590	3180	691	13517	
Omission	0.0184	0.0849	0.2077	0.3716	0.2531	0.0347		0.2060

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9804
FOR	0.9134	0.8816
AGR	0.5895	0.7632
RES	0.6028	0.5139
COM	0.7156	0.6736
MIN	0.6400	0.9625

Overall Kappa = 0.7369

D. Spectral bands and 110 meter texture

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	347	0	0	0	0	0	347	0.0000
FOR	0	1692	6	80	14	0	1792	0.0558
AGR	0	3	571	385	20	0	979	0.4168
RES	1	0	10	1661	502	0	2174	0.2360
COM	0	0	0	331	1440	0	1771	0.1869
MIN	0	0	0	18	350	392	760	0.4842
Total	348	1695	587	2475	2326	392	7823	
Omission	0.0029	0.0018	0.0273	0.3289	0.3809	0.0000		0.2199

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9970
FOR	0.9288	0.9977
AGR	0.5494	0.9688
RES	0.6548	0.5445
COM	0.7340	0.5076
MIN	0.4902	1.0000

Overall Kappa = 0.7177

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	801	0	0	0	0	0	801	0.0000
FOR	0	3612	58	183	26	0	3879	0.0688
AGR	0	84	1112	763	32	0	1991	0.4415
RES	13	192	148	2156	814	2	3325	0.3516
COM	0	28	1	457	1896	17	2399	0.2097
MIN	0	7	0	31	412	672	1122	0.4011
Total	814	3923	1319	3590	3180	691	13517	
Omission	0.0160	0.0793	0.1569	0.3994	0.4038	0.0275		0.2418

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9830
FOR	0.9030	0.8888
AGR	0.5108	0.8160
RES	0.5213	0.4702
COM	0.7258	0.5091
MIN	0.5773	0.9700

Overall Kappa = 0.6930

E. Spectral bands and 210 meter texture

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	347	0	0	0	0	0	347	0.0000
FOR	0	1669	2	66	6	0	1743	0.0425
AGR	0	0	488	145	2	0	635	0.2315
RES	1	26	97	2022	271	0	2417	0.1634
COM	0	0	0	240	1987	0	2227	0.1078
MIN	0	0	0	2	60	392	454	0.1366
Total	348	1695	587	2475	2326	392	7823	
Omission	0.0029	0.0153	0.1687	0.1830	0.1457	0.0000		0.1173

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9970
FOR	0.9458	0.9803
AGR	0.7497	0.8164
RES	0.7609	0.7351
COM	0.8466	0.7963
MIN	0.8562	1.0000

Overall Kappa = 0.8453

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	800	0	0	0	0	0	800	0.0000
FOR	0	3412	10	158	12	0	3592	0.0501
AGR	0	21	825	376	5	0	1227	0.3276
RES	14	425	484	2772	477	13	4185	0.3376
COM	0	62	0	273	2611	20	2966	0.1197
MIN	0	3	0	11	75	658	747	0.1191
Total	814	3923	1319	3590	3180	691	13517	
Omission	0.0172	0.1303	0.3745	0.2279	0.1789	0.0478		0.1804

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9817
FOR	0.9294	0.8226
AGR	0.6369	0.5881
RES	0.5403	0.6700
COM	0.8435	0.7708
MIN	0.8744	0.9494

Overall Kappa = 0.7668

F. Spectral bands and 220 meter texture

Interior Pixels

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	346	0	0	0	0	0	346	0.0000
FOR	0	1684	2	66	10	0	1762	0.0443
AGR	0	0	447	261	3	0	711	0.3713
RES	2	11	138	1901	296	0	2348	0.1904
COM	0	0	0	242	1791	0	2033	0.1190
MIN	0	0	0	5	226	392	623	0.3708
Total	348	1695	587	2475	2326	392	7823	
Omission	0.0057	0.0065	0.2385	0.2319	0.2300	0.0000		0.1613

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9940
FOR	0.9435	0.9916
AGR	0.5986	0.7377
RES	0.7215	0.6686
COM	0.8306	0.6892
MIN	0.6097	1.0000

Overall Kappa = 0.7895

Entire Class

	WTR	FOR	AGR	RES	COM	MIN	Total	Commission
WTR	797	0	0	0	0	0	797	0.0000
FOR	0	3570	20	149	19	0	3758	0.0500
AGR	0	22	751	497	7	0	1277	0.4119
RES	17	225	547	2636	474	1	3900	0.3241
COM	0	103	1	296	2422	25	2847	0.1493
MIN	0	3	0	12	258	665	938	0.2910
Total	814	3923	1319	3590	3180	691	13517	
Omission	0.0209	0.0900	0.4306	0.2657	0.2384	0.0376		0.1980

Kappa Index of Agreement (KIA)

Class	Commission	Omission
WTR	1.0000	0.9778
FOR	0.9295	0.8754
AGR	0.5436	0.5244
RES	0.5587	0.6265
COM	0.8048	0.6980
MIN	0.6933	0.9596

Overall Kappa = 0.7451

Christopher John Sebastian Ferro

EDUCATION

1996 - 1998
Master of Arts in Geography in the Department of Geology and Geography at West Virginia University in Morgantown, WV
Thesis Title: Scale and Texture in Digital Image Classification

1992 - 1996
Bachelor of Science *magna cum laude* in Geography in the Department of Earth Sciences and Geography at Bridgewater State College in Bridgewater, MA
Minor in Art (Graphic Design and Art History)
Member of Gamma Theta Upsilon (International Honor Society for Geographers)

EMPLOYMENT
1998 Graduate Research Assistant West Virginia State GIS Technical Center
Morgantown, WV

1996 - 1998 Graduate Teaching Assistant, West Virginia University Department of Geology and Geography
Morgantown, WV

1997 Graduate Research Assistant, National Center for Geographic Information and Analysis, Initiative 19
Morgantown, WV

1996 GIS Support/Intern, Taunton River Watershed Connections Project
Bridgewater, MA

SOFTWARE/COMPUTER SKILLS

ERDAS Imagine	ArcView	Arc/Info	Windows
HTML	XML	Idrisi	Microsoft Office