2006

Detecting the socioeconomic conditions of urban neighborhoods through wavelet analysis of remotely sensed imagery

Guiyun Zhou
Louisiana State University and Agricultural and Mechanical College, gzhou1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations
Part of the Social and Behavioral Sciences Commons

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/485

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradtd@lsu.edu.
DETECTING THE SOCIOECONOMIC CONDITIONS
OF URBAN NEIGHBORHOODS THROUGH WAVELET ANALYSIS
OF REMOTELY SENSED IMAGERY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Geography and Anthropology

by
Guiyun Zhou
B.S., Beijing University, Beijing, China, 1999
M.S., Beijing University, Beijing, China, 2002
December, 2006
Acknowledgements

I am deeply grateful to my major professor, Dr. Nina Lam, for her advice on my study and research, and her financial support throughout my study at LSU since the fall of 2002. Her advice will always be a great asset in my future career. I thank Dr. Tyler, Dr. Curtis, Dr. Leitner, Dr. Bosworth, and Mr. Braud for serving on my dissertation committee. Their valuable comments and advice greatly helped to improve my dissertation. I profoundly appreciate DeWitt Braud for providing 1-foot resolution aerial images of Baton Rouge, Louisiana as well as his advice for this study.

I am indebted to my fellow students at LSU. Their friendship will always be greatly cherished and be inscribed in my memory. Their help and support made my life at LSU more comfortable and exciting.

This research was partially supported by the Doctoral Dissertation Research Improvement (DDRI) Awards from the National Science Foundation (BCS-0602111), IGIF(International Geographic Information Fund) Student Travel Grant from the Association of American Geographers, and the Robert C. West field research grant from the Department of Geography and Anthropology, LSU.
Table of Contents

Acknowledgements................................................................................................................................... ii

List of Tables .................................................................................................................................................... v

List of Figures.................................................................................................................................................... vii

Abstract............................................................................................................................................................. x

Chapter 1 Introduction ..................................................................................................................................... 1
  1.1 Remote Sensing of Urban Areas................................................................................................................ 1
  1.2 Textural Analysis...................................................................................................................................... 3
  1.3 Problem Statement................................................................................................................................. 4
  1.4 Research Objectives and Hypotheses.................................................................................................... 7
  1.5 Expected Significance........................................................................................................................... 8
  1.6 Chapter Organization............................................................................................................................. 9

Chapter 2 Reducing Edge Effects in the Classification of High-Resolution Imagery................................. 11
  2.1 Introduction............................................................................................................................................ 11
  2.2 The New Approach............................................................................................................................... 13
  2.3 A Worked Example ............................................................................................................................. 15
  2.4 Classification Scenarios....................................................................................................................... 17
    2.4.1 Classification of an Artificial Image.............................................................................................. 19
    2.4.2 Classification of a Mosaic Image .................................................................................................. 27
    2.4.3 Identification of Non-agricultural Land Use by Logistic Regression ........................................ 33
  2.5 Discussion............................................................................................................................................. 43
  2.6 Conclusion ........................................................................................................................................... 45

Chapter 3 Improving High-Resolution Image Segmentation by Combining Segmentations Using Different Textural Measures.................................................................................................. 46
  3.1 Introduction............................................................................................................................................ 46
  3.2 Textural Description ............................................................................................................................. 47
    3.2.1 Histogram ....................................................................................................................................... 47
    3.2.2 Local Binary Pattern and Contrast Measure.................................................................................. 47
    3.2.3 Saturation/Hue Distribution........................................................................................................ 49
  3.3 Similarity Index of Measure Distributions ............................................................................................ 49
    3.3.1 Histogram Intersection .................................................................................................................. 50
    3.3.2 Histogram Correlation .................................................................................................................. 50
    3.3.3 G-statistic......................................................................................................................................... 50
  3.4 Segmentation Procedures....................................................................................................................... 51
    3.4.1 Hierarchical Splitting.................................................................................................................... 51
    3.4.2 Agglomerative Merging................................................................................................................ 53
    3.4.3 Pixelwise Refinement................................................................................................................... 54
  3.5 Post-segmentation Integration.................................................................................................................. 54
  3.6 Experiments............................................................................................................................................ 55
List of Tables

Table 2.1 Means and standard deviations of 4 Gaussian distributions........................................... 21
Table 2.2 Overall classification accuracy of the artificial image (%)............................................... 23
Table 2.3 Overall classification accuracy of the mosaic image (%)................................................. 29
Table 2.4 Overall classification accuracy using the logistic regression (%)................................. 39
Table 4.1 Land use/land cover classification scheme................................................................. 89
Table 4.2 Overall classification accuracy of textural samples with window size of 64 by 64 .... 96
Table 4.3 Kappa Z statistics with window size of 64 by 64 applied to textural samples ............. 97
Table 4.4 The error matrix by the neural network classifier with the mean measure and the db2 wavelet at two decomposition levels applied to textural samples ........................................ 100
Table 4.5 The error matrix by discriminant analysis with the mean measure and the db2 wavelet at two decomposition levels applied to textural samples ........................................ 101
Table 5.1 Dimensions of subsets of six urban neighborhoods at 1-foot resolution and their median household income in 1999...................................................................................... 113
Table 5.2 List of means of the two measures of four subimages at the first decomposition level using Band 1, a window size of 85m by 85m and the db8 wavelet................................. 120
Table 5.3 List of means of the two measures of four subimages at the first decomposition level using Band 2, a window size of 85m by 85m and the db8 wavelet................................. 121
Table 5.4 List of means of the two measures of four subimages at the first decomposition level using Band 3, a window size of 85m by 85m and the db8 wavelet................................. 122
Table 5.5 Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 45m by 45m window size ...................................................................................... 127
Table 5.6 Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 65m by 65m window size ...................................................................................... 128
Table 5.7 Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 85m by 85m window size ...................................................................................... 129
Table 5.8 Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 45m by 45m window size ...................................................................................... 130
Table 5.9 Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 65m by 65m window size ...................................................................................... 131
Table 5.10 Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 85m by 85m window size ................................................................. 132

Table 5.11 Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 45m by 45m window size ................................................................. 134

Table 5.12 Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 65m by 65m window size ................................................................. 135

Table 5.13 Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 85m by 85m window size ................................................................. 136

Table 5.14 Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 45m by 45m window size. ....................... 139

Table 5.15 Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 65m by 65m window size ....................... 140

Table 5.16 Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 85m by 85m window size ....................... 141

Table 5.17 Overall accuracies (%) for neighborhood discrimination at three window sizes at three bands and at three resolutions ................................................................. 144
List of Figures

Figure 2.1 Illustration of edge effects............................................................................................................. 14
Figure 2.2 A worked example of the proposed approach...................................................................................... 16
Figure 2.3 An artificial image generated from 4 Gaussian distributions............................................................ 20
Figure 2.4 Pixelwise classified image using maximum-likelihood classifier................................................... 22
Figure 2.5 Classification accuracy with different window sizes for the artificial image by the traditional and the new methods.................................................................................................................. 24
Figure 2.6 Classified results of the artificial image with a window size of 33×33................................................. 26
Figure 2.7 Mosaic image of six land use classes from the panchromatic band of an IKONOS image of Atlanta, GA............................................................................................................................................. 28
Figure 2.8 Classification accuracy with different window sizes for the artificial image by the traditional and the new methods.................................................................................................................. 28
Figure 2.9 Classified results of the mosaic image with a window size of 85×85.................................................... 30
Figure 2.10 A subset of IKONOS panchromatic band with agricultural and non-agricultural land uses...................................................................................................................................................... 32
Figure 2.11 Function curve between the probability and the logit........................................................................... 34
Figure 2.12 Overall classification accuracy using logistic regression............................................................... 36
Figure 2.13 Classified images overlaid on the reference map for non-agricultural land use the two methods...................................................................................................................................................... 38
Figure 3.1 Computation of local binary pattern and contrast measure........................................................................ 48
Figure 3.2 Illustration of the segmentation algorithms.......................................................................................... 52
Figure 3.3 The subset of a panchromatic IKONOS image in northern Alabama................................................ 56
Figure 3.4 Clustering results using ISODATA procedure and pixel values only.................................................. 57
Figure 3.5 Segmentation of the panchromatic subset using the intensity distribution...................................... 58
Figure 3.6 Segmentation of the panchromatic subset using the LBP/C distribution........................................... 59
Figure 3.7 Combination of segmentations of the panchromatic subset from using intensity distribution and using LBC/C distribution.................................................................................................................. 60
Figure 3.8 The subset of the red band of an aerial photo in Baton Rouge, LA..................................................... 63
Figure 3.9 Segmentation of the subset of the photo using the intensity distribution................. 64
Figure 3.10 Segmentation of the subset of the photo using the intensity distribution.............. 65
Figure 3.11 Combination of segmentations of the subset of the aerial photo from using intensity distribution and using LBC/C distribution............................................................................. 66
Figure 3.12 Left: Segmentation of a subset of a panchromatic image (one-foot resolution) using the intensity distribution. Right: Segmentation using the LBP/C distribution ....................... 67
Figure 3.13 The color subset of an aerial photo (0.9m in resolution, 512 by 512 pixels)............ 69
Figure 3.14 Segmentation of the color subset using the hue/saturation distribution ............... 70
Figure 3.15 Segmentation of the subset of the photo using the LBP/C distribution .................. 71
Figure 3.16 Combination of segmentations of the color subset of the aerial photo from using intensity distribution and using LBC/C distribution......................................................... 72
Figure 3.17 Segmentation of a subset of an aerial photo using the LBP/C distribution .......... 73
Figure 3.18 Segmentation of a subset of an aerial photo using the saturation/hue distribution ... 74
Figure 4.1 Four one-level decomposed images with Harr wavelet. ........................................... 81
Figure 4.2 One-level two-dimensional discrete wavelet transform............................................. 82
Figure 4.3 The structure of a three-layer multilayer perceptron neural network...................... 85
Figure 4.4 The metropolitan area of Atlanta, Georgia, USA and the location of the image ...... 87
Figure 4.5 Samples of five land-use/land-cover classes from the panchromatic band of the IKONOS image of Atlanta............................................................................................................. 90
Figure 4.6 Daubechies wavelets of order 2 and order 8 ............................................................. 92
Figure 4.7 Mean and standard deviation plots of measures of four subimages at the first decomposition level using a window size of 64 by 64 and the db2 wavelet.. ....................... 95
Figure 4.8 The overall classification accuracy of three window sizes: 64×64, 48×48, and 32×32 with the db2 wavelet and the db8 wavelet by the neural network approach ..................... 102
Figure 4.9 The overall classification accuracy of three window sizes: 64×64, 48×48, and 32×32 with the db2 wavelet and the db8 wavelet by discriminant analysis ......................... 103
Figure 4.10 A mosaic image composed of 5 land-use classes and its classified results......... 104
Figure 4.11 User’s accuracy and producer’s accuracy of the mosaic image....................... 105
Figure 5.1. Locations of six urban/suburban neighborhoods in Baton Rouge Louisiana ........... 110
Figure 5.2 Image subsets of urban neighborhoods. ............................................................ 114
Figure 5.3 Image subsets of urban neighborhoods ............................................................ 115
Figure 5.4 Image subsets of urban neighborhoods ............................................................ 116
Figure 5.5 Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 1 .......................................................... 123
Figure 5.6 Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 2 .......................................................... 124
Figure 5.7 Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 3 .......................................................... 125
Figure 5.8 Overall accuracies of three window sizes using each individual band of one-foot aerial image ............................................................................................................. 137
Figure 5.9 Overall accuracies of three window sizes using each individual band of an aerial image of 0.9 meter. ............................................................................................... 142
Figure 5.10 Overall accuracies of three window sizes using each individual band of an aerial image of 2.7m. ............................................................................................... 143
Abstract

Wavelet analysis is an efficient approach to studying textural patterns at different scales. Artificial neural networks can learn very complex patterns in the data and could be an efficient classifier. However, whether wavelet analysis, in combination with artificial neural networks or other classifiers, can be used to detect the social-economic conditions of urban neighborhood is a key research question that needs further study. The hypotheses of this study were: 1) neural networks yielded higher classification accuracy than linear discriminant analysis and the minimum-distance classifier based on wavelet measures of urban land covers; 2) wavelet textural measures could be used to efficiently discriminate among urban neighborhoods of different social-economic conditions; 3) image resolution had great influences on the discrimination of urban neighborhoods; and 4) window size had great influences on the discrimination of urban neighborhoods. In addition, two technical problems related to the application of textural approach, including the edge effect and image segmentation problem, were examined.

The results show that the new approach developed to reducing edge effects consistently achieved higher accuracy than the traditional moving-window approach. The post-segmentation integration scheme in the region-based splitting-and-merging segmentation procedures reflected all the segmented clusters identified by two or more textural measures and was helpful in identifying homogeneous regions in an image. Regarding the four hypotheses, (1) The minimum-distance classifier performed the worst. Neural networks were found to generally yield slightly better results than discriminant analysis but the difference was not statistically significant. The first hypothesis was shown to be invalid. (2) With a window size of 85m by 85m, an overall accuracy of 93.00% was achieved using band 2 and an overall accuracy of 96.83% was achieved using combination of band 2 and band 3. (3) The 1-foot resolution subsets were found to yield
higher classification accuracy than the 0.9m resolution subsets and the 2.7m resolution subsets for band 2 and band 3 for the six neighborhoods in Baton Rouge, Louisiana. The differences were generally over 5%. (4) Window size was found to have great influences on the discrimination of urban neighborhoods. The larger the window size, the higher the classification accuracy.
Chapter 1
Introduction

1.1 Remote Sensing of Urban Areas

Today, nearly half of the world’s population lives in urban areas (United Nations 2002). Cities are playing more and more significant role in the world’s social, economic, and political systems. The rapid urbanization in developing countries and suburbanization in developed countries invite complex social-economic and environmental problems, profoundly impacting the health and quality of life of human beings. Timely, accurate, and detailed information about cities and their suburbs are required by many government agencies and private companies such as planning agencies and councils of government, Departments of Commerce and Transportation, Tax Assessor offices, Public Service Commissions, private real estate companies, private utility companies, and residential and commercial developers (Jensen and Cowen 1999).

The advent of remote sensing technology makes periodic data about urban land cover practicable and affordable. Today, high-resolution data acquired by such sensors as IKONOS and QuickBird open up new opportunities and present new challenges for urban applications. The availability of high-resolution remotely sensed images in urban context call for new approaches to automatically extract land-cover information (Jensen et al. 2004). As image spatial resolution becomes finer, texture information becomes more dominant, hence critical and indispensable for the interpretation of remotely sensed images. The classification algorithms will need to be able to deal with the heterogeneous nature of urban environment at high resolution and will need to incorporate both the spectral and spatial information.

Spatial resolution is one of the most crucial issues in urban remote sensing (Welch 1982, Jensen and Cowen 1999). Among the earliest commercial satellite sensors, Landsat MSS has a nominal resolution of 80 meters, Landsat TM 30 meters, SPOT HRV 10 meters for panchromatic
band and 20 meters for multispectral band. According to Jensen and Cowen (1999), Landsat MSS will allow Level I classification in the USGS (U.S. Geological Survey) Land Use/Land Cover classification system (Anderson et al. 1976), whereas TM and SPOT HRV will allow Level II classification. A minimum spatial resolution of 1-5 meters, such as data acquired from IKONOS and QuickBird, enables the identification of urban classes down to Level III, such as the discrimination between single-family and multi-family housing.

Fundamental to urban planning and decision-making is the efficient characterization of urban landscape, which “facilitates the study of local and regional environmental processes in the broader context of global environmental change and the sustainability of cities and their hinterlands” (de Sherbinin et al. 2002). Urban environment is characterized by heterogeneous patterns, a variety of mixed land uses such as residential buildings, commercial infrastructures, transportation networks, utilities, and recreational parks. These land uses are composed of a wide range of materials including concrete, asphalt, metal, plastic, glass, shingles, water, grass, shrubs, trees, and soil (Jensen and Cowen 1999).

Whether it is with high- or low-resolution remotely sensed data, urban areas present one of the most challenging issues in image classification. Data of higher spatial resolution are supposed to be more efficient in characterizing urban environment, leading to higher classification accuracy. Yet it has been reported that higher resolution leads to lower classification accuracy in urban areas by conventional per-pixel classification methods (Toll 1985, Hacck 1987, Chen et al. 2004). The reason lies in that improved spatial resolution has two competing effects. One effect is the decrease of the number of mixed pixels, which increases classification accuracy. On the other hand, the increased resolution increases within-class variance, which contributes to the violation of the assumptions of classical statistical classifiers such as maximum-likelihood classifier. Due to the diverse land uses and materials, this issue is particularly salient in urban areas in contrast to non-urban areas.
In view of the extreme heterogeneity of surface materials at both the inter-pixel and intra-pixel scales, a number of difficulties arise with the complexity of urban landscape, necessitating the incorporation of texture information into the classification of urban environments.

1.2 Textural Analysis

Remote sensing experts and professionals utilize “the tone, color, texture, shape, size, orientation, pattern, shadow silhouette, site, and situation of objects in the urban landscape to identify and judge their significance” (Jensen and Cowen 1999). Textural information has been employed to increase classification accuracy (van Teeffelen et al. 2001, Dekker 2003, Shaban and Dikshit 2001, Magnussen et al. 2004, Emerson et al. 2005, Myint and Lam 2005).

Texture is elusive to define (Rao 1990). A texture pattern is comprised of a large amount of simple primitive patterns as well as spatial relationships among them (Ballard and Brown 1982, p168). A texture primitive is a basic texture element that occurs in different positions, deformations, or orientations inside a texture pattern. In either a deterministic or a stochastic process, the primitive patterns are repeated to form the texture pattern. The primitive pattern has similar properties at different locations. The degree to which primitive patterns differ and the manner in which they are spaced determines the texture pattern.

Image spatial resolution is a crucial factor in texture patterns (Ballard and Brown 1982, p170). There is a range of resolutions at which texture patterns are discerned. Low spatial resolution tends to smooth an image and form a continuous gray-tone. High spatial resolution tends to highlight the details and form varied patterns. A forest is a continuous tone or color in a Landsat MSS (80m resolution) or TM image (30m resolution). The tree patterns begin to appear in an IKONOS panchromatic image with a pixel resolution of one meter. As the resolution becomes finer, leaf patterns can be discerned. This issue of pixel resolution is a special case of the more general scale problem (Lam et al. 2004), which is an important research priority in
many disciplines. Finding a range of characteristic scales suitable for the study of a particular geographic issue is a prime research area.

To efficiently analyze texture patterns, we need to extract features from textures. Feature extraction is the process of mapping the original measurements into more effective features (Benediktsson and Sveinsson 1997). For texture, we need to find a set of measurements, which can be associated with textural patterns and be used to characterize textural patterns.

Currently, there are no widely accepted ways to characterize texture. The measures used to characterize texture vary from relatively simple indices of image texture to those derived from complex mathematical transformations. The widely used tools to extract textural information include first-order and second-order statistics, gray-level co-occurrence matrix (Haralick 1973, Ballard and Brown 1982, p186, de Martino et al. 2003), fractal geometry (Quattrochi et al. 2001, Lam et al. 2002, Emerson et al. 2005), lacunarity (Dong 2000, Greenhill et al. 2003, Myint and Lam 2005), wavelet analysis (Zhu and Yang 1998, Myint 2001, Myint et al. 2002 and 2004), geostatistical analysis (Brivio and Zilioli 2001), mathematical morphology (Pesaresi and Bianchin 2001, Benediktsson et al. 2001, Benediktsson et al. 2003), and spatial autocorrelation (Read and Lam 2002, Myint et al 2004). Among them, wavelet transform has been shown to be highly efficient for texture analysis.

1.3 Problem Statement

Remote sensing technology has traditionally been used extensively in earth sciences and related fields. Since the late 1980s, remote sensing has seen considerable increase in social science applications (Liverman et al. 1998, de Sherbinin et al. 2002). Typical social science fields where remote sensing can make contributions include demography, human health and epidemiology, archeology and anthropology, land-use change and sustainability trajectories, and urban studies (de Sherbinin et al. 2002, Mennis and Liu 2005).
The identification of urban neighborhoods of various social-economic conditions is important to many governmental agencies and private companies. For the past half a century, urban sprawl has been a major trend of the urbanization process across the United States (Gillham 2002). The massive exodus of the middle and upper classes to the suburbs left behind run-down neighborhoods at the centers of cities. Many cities have taken steps to revitalize urban neighborhoods through such measures as renovation and gentrification. As a result, neighborhoods of different social-economic statuses co-exist in urban areas, ranging from run-down neighborhoods through renovated and gentrified neighborhoods to well-managed middle- and upper-class neighborhoods. Using census block group median income as a criterion, the textural samples of different income-level urban neighborhoods can be extracted from high-resolution remotely sensed images. Distinct textural patterns can be observed, where low-income neighborhoods tend to have dense housing patterns, whereas middle- and upper-income neighborhoods are less compacted. Moreover, single-family housing generally has more tree cover than multifamily apartment housing. Multifamily housing, on the other hand, exhibits more regular linear or gridded patterns. These observable differences among different neighborhoods provide the major hypothesis of this research, that is, they are expected to be detectable with advanced image analysis techniques. The identification of neighborhoods of divergent social-economic conditions provides further information beyond the USGS Level III land-use/land-cover classes and corresponds to the details at Level IV. Through the study of textures, the identification offers a way to link remote sensing analysis to the social-economic aspects of human activities, i.e., to "socialize the texture", following the phrase "socializing the pixel" (Geoghegan et al. 1998)

However, very little research has focused on applying remote sensing to distinguish different urban neighborhoods in terms of their socio-economic conditions. Urban neighborhoods are an important part of urban landscape. Successful detection of neighborhoods at different socio-
economic statuses will shed light on how the spatial “form” of urban neighborhood, as manifested in remote sensing imagery, is related to its “process”. For example, well-managed or upper-class neighborhoods may have more trees than run-down neighborhoods, hence forming a spectral as well as spatial signal that can be detected by high-resolution remotely sensed images.

The goal of this research is to develop innovative methods for detecting the socio-economic conditions of neighborhoods (e.g., poverty level) through the use of remote sensing imagery. In particular, advanced textural and numerical methods including wavelets and artificial neural networks (ANNs) will be examined.

Geographic phenomena operate at different scales (Lam et al. 2004). An image may encompass patterns at different scales. A multiresolution approach to deal with texture and spatial patterns meets the nature of geographic phenomena. Multiresolution wavelet analysis (Mallat 1989, Daubechies 1992, Frazier 1999) has proven to be an effective tool for the characterization of textures and spatial patterns at different scales (Zhu and Yang 1998, Myint et al. 2004). However, as far as the classification of wavelet signatures is concerned, many researchers used simple classifiers such as minimum-distance classifier (Zhu and Yang 1998, Zhao 2001) and linear discriminant analysis (Myint et al. 2004). Minimum-distance classifier takes only distance into account and discriminant analysis has many assumptions about the input data, which may not be satisfied.

Artificial Neural Networks (ANNs) have long been employed as an effort to improve the classification accuracy of remote sensing images (Paola and Schowengerdt 1995, Atkinson and Tatnal 1997). The greatest strength of ANNs lies in that they make no assumptions on the input data and tend to be more robust to the selection of training sets and heterogeneous classes. Therefore, it is expected that ANN will perform better than minimum-distance classifier and linear discriminant analysis when applied to wavelet signatures, which may not follow any distributions.
Although wavelet analysis and ANNs have been employed separately to analyze remotely sensed data, studies that bring them together are rare and the synergy is yet to be explored. This study will examine whether texture patterns are associated with the social-economic conditions of urban neighborhoods and whether such patterns can be efficiently detected using a combined methodology of wavelets and ANNs. The effects of scale on these methods will also be investigated.

1.4 Research Objectives and Hypotheses

Based upon the above discussion, the objectives of this research included:

• To determine whether ANNs in combination with wavelet analysis are effective for the characterization and classification of urban land covers;

• To determine whether wavelet textural measures can be associated with the social-economic conditions of urban neighborhoods;

• To explore the effects of image resolution and window size on the discrimination of urban neighborhoods.

Accordingly, the hypotheses of the study were:

• ANNs yielded higher classification accuracy than linear discriminant analysis and the minimum-distance classifier based on wavelet measures of urban land covers;

• Wavelet textural measures could be used to efficiently discriminate among urban neighborhoods of different social-economic conditions;

• Image resolution had great influences on the discrimination of urban neighborhoods;

• Window size had great influences on the discrimination of urban neighborhoods.

This dissertation also covered two topics encountered during the process of the study. The first topic was on reducing edge effects when a traditional moving-window method was used to collect textural information on high-resolution imagery. A new approach was proposed to
mitigate the edge effects. The second topic was on improving image segmentation by combining segmentations derived from using different textural measures. A post-segmentation combination method was proposed to take advantage of the discriminative abilities of different textural measures.

1.5 Expected Significance

The research is innovative in two aspects. First, although wavelet analysis and artificial neural networks have been employed separately to analyze remote sensing images, studies that bring them together are rare and the synergy is yet to be explored. This research will increase our knowledge of textures, texture methods, and their relationships with real features on the ground. The research will illustrate the importance of utilizing spatial and texture features for more accurate urban classification. It will gain insights into the relationships among window size, wavelet type, decomposition, and classification accuracy. The research will enhance our appreciation of the scale effect on the classification of heterogeneous urban environment from high-resolution imagery. Second, this study is among the first to examine the feasibility of relating textures from images to the economic and social conditions on the ground. The identification and characterization of urban neighborhoods of different social-economic statuses holds the promises of opening up a new avenue to link remotely sensed images to the social-economic aspects of human activities, i.e., to “socialize the texture”.

Although the proposed research focuses on income in urban neighborhoods, the proposed methodology can be extended to other classification scenarios, i.e., non-urban, low- and medium-resolution applications (e.g., healthy and unhealthy vegetation). The MATLAB modules developed for this research can be widely disseminated and will be useful to future studies of other urban as well as non-urban areas. Successful application of the proposed methodology could open up many new areas of applications, such as using the methodology to infer different
sanitary or health conditions on the ground. This will add significantly to the field of remote sensing, image processing, environmental assessment and monitoring, and urban analysis. Successful detection of neighborhoods with different socio-economic status will be useful to governmental agencies for urban planning and monitoring.

1.6 Chapter Organization

The dissertation comprises six chapters, including this introduction.

Chapter 2 deals with the edge effects when textural information is incorporated in the classification process. In this Chapter, a new approach was proposed to reduce the edge effects encountered when the traditional moving-window method was used to incorporate textural measures computed in the classification. Three experiments were carried out to illustrate the abilities of the proposed approach in reducing edge effects.

Chapter 3 covers the topic of region-based split-and-merge image segmentation. A post-segmentation integration scheme was proposed to improve high-resolution image segmentation by combining segmentations derived from using different textural measures. A series of experiments including both gray-scale images and color images were conducted to illustrate how the proposed post-segmentation integration scheme helped to achieve better results.

In Chapter 4, a combined method of wavelet analysis and artificial neural networks was used to improve textural discrimination of urban areas. Wavelet transforms were used to extract signatures from urban land-use/land-cover classes and neural networks were used to carry out the classification of wavelet signatures. The performance of neural networks as a textural classifier was compared with that of discriminant analysis and minimum distance classifier, which had been used in previous literature.

Chapter 5 is concerned with the classification of urban/suburban neighborhoods of various social-economic statuses. Wavelet measures were extracted from textural samples of urban
neighborhoods and classified using linear discriminant analysis. Various factors were examined in this Chapter, including window size, band combination and image resolution. Six urban/suburban neighborhoods, a cross-section of urban/suburban neighborhoods in East Baton Rouge parish in terms of median household income, were studied.

Chapter 6 summarizes the results in previous Chapters and provides conclusions and recommendation for future research.
Chapter 2
Reducing Edge Effects in the Classification of High-Resolution Imagery

2.1 Introduction

Texture information has long been employed to improve the classification accuracy of remotely sensed imagery (Hsu 1978, Gong and Howarth 1992, St-Onge and Cavayas 1995). As image resolution increases, land-use classes become more and more heterogeneous, and the statistical distributions of the pixel values are seldom normal. This has made traditional spectral-based classifiers such as the maximum-likelihood classifier yield unsatisfactory results, as it violates their assumptions such as the multivariate normal distribution (Haack et al. 1987, Chen et al. 2004).

The incorporation of textural information in the mapping and classification process brings about new problems when using the traditional moving window method (Ferro and Warner 2002). Large windows produce stable textural measures but large edge effects as well. Small windows have reduced edge effects but less stable textural measures. The trade-off between edge effect and window size is hard to be predetermined before the accuracy assessment. The edge issue is blamed for most of the classification errors (Maillard 2003, Warner and Stelnmaus 2005, Pearlstine et al. 2005). As a result, many researchers avoided the edge pixels in accuracy assessment and an overly optimistic result might be obtained (Ferro and Warner 2002).

Studies to find the optimum window size prior to image classification have been going on for a long time and contradictory results were reported in the literature. Nellis and Briggs (1989) suggested that small window sizes might be appropriate in complex landscapes and large window sizes in homogeneous landscapes. Marceau et al. (1990), in a study using SPOT imagery, found that the window size that maximized classification accuracy depended on each specific land-use class and the average window sizes of $17 \times 17$ (340m by 340m) and $25 \times 25$ (500m by 500m) achieved satisfactory classification accuracy for more than one land-use type.
Gong et al. (1992) indicated that windows larger than 7×7 (140m by 140m) yielded unsatisfactory classification accuracy using SPOT images. Gong and Howarth (1992) found that the optimum window size depended on the land-use classes studied. Clearly, using the traditional moving window method, the optimum window size depends on many factors, including pixel resolution, size of land-use polygons, and the homogeneity of land-use classes. Despite the above studies, there are no set rules on how to determine the optimal window size prior to image classification (Hodgson 1998). Gong (1994) used a simple thresholding and region-growing techniques to reduce edge effects. Maillard (2003) suggested extracting edges prior to classification.

Beside the supervised moving-window method, another approach to the incorporation of textural information in the classification is the split-and-merge segmentation algorithms (Ojala and Pietikäinen 1999, Hu et al. 2005, Lucieer et al. 2005), which have been successfully applied. The performance of these procedures depends on the textural measures, various parameters which may greatly differ from one image to another image, and the complexity of the images. These algorithms recursively divide an image into homogeneous regions in terms of textural patterns. To have stable textural measures, these procedures usually impose a minimum size on the sub-blocks and then apply a boundary-refining procedure, which improve the localization of the boundaries to some degree. The performances of these procedures depend on many parameters, which have to be decided experimentally.

From a cognitive perspective, Hodgson (1998) found that classification accuracy from visual analysis conducted by human interpreters increased monotonically with increasing window size, unaffected by edge effects present in purely automated classification. This paper presents a new, automated approach to reducing edge effects in image classification. The performance of the new approach was compared with the traditional moving window approach in three classification scenarios.
2.2 The New Approach

In the traditional method, a moving window of size $m \times n$ (usually $m = n$) is centered on each pixel and textural measures are computed from this window. The class membership of the center pixel is then decided based on a classification rule. Pixels close to the edges of two or more classes tend to be misclassified because of the confusion caused by mixed classes in the window. One problem with the traditional moving-window method is that information collected from the entire moving window is only useful to the center pixel and is not utilized by other pixels in the window. In the new approach, for each window, a similarity index is calculated, which measures the highest degree of similarity between the textural measures of the window and the textural measures of the training samples of all target classes. All pixels in the window will utilize the information and record the similarity index and its corresponding target class. Each pixel, except those near the border, gets $m \times n$ similarity indices since it is included within $m \times n$ windows. From the $m \times n$ similarity indices, each pixel will record the highest similarity index and its corresponding class. As shown below, possible similarity indices include the probability density if the maximum-likelihood classifier is used and the negative or the inverse of the distance if the minimum-distance classifier is used.

Note that in this new method, all pixels in a window obtain the same similarity index from that particular window. Thus, after being included and processed in $m \times n$ overlapping windows, a pixel close to the class boundaries will be more likely to get the highest similarity index from a window completely located within the class it really belongs to. This will surely mitigate the edge effects considerably. This is illustrated in Figure 2.1, in which the center pixel (the black point) is close to the boundary between class A and class B. If the traditional moving window method is used, the pixel will be classified based on texture within window 2 only. Since window 2 contains two classes, error and confusion are likely to arise. But with the proposed method, the pixel will have a much higher chance of being correctly classified into class A since
Figure 2.1. Illustration of edge effects. The pixel (black point) is highly likely to be misclassified using the textural measures in window 2 as it is the center pixel, whereas it has a much better chance of being correctly classified if window 1 is used. The dashed line is the class boundary.
it is included in window 1 and texture in window 1 will have a better chance of being correctly
categorized as class A texture because window 1 contains only class A texture.

The assumption underlying the new approach is that texture containing mixed classes of
pixels is dissimilar to texture containing only one class. The idea comes from the way human
image interpreters deal with boundaries. Humans recognize the dissimilarity between texture
containing mixed classes and texture containing one single class, and a pixel close to the
boundary is classified by humans based on its surrounding homogenous texture instead of by its
surrounding window which may cover two or more classes.

The new approach differs from the traditional method only in the way the similarity index
computed from a window is used. It requires no change in a classifier since the comparison
based on a similarity index is already included in most classifiers. The method does not involve
thresholding either, which has to be decided subjectively.

2.3 A Worked Example

Figure 2.2 shows a worked example to better illustrate the proposed approach using the
minimum-distance classifier. The larger the distance, the more dissimilar between textural
measures of a window and textural measures of a class. Therefore, for the new approach, the
similarity index is chosen as the negative of the distance (the inverse of the distance can also be
used). Figure 2a is a 6×6 image with 3 three uniform classes distinguished by their pixel
values 1, 3 and 5. The mean values of the three classes are used as the textural measure to
discriminate among them. The mean values of a 3×3 moving window, rounded to one decimal
point, are shown in Figure 2b. Figure 2c is the classified image by the traditional moving
window method using the minimum-distance classifier. Four edge pixels, two of which
belonging to class 1 and two of which belonging to class 5, are misclassified as class 3 pixels.
Also note that the pixels on the border of the image are not classified because there is no window
Figure 2.2. A worked example of the proposed approach. (a) a 6×6 image with 3 classes (1, 3, 5); (b) the average image using a 3×3 moving window; (c) classified image using minimum distance classifier by traditional moving window method; (d) the intermediate results of the maximum similarity indices with corresponding class in parentheses when the window is located on row 2 and column 4; (e) the intermediate results when the window is located on row 2 and column 5; (f) the final results of the new approach.
which is centered on them and also falls entirely within the image. Using the new approach, when the moving window is centered at position \( (2,2) \), i.e., the second row and the second column, the window has a mean value of 1 and the mean value is closest to class 1. Therefore, the window obtains from class 1 its highest similarity index of 0, which is the negative of the distance. All nine pixels are assigned this similarity index, as shown in Figure 2d. Since none of them is assigned a similarity index before, there is no comparison in this step. When the window moves to position \( (2,3) \), the window has a mean of 2.3 and the mean is closest to class 3. Therefore, the window obtains from class 3 its highest similarity index of -0.7, which is the negative of the distance. After comparison, the six pixels, which fall within both the current window and the previous window and already have a similarity index of 0, keep 0 as their similarity indices since 0 is larger than -0.7. The remaining 3 pixels in the current window obtain their similarity indices of -0.7 from class 3. Figure 2d also shows the intermediate results when the window moves to position \( (2,4) \). Figure 2e shows the intermediate results after the window moves to position \( (2,5) \). The window has a highest similarity index of 0 with class 5. After comparison, the six pixels, which fall within both the current window and the previous window and already have a similarity index of -0.7 from class 3, obtain new similarity index of 0 from class 5 since 0 is larger than -0.7. Through this comparison, the six pixels obtain similarity indices from a class which they actually belong to and therefore are going to be correctly classified. After the window moves to the last position \( (5,5) \), all pixels are assigned the largest similarity indices from the class they really belong to. The edge effect is automatically eliminated during the process in this artificial example.

### 2.4 Classification Scenarios

In this section, three classification scenarios were presented and compared the results generated by the new approach and the traditional method.
In all three scenarios, three textural measures were used, i.e., \textit{mean}, \textit{standard deviation (std)}, and \textit{entropy}. They were computed using the following formula:

\[ \text{mean} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P(i,j) \]  \hspace{1cm} (1)

\[ \text{std} = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (P(i,j) - \text{mean})^2} \]  \hspace{1cm} (2)

\[ \text{entropy} = -\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Q(i,j) \log[Q(i,j)] \quad (Q(i,j) \neq 0) \]  \hspace{1cm} (3)

where \( Q(i,j) = \frac{|P(i,j)|^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P(i,j)^2} \)  \hspace{1cm} (4)

\( M \) and \( N \) were the height and width of the moving window, and \( P(i,j) \) was the pixel value at position \((i,j)\) within the window.

The minimum-distance classifier and the maximum-likelihood classifier were used in the first two scenarios and the logistic regression was used in the last scenario. For the minimum-distance classifier, the distance \( d \) was calculated as:

\[ d = \sqrt{(x-\overline{x})^2 + (y-\overline{y})^2 + (z-\overline{z})^2} \]

where \( x, y, z \) were the mean, standard deviation and entropy calculated from the moving window, and \( \overline{x}, \overline{y}, \overline{z} \) were the mean values of the three measures for a particular class. As in the worked example, for the new approach, the similarity index was chosen as the negative of the distance, i.e., \(-d\). For the maximum-likelihood classifier, the similarity index in the new approach was the probability density of a multivariate normal distribution:

\[ p(x) = (2\pi)^{-n/2} \left| \Sigma \right|^{-1/2} \exp(-\frac{1}{2} (x - \mu)^\top \Sigma^{-1} (x - \mu)) \]  \hspace{1cm} (5)

where \( x \) was the vector comprising of the three textural measures; \( \mu \) was the mean vector of
textural measures; $\Sigma$ was the variance-covariance matrix of textural measures.

In all three scenarios, one hundred random textural samples for each class were collected to compile the statistics needed for the two classifiers.

### 2.4.1 Classification of an Artificial Image

In the first classification scenario, an artificial image was generated, which consisted of 4 subimages (Figure 2.3a). Each subimage was of size of $100 \times 100$ pixels, generated from a Gaussian distribution with different parameters. The mean and standard deviation of each Gaussian distribution are presented in Table 2.1. Figure 2.3b shows the probability density curves of the four distributions. It is apparent that there was considerable overlapping in the range of pixel values among the four subimages.

Figure 2.4 shows the per-pixel classified image using the maximum-likelihood classifier. The statistics of each class, i.e., mean and standard deviation, were calculated from 100 random pixels. The overall accuracy was 40.00% and the kappa coefficient was 21.34%. It was apparent that pixelwise classification did not provide satisfactory results in this case. However, each subimage showed a distinctive textural pattern, which could be used to distinguish itself from other subimages. Table 2.2 lists the classification accuracies of the image by the two approaches using both the minimum-distance classifier and the maximum classifier with window sizes varying from $3 \times 3$ to $33 \times 33$. The accuracies are also plotted in Figure 2.5. In the case of the traditional approach, the accuracy increased first and then decreased steadily. The cutoff point of the window size was 9. This reflects the common dilemma in textural analysis. Small windows are associated with unstable textural measures while large windows are associated with large confusion along class boundaries. In comparison, the overall accuracy achieved by the new approach generally increased as the window size increased and leveled off when the window size was large enough (larger than $15 \times 15$). For the two smallest window sizes ($3\times3$ and $5\times5$), the
Figure 2.3. (a) an artificial image consisting of four subimages generated from different Gaussian distributions; (b) probability density curves of the four Gaussian distributions.
Table 2.1. Means and standard deviations of 4 Gaussian distributions used to generate subimages.

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>Top Left</td>
<td>150</td>
<td>10</td>
</tr>
<tr>
<td>Class B</td>
<td>Top Right</td>
<td>160</td>
<td>20</td>
</tr>
<tr>
<td>Class C</td>
<td>Bottom Left</td>
<td>170</td>
<td>30</td>
</tr>
<tr>
<td>Class D</td>
<td>Bottom Right</td>
<td>180</td>
<td>40</td>
</tr>
</tbody>
</table>
Figure 2.4. Pixelwise classified image using maximum-likelihood classifier without using textural patterns with an overall accuracy of 40%. 
Table 2.2. Overall classification accuracy of the artificial image (%).

<table>
<thead>
<tr>
<th>Window size</th>
<th>Minimum-distance classifier</th>
<th>Maximum-likelihood classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional method</td>
<td>New method</td>
</tr>
<tr>
<td>3</td>
<td>75.48</td>
<td>77.98</td>
</tr>
<tr>
<td>5</td>
<td>90.46</td>
<td>94.01</td>
</tr>
<tr>
<td>7</td>
<td>94.81</td>
<td>97.98</td>
</tr>
<tr>
<td>9</td>
<td>96.62</td>
<td>98.87</td>
</tr>
<tr>
<td>11</td>
<td>96.54</td>
<td>98.51</td>
</tr>
<tr>
<td>13</td>
<td>96.21</td>
<td>98.65</td>
</tr>
<tr>
<td>15</td>
<td>95.75</td>
<td>98.19</td>
</tr>
<tr>
<td>17</td>
<td>95.03</td>
<td>99.28</td>
</tr>
<tr>
<td>19</td>
<td>94.39</td>
<td>99.18</td>
</tr>
<tr>
<td>21</td>
<td>93.83</td>
<td>98.97</td>
</tr>
<tr>
<td>23</td>
<td>93.27</td>
<td>99.38</td>
</tr>
<tr>
<td>25</td>
<td>92.57</td>
<td>99.12</td>
</tr>
<tr>
<td>27</td>
<td>91.81</td>
<td>99.33</td>
</tr>
<tr>
<td>29</td>
<td>91.11</td>
<td>99.07</td>
</tr>
<tr>
<td>31</td>
<td>90.28</td>
<td>99.13</td>
</tr>
<tr>
<td>33</td>
<td>89.41</td>
<td>99.33</td>
</tr>
</tbody>
</table>
Figure 2.5. Classification accuracy with different window sizes for the artificial image by the traditional and the new methods. (a) using the minimum-distance classifier; (b) using the maximum-likelihood classifier.
new approach achieved lower accuracy than the traditional approach when the maximum-likelihood classifier was used. It could be explained that small window size did not provide stable texture measures, causing the maximum-likelihood classifier unable to discriminate successfully among different classes. The minimum-distance classifier was less affected by this factor because it utilized only the means calculated from the samples while the maximum-likelihood classifier made use of not only the means but also the variances and covariances.

Figure 2.6 shows the classification results of the two approaches with a window size of 33 by 33 by the two classifiers. Some observations can be made. First, the traditional method leaf a strip of pixels along the border of the image unclassified using both classifiers. The width of the strip was approximately half the window size. The existence of the strip was due to the fact that windows centered upon the pixels in the strip had parts outside the image region. It was worth noting that the overall accuracy used here was based upon pixels not in the strip. For the new approach, all pixels were classified and the overall accuracy was based upon all pixels in the image. Second, the confusion along the class boundaries was much larger for the traditional approach than for the new approach. For the traditional approach, the confusion took place mainly along the horizontal middle part. In the middle left part, pixels tended to be misclassified as belonging to class B. This was because the mean and standard deviation of class B were between those of class A and class C, and a window along the boundary of class A and C tended to have a mean and standard deviation close to the averages of classes A and C, which were then similar to those of class B. In the middle right part, the confusion arised for similar reasons. For the new approach, the confusion was much less and could be mainly attributed to the random fluctuations in the texture. Both classifiers achieved very satisfactory results using the new approach.
Figure 2.6. Classified results of the artificial image with a window size of 33×33. Top left: the by minimum-distance classifier using the traditional method. Bottom left: by minimum-distance classifier using the new method. Top right: by the maximum-likelihood classifier using the traditional method. Bottom right: by the maximum-likelihood classifier using the new method.
2.4.2 Classification of a Mosaic Image

A mosaic image (Figure 2.7) was created, which included six land-use samples: commercial, industrial, water, single-family, multifamily, and forest land uses. Each land-use sample was of size of 200 by 200 pixels. These land-use samples were taken from the panchromatic band of an IKONOS image, which covered part of metropolitan Atlanta, Georgia. The IKONOS image was taken on 29 December 2000. The projection was Universal Transverse Mercator, Zone 16. The datum was WGS84.

Table 2.3 lists the overall accuracies by the two approaches using the minimum-distance classifier and the maximum-likelihood classifier with window size varying from 17 to 99. Figure 2.8a shows the overall classification accuracy by the two approaches using the minimum-distance classifier. Because of the wide variation within each class, the window size began at 17. For the traditional method, the overall accuracy by the traditional method increased as the window size increased, reaching a maximum of about 82% at the window size of 67×67, then dropped gradually by a small amount as the window size further increased. For the new approach, the overall accuracy generally increased as the window size increased until it leveled off. There were only a few minor exceptions, which deviated from the trend only by a negligible amount. The new approach achieved higher accuracy than the traditional method irrespective of the window size used.

Figure 2.8b presents the overall classification accuracy by the two approaches using the maximum-likelihood classifier. When the window size was less than 55×55, the traditional approach produced higher accuracy than the new approach. This result was different that using the minimum-distance classifier. It could be explained that small window sizes do not provide stable texture measures, causing the maximum-likelihood classifier unable to discriminate successfully among different classes. As a result, the new approach yielded lower accuracy than the traditional method. When the window size was larger than 55×55, the traditional approach
Figure 2.7. Mosaic image of six land use classes from the panchromatic band of an IKONOS image of Atlanta, GA. Top: commercial, industrial, and water. Bottom: single-family, multifamily, and forest.
Table 2.3 Overall classification accuracy of the mosaic image (%)

<table>
<thead>
<tr>
<th>Window size</th>
<th>Minimum-distance classifier</th>
<th></th>
<th></th>
<th>Maximum-likelihood classifier</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>66.47</td>
<td>68.26</td>
<td>68.13</td>
<td>61.25</td>
<td>63.60</td>
<td>67.06</td>
</tr>
<tr>
<td>19</td>
<td>67.50</td>
<td>70.97</td>
<td>70.81</td>
<td>66.37</td>
<td>69.16</td>
<td>69.48</td>
</tr>
<tr>
<td>21</td>
<td>68.53</td>
<td>71.49</td>
<td>72.18</td>
<td>66.35</td>
<td>69.52</td>
<td>69.52</td>
</tr>
<tr>
<td>23</td>
<td>70.30</td>
<td>71.49</td>
<td>73.35</td>
<td>68.30</td>
<td>71.24</td>
<td>71.24</td>
</tr>
<tr>
<td>25</td>
<td>71.32</td>
<td>72.02</td>
<td>74.26</td>
<td>69.25</td>
<td>72.29</td>
<td>72.29</td>
</tr>
<tr>
<td>27</td>
<td>72.25</td>
<td>72.81</td>
<td>74.60</td>
<td>68.38</td>
<td>70.18</td>
<td>70.18</td>
</tr>
<tr>
<td>29</td>
<td>73.18</td>
<td>74.71</td>
<td>75.12</td>
<td>69.84</td>
<td>70.18</td>
<td>70.18</td>
</tr>
<tr>
<td>31</td>
<td>74.04</td>
<td>76.32</td>
<td>75.66</td>
<td>70.18</td>
<td>70.18</td>
<td>70.18</td>
</tr>
<tr>
<td>33</td>
<td>74.80</td>
<td>78.12</td>
<td>76.44</td>
<td>70.94</td>
<td>70.18</td>
<td>70.18</td>
</tr>
<tr>
<td>35</td>
<td>75.43</td>
<td>79.33</td>
<td>77.01</td>
<td>71.24</td>
<td>72.29</td>
<td>72.29</td>
</tr>
<tr>
<td>37</td>
<td>76.33</td>
<td>81.25</td>
<td>77.92</td>
<td>72.29</td>
<td>72.29</td>
<td>72.29</td>
</tr>
<tr>
<td>39</td>
<td>77.25</td>
<td>82.84</td>
<td>78.36</td>
<td>72.09</td>
<td>72.09</td>
<td>72.09</td>
</tr>
<tr>
<td>41</td>
<td>77.86</td>
<td>84.51</td>
<td>78.79</td>
<td>74.23</td>
<td>74.23</td>
<td>74.23</td>
</tr>
<tr>
<td>43</td>
<td>78.49</td>
<td>84.89</td>
<td>79.14</td>
<td>74.91</td>
<td>74.91</td>
<td>74.91</td>
</tr>
<tr>
<td>45</td>
<td>78.90</td>
<td>86.12</td>
<td>79.58</td>
<td>76.80</td>
<td>76.80</td>
<td>76.80</td>
</tr>
<tr>
<td>47</td>
<td>79.42</td>
<td>86.70</td>
<td>80.01</td>
<td>76.49</td>
<td>76.49</td>
<td>76.49</td>
</tr>
<tr>
<td>49</td>
<td>80.04</td>
<td>86.96</td>
<td>80.50</td>
<td>78.52</td>
<td>78.52</td>
<td>78.52</td>
</tr>
<tr>
<td>51</td>
<td>80.49</td>
<td>87.75</td>
<td>80.81</td>
<td>80.06</td>
<td>80.06</td>
<td>80.06</td>
</tr>
<tr>
<td>53</td>
<td>80.97</td>
<td>87.86</td>
<td>80.87</td>
<td>81.18</td>
<td>81.18</td>
<td>81.18</td>
</tr>
<tr>
<td>55</td>
<td>81.29</td>
<td>88.80</td>
<td>80.88</td>
<td>83.15</td>
<td>83.15</td>
<td>83.15</td>
</tr>
<tr>
<td>57</td>
<td>81.45</td>
<td>89.55</td>
<td>80.75</td>
<td>83.89</td>
<td>83.89</td>
<td>83.89</td>
</tr>
<tr>
<td>59</td>
<td>81.64</td>
<td>90.80</td>
<td>80.53</td>
<td>85.44</td>
<td>85.44</td>
<td>85.44</td>
</tr>
<tr>
<td>61</td>
<td>81.84</td>
<td>91.34</td>
<td>80.40</td>
<td>87.63</td>
<td>87.63</td>
<td>87.63</td>
</tr>
<tr>
<td>63</td>
<td>82.08</td>
<td>91.95</td>
<td>80.15</td>
<td>87.74</td>
<td>87.74</td>
<td>87.74</td>
</tr>
<tr>
<td>65</td>
<td>82.18</td>
<td>92.34</td>
<td>79.78</td>
<td>87.92</td>
<td>87.92</td>
<td>87.92</td>
</tr>
<tr>
<td>67</td>
<td>82.11</td>
<td>92.38</td>
<td>79.37</td>
<td>89.83</td>
<td>89.83</td>
<td>89.83</td>
</tr>
<tr>
<td>69</td>
<td>82.15</td>
<td>92.44</td>
<td>78.94</td>
<td>90.16</td>
<td>90.16</td>
<td>90.16</td>
</tr>
<tr>
<td>71</td>
<td>81.98</td>
<td>93.17</td>
<td>78.48</td>
<td>92.14</td>
<td>92.14</td>
<td>92.14</td>
</tr>
<tr>
<td>73</td>
<td>81.74</td>
<td>94.46</td>
<td>78.02</td>
<td>92.32</td>
<td>92.32</td>
<td>92.32</td>
</tr>
<tr>
<td>75</td>
<td>81.50</td>
<td>95.65</td>
<td>77.56</td>
<td>92.84</td>
<td>92.84</td>
<td>92.84</td>
</tr>
<tr>
<td>77</td>
<td>81.33</td>
<td>96.14</td>
<td>77.13</td>
<td>93.16</td>
<td>93.16</td>
<td>93.16</td>
</tr>
<tr>
<td>79</td>
<td>81.14</td>
<td>95.99</td>
<td>76.63</td>
<td>94.39</td>
<td>94.39</td>
<td>94.39</td>
</tr>
<tr>
<td>81</td>
<td>80.91</td>
<td>95.92</td>
<td>76.08</td>
<td>94.53</td>
<td>94.53</td>
<td>94.53</td>
</tr>
<tr>
<td>83</td>
<td>80.78</td>
<td>95.93</td>
<td>75.48</td>
<td>94.85</td>
<td>94.85</td>
<td>94.85</td>
</tr>
<tr>
<td>85</td>
<td>80.61</td>
<td>95.76</td>
<td>74.90</td>
<td>95.19</td>
<td>95.19</td>
<td>95.19</td>
</tr>
<tr>
<td>87</td>
<td>80.46</td>
<td>95.70</td>
<td>74.29</td>
<td>95.36</td>
<td>95.36</td>
<td>95.36</td>
</tr>
<tr>
<td>89</td>
<td>80.23</td>
<td>95.46</td>
<td>73.69</td>
<td>95.79</td>
<td>95.79</td>
<td>95.79</td>
</tr>
<tr>
<td>91</td>
<td>80.02</td>
<td>95.55</td>
<td>73.05</td>
<td>95.98</td>
<td>95.98</td>
<td>95.98</td>
</tr>
<tr>
<td>93</td>
<td>79.74</td>
<td>94.94</td>
<td>72.41</td>
<td>95.95</td>
<td>95.95</td>
<td>95.95</td>
</tr>
<tr>
<td>95</td>
<td>79.43</td>
<td>94.73</td>
<td>71.77</td>
<td>96.02</td>
<td>96.02</td>
<td>96.02</td>
</tr>
<tr>
<td>97</td>
<td>79.11</td>
<td>94.63</td>
<td>71.11</td>
<td>96.03</td>
<td>96.03</td>
<td>96.03</td>
</tr>
</tbody>
</table>
Figure 2.8. Classification accuracy with different window sizes for the artificial image by the traditional and the new methods. (a) using the minimum-distance classifier; (b) using the maximum-likelihood classifier.
yielded lower and lower accuracy as the window size increased whereas for the new approach, the overall accuracy kept increasing until it levels off. When the window size was large enough (larger than 55×55) to generate stable measures, the edge effect was more prominent by the maximum-likelihood classifier than by the minimum-distance classifier, as evidenced by the lower accuracy achieved by the maximum-likelihood classifier using the traditional moving window method. This was on the grounds that the maximum-likelihood classifier performed better on stable textural measures than the minimum-distance classifier.

Figure 2.9 shows the classified images for a window size of 85×85. For both classifiers, it can be seen that the new approach reduced edge effects substantially. The edges between water and forest, water and industrial land use, multifamily and forest land uses were clear by the new approach. With the maximum likelihood classifier, the boundaries between all classes were very close to the true boundaries using the new approach. Some errors arose, which could be ascribed to the considerable variation shown in the mosaic image.

For the new approach, the smoothing effect was clear in both cases. Some mall isolated clusters of pixels, present in the two images by the traditional method, were replaced with surrounding large classes. This smoothing effect was caused by the way the new approach utilized the textural information in the moving window. The small isolated clusters were large deviations from the surrounding pixels in terms of textural patterns they presented. An entire small cluster could be removed if its separate parts were removed by many neighboring moving windows using the criterion of highest similarity index. The smoothing effect was generally preferred because the smoothed map would easily be converted into areal units in vector format in subsequent analysis. In many traditional post-classification processing, a majority operation is applied to reduce the salt-and-pepper pattern and to obtain a smoothed map.
Figure 2.9. Classified results of the mosaic image with a window size of 85×85. Top left: the by minimum-distance classifier using the traditional method. Bottom left: by minimum-distance classifier using the new method. Top right: by the maximum-likelihood classifier using the traditional method. Bottom right: by the maximum-likelihood classifier using the new method.
2.4.3 Identification of Non-agricultural Land Use by Logistic Regression

In the third scenario, an IKONOS image, acquired on 18 May 2001, covering part of rural northeastern Alabama was used. The projection was Universal Transverse Mercator, Zone 16. The datum was WGS84. A subset of the image (Figure 2.10) was extracted (600x300 pixels) and the subset contained two types of land covers, agricultural land use and non-agricultural land use. The agricultural land had a smooth pattern while the non-agricultural land had a rough and coarse pattern. The purpose of this experiment was to delineate the area of non-agricultural land use from agricultural land use. In the first two scenarios, the shape of the subimages matched the shape of the window, which might help to achieve favorable results. In this scenario, the class boundaries were irregular.

In this experiment, logistic regression (McCullagh and Nelder 1989, Pampel 2000) was used to distinguish between the two types of land uses. Logistic regression is a useful tool when the dependent variable is dichotomous where the ordinary least squares regression is inappropriate. In logistic regression, the logit, which is the logarithm of the odds of an event (the ratio of the probability that an event occurs to the probability that it fails to occur), is treated as a latent variable and assumed to have a linear relationship with the input data:

\[
\text{logit} = \log \frac{p}{1-p} = a + \tilde{b}^\top \tilde{x}
\]  \hspace{1cm} (6)

so that:

\[
p = \frac{\exp(a + \tilde{b}^\top \tilde{x})}{1 + \exp(a + \tilde{b}^\top \tilde{x})}
\]  \hspace{1cm} (7)

where \(a\) is a constant; \(\tilde{b}\) is the parameter vector; and \(\tilde{x}\) is the input vector, comprising of the three textural measures in this experiment. Based on training data, maximum-likelihood estimates of the parameters (\(a\) and \(\tilde{b}\)), in the above equation can be obtained through an iterative process by the Newton-Raphson algorithm (McCullagh and Nelder 1989, Pampel 2000).
Figure 2.10. A subset of IKONOS panchromatic band with agricultural and non-agricultural land uses.
The parameters can then be used to compute the probability of an event. Figure 2.11 plots the curve of the probability versus the logit. Note that the logit can vary from negative infinity to positive infinity, whereas the probability can only vary between 0 and 1 due to the functional relationship between the two variables.

The maximum-likelihood estimates of the parameters are derived in the following way. Given the augmented predictor data matrix:

\[
X = \begin{pmatrix}
1 & x_{11} & x_{12} & \cdots & x_{1p} \\
1 & x_{21} & x_{22} & \cdots & x_{2p} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_{N1} & x_{N2} & \cdots & x_{Np}
\end{pmatrix}
\]

and the response vector

\[
y = \begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_N
\end{pmatrix}
\]

where \( y_i \) (i=1 to N) takes the value of either 0 or 1. The probability of presence (taking value 1) of the response for a given \( x \) is

\[
p = \frac{\exp(\beta'x)}{1 + \exp(\beta'x)}
\]

and the probability of absence (taking value 0) of the response is

\[
p = 1 - \frac{\exp(\beta'x)}{1 + \exp(\beta'x)} = \frac{1}{1 + \exp(\beta'x)}
\]

Then,

\[
\log(p) = y \log(p) + (1-y) \log(1-p) \\
= y(\beta'x - \log(1+\exp(\beta'x))) - (1-y) \log(1+\exp(\beta'x)) \\
= y\beta'x - \log(1+\exp(\beta'x))
\]

for either \( y = 0 \) or \( y = 1 \).

The log-likelihood function is

\[
l(\beta) = \sum_{i=1}^{N} \log(p_i) = \sum_{i=1}^{N} [y_i\beta'x_i - \log(1+\exp(\beta'x_i))].
\]
Figure 2.11 Function curve between the probability and the logit
maximize $l(\beta)$, the Newton-Raphson algorithm is used, i.e.,

$$\Delta \beta = \left( \frac{\partial^2 l}{\partial \beta \partial \beta'} \right)^{-1} \frac{\partial l}{\partial \beta}.$$ 

\[
\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{N} y_i x_{ij} - \sum_{i=1}^{N} x_i e^{\beta x_i} \\
= \sum_{i=1}^{N} y_i x_{ij} - \sum_{i=1}^{N} p_i x_{ij} \\
= \sum_{i=1}^{N} x_{ij} (y_i - p_i)
\]

and in matrix form, \[
\frac{\partial l}{\partial \beta} = \sum_{i=1}^{N} x_i (y_i - p_i) = X'(y - p), \text{ where } p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{pmatrix}.
\]

\[
\frac{\partial^2 l}{\partial \beta_j \partial \beta_k} = -\sum_{i=1}^{N} \frac{(1 + e^{\beta x_i}) e^{\beta x_i} x_{ij} x_{ik} - (e^{\beta x_i})^2 x_{ij} x_{ik}}{(1 + e^{\beta x_i})^2} \\
= -\sum_{i=1}^{N} x_{ij} x_{ik} p_i - x_{ij} x_{ik} p_i^2 \\
= -\sum_{i=1}^{N} x_{ij} x_{ik} p_i (1 - p_i)
\]

and in matrix form, \[
\frac{\partial^2 l}{\partial \beta \partial \beta'} = -\sum_{i=1}^{N} x_{ij} x_{ij}' p_i (1 - p_i) = -X'WX,
\]

where $W$ is a diagonal matrix $W = \begin{pmatrix}
p_1(1 - p_1) & 0 & \cdots & 0 \\
0 & p_1(1 - p_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & p_N(1 - p_N)
\end{pmatrix}$.

Thus, \[
\Delta \beta = -\left( \frac{\partial^2 l}{\partial \beta \partial \beta'} \right)^{-1} \frac{\partial l}{\partial \beta} = (X'WX)^{-1} X'(y - p).
\]

The iterative process stops until some predetermined stopping criteria are met.

Logistic regression has been successfully applied to map one particular type of phenomenon across an entire image. Pu and Gong (2004) compared the performance of logistic regression and
artificial neural networks in the detection of burned scars from post-fire Landsat 7 Enhanced Thematic Mapper plus (ETM+) images in mountainous areas of northern California. They concluded that the logistic regression was more efficient than the neural network model in predicting burned scars. Pearlstine et al. (2005) derived textural features from first and second-order statistics and edge components in high-resolution digital color infrared images to test their ability to discriminate an exotic, invasive plant in Florida. Logistic regression found a best subset combination of textural features that consistently identified core areas of the exotic, invasive plant.

In this experiment, the non-agricultural class was treated as the event. So the probability of 1 was indicative of complete non-agricultural texture and the probability of 0 was indicative of complete agricultural texture. The similarity index was: p-1 if p >=0.5 or -p if p<0.5. The value of 0.5 was used as the cutoff point. Zero is the largest possible similarity index. This similarity index was the negative of distance in terms of probability.

Table 2.4 lists the overall accuracies using the two approaches with window sizes varying from 17×17 to 51×51. The accuracies are also plotted in Figure 2.12. As in previous scenarios, the new approach achieved higher accuracy over the traditional method. As the window size increased, the traditional method generally produced lower and lower accuracy due to the edge issue whereas the accuracy by the new approach increased until it levels off.

The classified images for the non-agricultural land use with the window size of 33×33 are shown in Figure 2.13. It can be seen that for the traditional method, the commission errors largely occurred along the entire boundary between the agricultural and non-agricultural land uses, with some commission errors occurring as one small isolated clump. For the new method, the isolated small cluster of pixels was gone because of the smoothing effect discussed above. For the new method, the commission errors largely occurred on the top-middle part of the image. A close examination of the top-middle part of the image revealed that there were some bright
<table>
<thead>
<tr>
<th>Window size</th>
<th>Traditional method</th>
<th>New method</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>94.35</td>
<td>95.72</td>
</tr>
<tr>
<td>19</td>
<td>94.46</td>
<td>96.39</td>
</tr>
<tr>
<td>21</td>
<td>94.15</td>
<td>96.78</td>
</tr>
<tr>
<td>23</td>
<td>94.04</td>
<td>97.07</td>
</tr>
<tr>
<td>25</td>
<td>94.35</td>
<td>97.72</td>
</tr>
<tr>
<td>27</td>
<td>94.11</td>
<td>97.60</td>
</tr>
<tr>
<td>29</td>
<td>94.01</td>
<td>97.80</td>
</tr>
<tr>
<td>31</td>
<td>94.06</td>
<td>98.27</td>
</tr>
<tr>
<td>33</td>
<td>93.84</td>
<td>98.56</td>
</tr>
<tr>
<td>35</td>
<td>93.51</td>
<td>98.48</td>
</tr>
<tr>
<td>37</td>
<td>93.20</td>
<td>98.26</td>
</tr>
<tr>
<td>39</td>
<td>92.70</td>
<td>98.44</td>
</tr>
<tr>
<td>41</td>
<td>92.54</td>
<td>98.63</td>
</tr>
<tr>
<td>43</td>
<td>92.20</td>
<td>98.54</td>
</tr>
<tr>
<td>45</td>
<td>91.92</td>
<td>98.51</td>
</tr>
<tr>
<td>47</td>
<td>91.73</td>
<td>98.45</td>
</tr>
<tr>
<td>49</td>
<td>91.50</td>
<td>98.42</td>
</tr>
<tr>
<td>51</td>
<td>91.27</td>
<td>98.38</td>
</tr>
</tbody>
</table>
Figure 2.12. Overall classification accuracy using logistic regression.
Figure 2.13. Top: Classified image overlaid on the reference map for non-agricultural land use with the traditional method (window size $33 \times 33$). Bottom: Classified image overlaid on the reference map for non-agricultural land use with the new method.
pixels standing out from surrounding pixels in the agricultural land use. This large anomaly might yield textural measures that were close to those of non-agricultural land use, thus causing errors. For the traditional method, the omission errors occurred as four small clusters in the image (the pixels along the border were not counted). But, three clusters were gone in the classified image by the new approach due to the smoothing effect. The new approach produced some small omission errors along the boundary.

Two computational considerations are noted here if logistic regression is used in this application as the basis for delineating dichotomous land covers. First, the maximum-likelihood estimates do not always exist for logistic regression (Albert and Anderson 1984, Santner and Duffy 1986). In that case, the logistic model is questionable and should not be used. In the experiments, a convergence could not be found when the window size was larger than 51×51. Although popular statistical packages usually continue the process despite the failure to achieve convergence, the model, if used, may give unexpected results for both the traditional and the new methods.

Second, if the logit values are large enough (say larger than 50), the probability calculated from $\frac{\exp(\text{logit})}{1 + \exp(\text{logit})}$ will always be 1 due to the precision limits in the computer for two different logit values even when there is a noticeable difference between them. The same is true when the logit is too small and the probability will always be 0. As a result, a pixel along the boundary may be assigned the largest similarity index of zero when most of the moving window is agricultural and get the same largest similarity index of zero when most of the moving window is non-agricultural. In this case, the pixel has the chance of getting two maximum indices from two different categories and the order in which the pixel gets the indices affects the results. Since the probability increases monotonously as logit increases, this problem is eliminated if the logit values are compared directly when two logit values have the same sign and compare the logit
values directly when two logits have the same sign. When two logits have different signs, a little
calculation shows that the sum of the two logits only needs to be compared with zero. The logit
generally does not exceed the precision limits in the computer and can be compared stably.

2.5 Discussion

From the above experiments, it was shown that the new approach had the following
characteristics. First, the new approach consistently achieved higher accuracy with slight
fluctuations until it leveled off, whereas the accuracy of the traditional approach generally
decreased with increasing window size after the window size passed a threshold. Second, the
new approach was capable of classifying border pixels, while the traditional method left a strip
of border pixels unprocessed. Third, the new approach had a smoothing effect. Small isolated
clusters, present in classified images by the traditional method, tended to be eliminated in
classified images by the new approach. Fourth, the side length of the window could be an even
number in the new approach because it does not use the concept of the center of the window,
whereas in the traditional approach, the window size is usually odd to make sure the center pixel
is the exact center.

For the new approach to be effective, the following requirements should be met. First,
training samples should be representative as in any supervised classification scheme. Second, the
texture measures used to determine class membership should be able to discriminate between
homogenous samples and samples including more than one land-cover classes. This is important
because the underlying assumption of the new approach is that texture containing mixed classes
is dissimilar to texture containing one single class. If mixed texture yields a similarity index
higher than that by homogeneous texture, errors will arise along class boundaries because part of
pixels in the mixed texture will be misclassified. Third, similarity index should be comparable
among different windows, such as the probability density in the maximum-likelihood classifier
and the negative or the reverse of the distance in the minimum-distance classifier used in this study. In comparison, if the linear discriminant analysis is used, the discriminant score should not be used as the similarity index, but rather the probability values derived from the discriminant scores should be used because the linear discriminant scores are incomparable among different observations while the probability values are (Tatsuoka and Lohnes 1988, p369).

In this Chapter, the experiments were carried out on only the panchromatic band of the IKONOS images. However, the proposed method can easily be extended to multi-spectral imagery, where the similarity index can be calculated based on textural measures obtained from each individual band. The complexity will also be greatly increased as the dimension increases. The proposed approach could easily work together with other textural measures and classifiers.

The new approach is not a panacea. For pixels that have no chance of entirely falling in a window containing only one type of texture, the edge issues tend to persist and the pixels tend to be misclassified. This happens when the class boundaries are too complex or the land-cover polygons are smaller than the window size used. This problem is inherent to regularly shaped windows. No single one geometric shaped window provides universally best results for all boundary types. Although dynamic windows with changing shape and size according to local structures had been suggested as a possible solution (Hodgson 1998), the irregularity of the window shapes and sizes will make comparisons of their textural measures inconsistent and unstable, leading to unexpected results and difficult interpretation. When there are a large number of classes in an image, particularly in large high-resolution images, the class polygon size varies substantially from one class to another class and the probability of errors may increase in a non-monotonous fashion with the increase in the number of classes.
2.6 Conclusion

This Chapter proposes a new approach to reducing the edge effects in image classification. In the new approach, all pixels in a moving window make use of the textural information instead of only the center pixel as in the traditional moving window method. Results from three classification scenarios using one artificial image, one mosaic image, and one natural scene, show that the new approach generally produced higher accuracy with increasing window size and was much less affected by edge issues than the traditional moving window method. The new approach yielded satisfactory results as long as the window size does not exceed the size of the land cover polygons and the class boundaries are not too complex.
Chapter 3
Improving High-Resolution Image Segmentation by Combining Segmentations Using Different Textural Measures

3.1 Introduction

Image segmentation is the process of partitioning an image into several homogeneous regions with respect to some characteristics (Chen and Chen 2002). Automatic segmentation of high-resolution remotely sensed imagery is a potentially useful information extraction method (Hu et al. 2005). Once an image is correctly segmented into homogeneous areas, each region may then be assigned a meaningful label corresponding to its land use/land cover on the ground.

Compared with two of the most commonly used clustering procedures, i.e., Iterative Self-Organizing Data Analysis Technique (ISODATA) and K-means (Jensen 2005), image segmentation seeks the explicit goal of spatial contiguity, whereas it is only implicit in ISODATA or K-means procedures (Lucieer et al. 2005). The ISODATA and K-means procedures based solely upon pixel values are hardly expected to yield satisfactory results when they are applied to high-resolution remotely sensed images because of the great variations within each region.

Segmentation approaches can be classified as region-based or boundary-based (Haralick and Shapiro 1985, Reed and Du buf 1993, Pal and Pal, 1993). Textural pattern provides valuable information for the identification of objects in remote sensing images. Texture-based image segmentation holds the promise of dealing with the identification of objects in high-resolution remotely sensed imagery more successfully than boundary-based method.

In this Chapter, the region-based segmentation framework proposed by Ojala and Pietikainen (1999) and modified by Chen and Chen (2002) was used to test the performance of using different textural measures in segmenting remotely sensed images. The segmentations by different textural measures were then overlaid to make use of the results by different measures.
In doing so, clusters identified by one textural measure but not identified by another measure would be present in the final segmentation. This post-segmentation integration of results from different textural measures differs from pre-segmentation weighting scheme for textural measures (Hu et al. 2005) in that no weights are subjectively decided.

3.2 Textural Description

Characterization and discrimination of textural patterns are probably the most critical factor for the segmentation procedures to succeed (Ojala et al. 1996). A large body of research characterizes texture by single values such as means, standard deviations, etc. In doing so, the distribution of the measures is not utilized, which may be important for the successful discrimination of texture. In the segmentation methods to be described in the following section, researchers (Ojala and Pietikainen 1999, Chen and Chen 2002, Hu et al. 2005, Lucieer et al. 2005) made use of the measure distributions to segment images.

3.2.1 Histogram

The histogram of pixel values describes the distribution of spectral response patterns. As with the training stage of traditional pixelwise classification schemes, the histogram can provide valuable information to distinguish among different textural patterns. For multispectral images, the computation of the histogram can be based on the averages of all bands (Hu et al. 2005).

3.2.2 Local Binary Pattern and Contrast Measure

The local binary pattern (LBP) is proposed by Ojala et al. (1996) and has been shown to be efficient for textural discrimination. A 3×3 moving window (Figure 3.1a) is used. The pixel values within the window is compared with the center pixel value to get a binary neighborhood (Figure 3.1b), which are then multiplied by binominal weights (Figure 3.1c) given to the corresponding pixel. The resultant values (Figure 3.1d) are summed up to obtain the LBP value.

LBP depends on the relative values within the window. It describes the spatial structure of
Figure 3.1 Computation of local binary pattern and contrast measure. (a) An artificial image of 3 by 3 pixels. (b) Image of binary neighbor relative to the center. (c) Map of binomial weights. (d) Product of the weights and the binary values.

\[
\begin{align*}
\text{LBP} &= 2 + 8 + 64 + 128 = 202 \\
C &= (6 + 5 + 9 + 8)/4 - (1 + 4 + 3 + 2)/4 = 3.5
\end{align*}
\]
the local texture but not the contrast of the texture (Ojala and Pietikainen 1999). A contrast measure (C) is can be combined with LBP to jointly describe local patterns. C can be defined simply as the difference between the average value of the pixels which are no less than the center pixel and the average value of the pixels of the pixels which are smaller than the center pixel.

The LBP/C joint distribution is obtained after the window moves across a textural region. Note that the LBP/C joint distribution is intensity invariant. A histogram of the intensity values in this case can provide additional useful information. For multispectral images, the computation can be based on the averages of all bands (Hu et al. 2005).

The distribution can be approximated using a two-dimensional discrete histogram of size $256 \times b$, where $b$ is the number of bins for contrast measure C. $b$ is chosen experimentally such as 8 or 16.

The LBP/C measure is independent of the absolute intensity values, but dependent on rotation.

### 3.2.3 Saturation/Hue Distribution

Color is a great help for human interpreters to identify different regions. A joint distribution of saturation and hue can be used to characterize colors (Hu et al. 2005).

The saturation and hue can be calculated from the R, G, B component of an image as:

$$ Sat = \max(R, G, B) - \min(R, G, B) $$

$$ Hue = \arctan \left( \frac{\sqrt{3} (G - B)}{(R - G) + (R - B)} \right) $$

Hue values range from 0 to 360 degrees. The distribution can be approximated using a two-dimensional discrete histogram of size $8 \times 60$.

### 3.3 Similarity Index of Measure Distributions

Given two measure distributions $h_1(i)$ and $h_2(i)$, where $i$ can vary between 1 and the number of bins used in the calculation of the histogram, there exist some methods to compute the degree
of similarity or dissimilarity between them.

### 3.3.1 Histogram Intersection

This method (Swain and Ballard 1991) compares the similarity measure between two histograms using the following formula:

\[
I = \sum \min(h_1(i), h_2(i))
\]

\(h_1\) and \(h_2\) should be normalized by their sums respectively. The larger this index is, the more similar the two histograms are.

### 3.3.2 Histogram Correlation

The correlation coefficient between two \(h_1(i)\) and \(h_2(i)\) can be used as a similarity measure (Hu et al. 2005):

\[
I = \frac{\sum (h_1(i) - \overline{h_1})(h_2(i) - \overline{h_2})}{\sqrt{\sum (h_1(i) - \overline{h_1})^2 \sum (h_2(i) - \overline{h_2})^2}}
\]

where \(\overline{h_1}\) and \(\overline{h_2}\) are the means of the two histograms. The larger the correlation coefficient is, the more similar the two histograms are.

### 3.3.3 G-statistic

The G-statistic (Sokal and Rohlf 1987) is calculated as:

\[
I = 2(\sum_i h_1(i) \log(h_1(i)) + \sum_i h_2(i) \log(h_2(i))
- \sum_i (h_1(i) + h_2(i)) \log(h_1(i) + h_2(i))
- (\sum_i h_1(i) \log(\sum_i h_1(i)) + \sum_i h_2(i) \log(\sum_i h_2(i)))
+ \sum_i (h_1(i) + h_2(i)) \log(\sum_i (h_1(i) + h_2(i))))
\]

The G-statistic is a measure of the probability of the two sample distributions taken from the same population. The larger the G-statistic is, the less similar the two histograms are. Unlike previous two similarity measures, the G-statistic is actually a measure of dissimilarity between two histograms and can be much larger than 1.
3.4 Segmentation Procedures

In the method proposed by Ojala and Pietikainen (1999) and modified by Chen and Chen (2002), the segmentation involves three steps: hierarchical splitting, agglomerative merging, and pixelwise refinement. Figure 3.2 gives an example illustrating the procedures. The composite image was composed of five Brodatz (1966) texture samples with distinct textural patterns.

3.4.1 Hierarchical Splitting

In this step, an image is recursively split into square blocks of varying size (Figure 3.2b). In the first place, the image is divided into square blocks of size $S_{\text{max}}$. The square blocks will be split into four subblocks unless a stopping criterion is satisfied. The stopping criterion is based on a uniform test, which checks whether a square block is texturally homogeneous. This procedure is recursively applied to the four subblocks until a minimum block size $S_{\text{min}}$ is reached. To get stable textural descriptions necessitates a minimum block size because the block size must contain sufficient number of pixels.

The stopping criterion for a block is based on the ratio of the largest index and the smallest index. There are six pair-wise indices among the four subblocks. If the relative dissimilarity within the block is greater than a threshold, i.e.

$$R = \frac{I_{\text{max}}}{I_{\text{min}}} > X$$

where $X$ is chosen by the researcher.

Obviously, the stopping criterion plays a critical role in the success of this step. If the stopping criterion fails to detect heterogeneous patterns in a square block, an entire block may be treated as homogeneous and the majority of the pixels in the block will remain mis-segmented after the following two steps. It is therefore better to split too much than too little so that heterogeneous regions are fully split. In the mean time, a homogenous block may be unnecessarily split into many smaller blocks. In the merging phase, homogenous regions will be
Figure 3.2 Illustration of the segmentation algorithms. (a) A composite image of five Brodatz texture samples. (b) Results of hierarchical splitting. (c) Results of agglomerative merging. (d) Overlay of merging results and the original image. (e) Results of pixel-wise refinement. (f) Overlay of pixel-wise refinement results and the original image.
merged and heterogeneous regions will remain separate. Therefore, researchers should choose a small \( X \) rather than a large \( X \).

Ojala and Pietikainen (1999) pointed out that this step was not mandatory. However, they found out that it improved the convergence of the following merging algorithm.

### 3.4.2 Agglomerative Merging

After the image is divided into uniform blocks of various sizes, a merging procedure is applied to merge similar adjacent regions until a stopping criterion is satisfied (Figure 3.2c, 3.2d).

If the index \( I \) is a measure of similarity within the range of 0 and 1 (like those obtained from histogram intersection and correlation), the pair of adjacent regions with the largest merging importance value (\( MI \)) is merged at each iteration. \( MI \) is computed as:

\[
MI = \frac{1}{\sqrt{p}} \times I
\]

where \( p \) is the number of pixels in the smaller region. Merging is allowed until either of two stopping criteria is met:

\[
MIR1 = \frac{MI_{\text{cur}}}{MI_{\text{min}}} < Y
\]
\[
MIR2 = \frac{MI_{\text{cur}}}{MI_{\text{max}}} < Z
\]

where \( MI_{\text{min}} \) is the smallest of all the preceding \( MI \) values; \( MI_{\text{max}} \) is the largest; and \( MI_{\text{cur}} \) is the current \( MI \) value of current best pair of regions. The smaller \( Y \) and \( Z \) are, the less clusters in the segmentation.

If the index \( I \) is a measure of dissimilarity (like the G-statistic), the pair of adjacent regions with the smallest \( MI \) value is merged at each iteration, where \( MI \) is computed as:

\[
MI = p \times I
\]

The stopping criterion is
\[ MIR2 = \frac{MI_{\text{cur}}}{MI_{\text{max}}} > Y \]

The value of \( Y \) is determined experimentally and may vary for different images. The smaller \( Y \) is, the more clusters in the segmentation.

### 3.4.3 Pixelwise Refinement

After the successful application of the previous two steps, reliable estimates of the different regions have been obtained. In this last step, a pixelwise classification scheme is used to better resolve the boundaries of the regions (Figure 3.2e, 3.2f). A pixel is regarded a boundary point if it is on the boundaries of at least two regions.

In the first scan, only boundary pixels are checked to see whether they should be reclassified. To do so, a discrete disk with a radius \( r \) is placed on a boundary point \( P \) and textural measures are calculated for this disk. Assuming there are \( m \) different regions for the four neighbors of the point \( P \). \( m \) can take the value of 2, 3 or 4. The indices between the disk and the \( m \) regions are calculated. Point \( P \) will be relabeled to the most similar region \( w \) in terms of textural patterns if the label of \( P \) is different from \( w \). If the index \( I \) is a measure of heterogeneity, the most similar region has the smallest \( I \) with the disk. If the index \( I \) is a measure of homogeneity, the most similar region has the largest \( I \) with the disk.

In the following scans, only those pixels which have been relabeled in the previous step will be checked. The process stops when no pixels are relabeled or when the maximum number of scans is reached.

### 3.5 Post-segmentation Integration

The success of the segmentation described above depends primarily on the discriminatory power of textural measures used. However, a particular measure or its distribution can hardly be expected to accommodate the complex situations encountered in remotely sensed images. For example, intensity histograms cannot discriminate between regions with similar spectral
response patterns but distinct spatial patterns, like agricultural land uses of different cultivation patterns. In a similar way, the LBP/C distribution or other textural measures along may be able to distinguish among some regions, but fail to successfully segment other textural distinct regions discernable to human eyes.

Hu et al. (2005) integrated texture, intensity and color measures to enhance the segmentation performance. First, three similarity indices are calculated based upon LBP/C distribution, intensity histogram, and saturation/hue distribution. The three indices are integrated into one index by an adaptive weighting scheme. Although the adaptive scheme improves the results, only one weighted index is obtained and it may still fail to accommodate some complex scenarios. In addition, the weights are decided in some subjective way.

Different textural measures excel in the discrimination of some aspects of textural patterns. The integration of segmented images combines the strength of each measure and can provide better results.

**3.6 Experiments**

Figure 3.3 shows a subset of a panchromatic IKONOS image, which was acquired on May 18, 2001 and located in northern Alabama. It had one-meter resolution and was of 400 by 400 pixel dimension. The image contained one water body, one vegetated area, two different agricultural land uses. Figure 3.4 shows the clustering results of the ISODATA procedure when the pixel values were the input and four cluster were specified. The water body was satisfactorily identified. However, the heterogeneous vegetated area in the lower left part of the image was not correctly clustered as one clump because of the substantial variations in pixel values in this cluster, even though it had pretty homogeneous textural patterns. It was the case for the main agricultural land use. This example illustrated the inefficiency of pixel-based clustering procedures and highlighted the need of incorporating textural patterns in image segmentation.
Figure 3.3 The subset of a panchromatic IKONOS image in northern Alabama, acquired on May 18, 2001. (one-meter resolution, 400 by 400 pixels)
Figure 3.4. Clustering results using ISODATA procedure and pixel values only.
Figure 3.5 Segmentation of the panchromatic subset using the intensity distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.6 Segmentation of the panchromatic subset using the LBP/C distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.7 (a) Combination of segmentations of the panchromatic subset from using intensity distribution and using LBC/C distribution. (b) Combination overlaid on the original image.
For the region-based splitting and merging algorithm, the minimum window size used was 16 by 16 pixels and the maximum window size was 64 by 64 pixels. $X$ was chosen to be 1.1.

When the intensity measure was used and correlation coefficient was the similarity index, the image was segmented into three regions (Figure 3.5), which were distinct in terms of their intensity patterns. For the intensity measure, $Y$ was 0.7 and $Z$ was 0.05, $r$ was 10, and the maximum number of scan was set to be 50.

However, the intensity measure failed to discriminate the vegetated region from its neighboring agricultural land use due to their similar intensity distribution even though they were markedly different in their spatial patterns. When the LBP/C measure was used and G-statistic was the dissimilarity index, it successfully distinguished the vegetated region from the agricultural region (Figure 3.6). For the LBP/C measure, $Y$ was 1.9, $r$ was 10, and the maximum number of scan was set to be 50.

However, the LBPC/C distribution failed to isolate the water body from its neighboring agricultural region. This was because both the water body and its neighboring region were relatively smooth and their LBP/C patterns were similar since LBP/C was intensity-invariant and it detected local patterns only. When the two segmented results were overlaid together (Figure 3.7), the result reflected the original four regions very well.

There were some narrow strips after the overlay due to the discrepancies of the boundaries between the two segmentations. Those narrow strips were combined with its largest neighboring clumps to avoid too fragmented segmentation. This was done by specifying the minimum number of pixels in a clump.

Figure 3.8 shows a subset of the red band of an aerial photo in Baton Rouge, LA, acquired on March 22, 2002. It had one-foot resolution and was of 512 by 512 pixel dimension. The image contained one river, one road, some vegetated clusters, and etc. The minimum window size used was 16 by 16 pixels and the maximum window size was 64 by 64 pixels. $X$ was
chosen to be 1.1. When the intensity measure was used and correlation coefficient was the similarity index, the image was segmented into several regions (Figure 3.9), which were distinct in terms of their intensity patterns. For the intensity measure, \( Y \) was 0.5 and \( Z \) was 0.05, \( r \) was 10, and the maximum number of scan was set to be 50.

However, the intensity measure failed to discriminate the vegetated region along the river from its neighboring region due to their similar intensity distribution even though they were markedly different in their spatial patterns. When the LBP/C measure was used and G-statistic was the dissimilarity index, it successfully distinguished this vegetated region along the river from its adjacent region (Figure 3.10). For the LBP/C measure, \( Y \) was 1.7, \( r \) was 10, and the maximum number of scan was set to be 50.

However, the LBP/C distribution failed to discriminate the region along the road from its neighboring region further away from the road. This was because the two regions had similar local binary patterns. They were successfully separated from each other using the intensity measure because of their distinct intensity differences. When the two segmented results were overlaid together (Figure 3.11), the result reflected the original clusters very well. Some small strips after the overlay were merged with their largest adjacent clusters.

The above two examples show that the combination of segmentations helped to identify different clusters recognized by different measure distributions. This scheme, however, may result in too fragmented segmentation which may not be desirable. Figure 3.12 illustrates this case. The intensity distribution identified four clusters in the image. The LBP/C distribution identified one more cluster (marked with “A”). Cluster A differed from cluster B in terms of the orientations of their spatial patterns. If this difference is not desirable to the users, the two clusters may need to be merged. We may also choose some rotation-invariant spatial measures to avoid detecting the two sub-clusters. On the other hand, if the difference in the orientation is
Figure 3.8 The subset of the red band of an aerial photo in Baton Rouge, LA, acquired on March 22, 2002 (one foot in resolution, 512 by 512 pixels).
Figure 3.9 Segmentation of the subset of the photo using the intensity distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.10 Segmentation of the subset of the photo using the intensity distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.11 (a) Combination of segmentations of the subset of the aerial photo from using intensity distribution and using LBC/C distribution. (b) Combination overlaid on the original image.
Figure 3.12 Left: Segmentation of a subset of a panchromatic image (one-foot resolution) using the intensity distribution. Right: Segmentation using the LBP/C distribution.
important, the segmentation by the LBP/C distribution provides very satisfactory results.

The next example illustrates how the saturation/hue distribution can contribute to the successful segmentation of an image. Figure 3.13 shows a subset of a color aerial photo (red, green, blue) in Baton Rouge, LA, acquired on March 22, 2002. It had one-foot resolution and was of 512 by 512 pixel dimension. The minimum window size used was 16 by 16 pixels and the maximum window size was 64 by 64 pixels. $X$ was chosen to be 1.1. When the saturation/hue measure was used and correlation coefficient was the similarity index, the image was segmented into several regions (Figure 3.14), which were distinct in terms of their color patterns. For the saturation/hue measure, $Y$ was 0.5 and $Z$ was 0.085, $r$ was 10, and the maximum number of scan was set to be 50.

However, the saturation/hue measure failed to separate a relatively smooth region in the middle-left part of the image from its neighboring region due to their similar color distribution even though they were markedly different in their spatial patterns. When the LBP/C measure was used and G-statistic was the dissimilarity index, it successfully distinguished this vegetated region along the river from its adjacent region (Figure 3.15). For the LBP/C measure, $Y$ was 1.7, $r$ was 10, and the maximum number of scan was set to be 50.

However, the LBP/C distribution failed to discriminate two large smooth regions with different colors (green and rosy brown) from each other. This was because the two regions had similar local binary patterns. They were successfully separated from each other using the saturation/hue measure because of their distinct color differences. When the two segmented results were overlaid together (Figure 3.16), the result reflected the original clusters very well. Some small strips after the overlay were merged with neighboring large clumps.

The following two examples illustrate that in some cases, one measure distribution is good enough for the successful segmentation of an image. Figure 3.17 shows a subset of an aerial photo whose resolution was 1.8 meter whose dimension was of 400 by 400 pixels. The image
Figure 3.13  The color subset of an aerial photo (0.9m in resolution, 512 by 512 pixels)
Figure 3.14 Segmentation of the color subset using the hue/saturation distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.15 Segmentation of the subset of the photo using the LBP/C distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.16 (a) Combination of segmentations of the color subset of the aerial photo from using intensity distribution and using LBC/C distribution. (b) Combination overlaid on the original image.
Figure 3.17 Segmentation of a subset of an aerial photo using the LBP/C distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
Figure 3.18 Segmentation of a subset of an aerial photo using the saturation/hue distribution. (a) Results of agglomerative merging. (b) Overlay of merging results and the original image. (c) Results of pixel-wise refinement. (d) Overlay of pixel-wise refinement results and the original image.
contained four main clusters, i.e., a vegetated area, a bare ground, and two neighborhoods of different densities. The LBP/C distribution segmented the image very well and the boundaries between each cluster were well-resolved. Figure 3.18 shows a subset of another aerial photo whose resolution was 1.8 meter and was of 400 by 400 pixels in dimension size. There were five clusters in the image and each of them had a distinct color pattern. The saturation/hue measure satisfactorily identified the five clusters and resolved the boundaries very well.

3.7 Discussion and Conclusions

In this chapter, three textural measure distributions, i.e., the intensity distribution, the LBP/C distribution, and the saturation hue distribution were used in the discrimination of textural patterns. Each of them characterized one aspect of spatial patterns and proved to be useful in the detection of homogeneous regions in an image. The combination of segmentations by different measures proved to be helpful in identifying homogeneous regions in an image. Some clusters left out by one measure were picked up by another measure. The post-segmentation overlay reflected all the clusters identified by two or more measures. Compared with pre-segmentation weighting schemes of different measures, this method did not involve the determination of weighting coefficients subjectively.

In some complex scenarios, a textural measure is hardly expected to perform equally well across an entire image. A textural measure may identify clusters in some parts of an image very well while producing too fragmented segmentation (over-segmentation) or too coarse segmentation (under-segmentation) in other parts. In this case, to achieve a better combination of the segmentations by different measures, the users may have to manually combine some clusters in one segmentation while discarding some clusters in the same segmentation to meet their needs. This can be done by interactively working on two segmentations and keep the segmentations that are appropriate to the users.
The success of the splitting-and-merging segmentation procedures used in this Chapter was found to be sensitive to the selection of the parameters. This is particularly the case for complex remotely sensed images. For hierarchical splitting, the maximum widow size and the stopping criterion \( X \) were critical. For agglomerative merging, the stopping criteria affected the segmentation greatly, leading to over-segmentation or under-segmentation. The segmentation was less affected by the parameters in the pixel-wise refinement stage, which helped to resolve boundaries between adjacent clusters more accurately. The parameters had to be determined experimentally and varied for different images.

Another point that is worth noting is that the designated parameters are applied to an entire image without any adaptive scheme to take local variations into consideration. For some parts of an image, large parameters are needed to yield a satisfactory segmentation, whereas for other parts, small parameters are needed because of the difference in the complexity in difference regions. Global parameters are hardly expected to accommodate different complexities in different parts of an image. In future studies, local parameters may be used to achieve better segmentations.

Textural measures played a critical role in the success of the splitting-and-merging procedures. The LBP/C measure was very powerful in distinguishing textural patterns. The LBP/C measure varied with rotations and gray-scale linear contrast. Ojala et al. (2002) extended their work and developed a gray-scale and rotation-invariant local binary pattern textural descriptor. Whether gray-scale and rotation-invariant textural measures are desired, however, depends on different applications. Other textural measures may also be adopted in the segmentation framework.

In this Chapter, three bands were used to extract saturation/hue information based on average spectral values. Multispectral information can be incorporated in the segmentation
procedures very straightforwardly. This can be done by obtaining measure distributions from each band and getting an average similarity or dissimilarity index from all bands.
Chapter 4
Wavelet Feature Extraction and Artificial Neural Networks for Textural Discrimination of Urban Areas Using High-Resolution Satellite Imagery

4.1 Introduction

Urban remote sensing is one of the most fascinating and challenging realms in spaceborne remote sensing. To examine detailed types of land uses and land covers, high-resolution data are necessary (Welch 1982, Jensen and Cowen 1999). However, it has been reported that the use of higher spatial-resolution imagery does not necessarily lead to higher classification accuracy. On the contrary, higher spatial-resolution imagery has lower classification accuracy by conventional per-pixel classification methods in urban areas due to the heterogeneous nature of urban environment (Toll 1985, Hacck 1987, Chen et al. 2004).

A series of satellite sensors are now available to acquire images at resolutions ranging from more than one kilometer through tens of meters to less than one meter (Jensen et al. 2004). The availability of high spatial resolution remotely sensed images in the urban context calls for new approaches to extract land-use/land-cover information (Jensen et al. 2004). These new approaches should be able to deal with the heterogeneous nature of urban environment at high resolution. Moreover, as the spatial resolution becomes finer, textural information becomes more valuable and texture becomes critical and indispensable for the digital analysis of satellite images. Therefore, classification procedures that incorporate key spatial/textural information are expected to be useful and should be further explored for classifying heterogeneous urban land covers.

Geographic phenomena operate at different scales (Lam et al. 2004). A multiresolution approach to deal with texture and spatial patterns as manifested in an image meets the nature of geographic phenomena. Multiresolution wavelet analysis (Mallat 1989, Daubechies 1992, Frazier 1999) has proven to be an effective tool for the characterization of textures and spatial
patterns at different scales (Zhu and Yang 1998, Myint et al. 2004).

Artificial Neural Networks (ANNs) have long been employed as an effort to improve classification accuracy of remote sensing images (Paola and Schowengerdt 1995, Atkinson and Tatnal 1997). ANNs require no assumptions about the data distribution and are expected to be more efficient in classifying wavelet measures. Although wavelet analysis and ANNs have been employed separately to analyze remotely sensed data, studies that bring them together are rare, with none being found in the remote sensing literature. It is expected that a combined methodology using both wavelet transform and ANN will improve the discriminating ability of high-resolution urban textures.

Previous studies have focused on using either the minimum-distance classifier or discriminant analysis for classifying wavelet measures. In this study, investigation on the use of neural networks as a means for classifying wavelet measures from an IKONOS image was conducted and the results were compared with those derived from the minimum-distance classifier and discriminant analysis.

4.2 Wavelet Analysis and Artificial Neural Networks

4.2.1 Wavelet Analysis

The past 20 years has seen a considerable increase in the study and application of wavelet analysis. The wavelet transform has a well-established mathematical foundation. Wavelets are translated and dilated versions of a common function, called mother wavelet. The translation corresponds to the location in geographic space and the dilation relates to different scales. By adjusting the translation and dilation, scale effects can be studied locally. There are two types of wavelet transforms: continuous versus discrete. Continuous wavelet transform is mostly used for theoretical study. The discrete wavelet transform is employed to do a non-redundant and reversible transformation. A one-dimensional discrete transformation is basically a change of orthonormal basis in the linear space. For a two-dimensional discrete wavelet transform, an
image \( f \) is decomposed to four subimages: the approximate image \( A \), the horizontally detailed image \( H \), the vertically detailed image \( V \), and the diagonally detailed image \( D \).

The coefficients of the four subimages are computed by equations (1)–(4):

\[
A(i, j) = \sum_k \sum_l h(k - 2i)h(l - 2j) f(k, l)
\]

\[
H(i, j) = \sum_k \sum_l h(k - 2i)g(l - 2j) f(k, l)
\]

\[
V(i, j) = \sum_k \sum_l g(k - 2i)h(l - 2j) f(k, l)
\]

\[
D(i, j) = \sum_k \sum_l g(k - 2i)g(l - 2j) f(k, l)
\]

where \( h(k) \), \( g(k) \) are the scaling filter and the wavelet filter, respectively; \( i, j \) are the row and column indices in the decomposed image and \( k, l \) are the row and column indices in the original image (Mallat 1989, Daubechies 1992, Frazier 1999). An example of wavelet decomposition is shown in Figure 4.1.

The decomposition is conducted through a pyramid algorithm (Figure 4.2). Traditionally, the approximate image can be decomposed further, resulting in a multi-resolution analysis. Measures, like mean, standard deviation, entropy, and energy, can be computed from each of these subimages to form a wavelet measure vector, which can be processed further.

The wavelet transform offers an efficient approach to studying textures in a multi-resolution manner. Wavelet measures have been shown to be an effective way for the characterization and classification of textures (Zhu and Yang 1998, Myint et al. 2004), urban object retrieval (Bian 2003), denoising and smoothing (Wang and Zhang 2003), image fusion (Chibani and Houacine 2003), image compression (Kiema and Bahr 2001), and so on. However, as far as the classification of wavelet measure is concerned, many researchers used simple classifiers such as the minimum-distance classifier (Zhu and Yang 1998, Zhao 2001) and linear discriminant analysis (Myint et al. 2004). The minimum-distance classifier takes only distance into account.
Figure 4.1. (a) The image of a fingerprint  (b) Four one-level decomposed images with Harr wavelet. Top left: approximate subimage; Top right: horizontally detailed subimage; Bottom left: vertically detailed subimage; Bottom right: diagonally detailed subimage. Image source: MATLAB 6.5 sample.
Figure 4.2. One-level two-dimensional discrete wavelet transform (after Mallat, 1989)
and linear discriminant analysis has many statistical assumptions about the input data, such as equality of variance-covariance matrices and normal distribution (Tatsuoka and Lohnes 1988), which may not be satisfied. A classifier such as ANN may alleviate these problems and help improve the classification of wavelet measures so that they can be used for subsequent classification of an entire image.

4.2.2 Artificial Neural Networks

ANNs have emerged as a powerful tool for various applications in many fields due to their ability to learn patterns and relationships in complex, multi-dimensional data sets. ANNs have been frequently applied in remote sensing since the late 1980s (e.g., Key et al. 1989, Hepner et al. 1990, Cherukuri 1997). The development of ANNs was inspired from biological model networks of human brains, but it has evolved such that it retains only a distant resemblance to real life cognitive systems (Amaratunga 2004).

An ANN is a computational model consisting of simple interwoven neurons called neurons. Each neuron operates as a computation neuron, taking input from other neurons and giving output that can be input to other neurons or becoming the final desired output. Once the architecture of an ANN is established, it is entirely determined by its constituent neurons and the connections among them. Generally, the computation of the output from an input is fast from an ANN. Nonetheless, the process of establishing the structure of the ANN may be time-consuming. Once properly trained, the ANN is expected to have generalization ability and to respond well to unknown data.

The popularity of ANNs lies in that they have no \textit{a priori} assumptions about the distribution of input data. When used as a classifier, an ANN learns the complex relationships among the input data through the training process and records the relationships in its connecting coefficients between connected neurons. In contrast, the maximum-likelihood classifier assumes a Gaussian
distribution of the input data and requires the estimation of parameters of the Gaussian
distribution from the input data. The validity of this assumption is often questioned and failure to
satisfy this assumption is held accountable for low classification accuracy (Paola and

The ANN classifiers have become an important alternative to the maximum likelihood
classifier (Augusteijn and Folkert 2002). The ANN approach has been reported to achieve
improved accuracy over traditional statistical classifiers (Erbet et al. 2004, Qiu and Jensen 2004).
The ANN method tends to be more robust to training site selection and class definition, and is
easier to accommodate heterogeneous categories, whereas the maximum likelihood algorithm is
sensitive to the purity of the class measures and performs poorly if these measures are not pure
(Paola and Schowengerdt 1995).

Despite its many strengths, an ANN has several drawbacks (Verbeke et al. 2004, Atkinson
and Tatnall 1997). There are a lot of parameters to be specified before an ANN is trained, such
as network architecture, input/output representation, initialization of weighting coefficients, and
learning algorithm and parameters. There are no universally accepted rules for setting these
parameters. As a result, the training is usually a trial-and-error process. After an ANN is trained,
it is a black-box model, which makes the interpretation of its behaviors difficult.

The most extensively used supervised ANN is called multilayer perception (MLP) (Bishop
1995). Shown in Figure 4.3 is the structure of a three-layer MLP. An MLP consists of an input
layer, one or more hidden layers, and one output layer. The layers are arranged in sequential
orders. The neurons between two consecutive layers are fully connected. Except the input layer,
neurons in the hidden layers and output layer receive weighted sum of outputs from their
immediately preceding layers plus biases as inputs and yield outputs through an activation
function. An MLP is typically trained by one of a host of algorithms such as gradient descent,
conjugate gradient, and Levenberg-Marquardt algorithms.
Figure 4.3. The structure of a three-layer multilayer perceptron neural network (after Bishop, 1995)
In this Chapter, the Levenberg-Marquardt algorithm (Hagan and Menhaj 1994) was used to train the network because of its excellent performance in achieving convergence (Demuth and Beale 2004). In the Levenberg-Marquardt algorithm, the change in the coefficients and biases are calculated through equation (5):

$$\Delta x = (J'(x)J(x) + \mu I)^{-1}J'(x)e(x)$$  \hspace{1cm} (5)

where $x$ is a vector composed of all the weighting coefficients and biases, $e(x)$ is the error vector, $J(x)$ is the Jacobian matrix of $e(x)$ with respect to $x$, $\mu$ is a scalar, and $I$ is the identity matrix. The neurons in $J(x)$ can be calculated through a back-propagation process. $\mu$ is decreased by a factor when a step would result in a drop in the performance function and is increased by a factor when the performance function would increase. Both wavelet analysis and neural network training and classification were done in Matlab™.

4.3 Methodology
4.3.1 Dataset and Classification Scheme

The metropolitan area of Atlanta, Georgia is the second fastest-growing city in the United States and has undergone rapid suburbanization and urban sprawl over the course of last 25 years (Lo and Choi 2004). A high-resolution IKONOS image (Space Imaging 2004) was acquired on 29 December 2000. The IKONOS image has one 1-meter resolution panchromatic band and four multispectral bands of a resolution of 4 meters. The image covers part of the metropolitan area of Atlanta, Georgia. It was encoded in 11 bits, and had been radiometrically and geometrically corrected. The projection was Universal Transverse Mercator, Zone 16. The datum was WGS84. The location of the study area is shown in Figure 4.4.

Five types of land uses and land covers were considered, including commercial land, industrial land, multifamily residential housing, single-family residential housing, and forest. This scheme follows the one used by the Atlanta Regional Commission (ARC), a regional planning and intergovernmental coordination agency for the Atlantic area. The ARC classified
Figure 4.4 The metropolitan area of Atlanta, Georgia, USA and the location of the image
land uses/land covers based on the USGS system (Anderson et al. 1976). The scheme used in this research excludes land-use/land-cover types that either did not exist in the study areas or covered too small areas so that no sufficiently large texture samples could be obtained. The descriptions of the classes and the number of training samples are shown in Table 4.1. Textural samples of the five classes are shown in Figure 4.5.

Two experiments were conducted. The first experiment involved the classification of textural samples. In the first experiment, to avoid the problem of mixed classes and boundaries, two subsets were taken from the panchromatic band of the image for each land-cover class, resulting in ten subsets in total. Each subset is approximately 640 by 320 pixels. These subsets were selected carefully with the help of ground surveying, a GIS dataset compiled by the ARC, and visual inspections of the image. Systematic sampling was carried out to collect samples from the subsets of each land-cover class. Table 4.1 lists the number of samples extracted for each land-cover class. In addition, for each sample, three window sizes were considered: 64×64, 48×48, and 32×32. The center of each sample was spaced from each other at 65 pixels in row and 65 pixels in column. Since the size of the maximum window considered in this study was 64×64, this sampling strategy ensured that the training textural samples did not overlap. The samples were then divided into three parts. The first part was used as the training set, the second part as the validation set, and the last part as the test set.

In the second experiment, a mosaic image (Figure 4.10) composed of the 5 land-cover samples (each with a size of 400×400 pixels) was extracted from the IKONOS panchromatic band. The mosaic image was classified using the ANN network which produced the highest accuracy in the first experiment and the corresponding wavelet parameters (i.e., window size, wavelet measures, and decomposition level). A moving window was centered on each pixel and the pixel was classified with respect to the measure vector computed from the decomposed images.
Table 4.1. Land use/land cover classification scheme, characteristics, and number of textural samples

<table>
<thead>
<tr>
<th>Land use/land cover classes</th>
<th>Descriptions</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial</td>
<td>Areas used predominantly for the sale of products and services, including office buildings, warehouses, driveways, sheds, parking lots, landscaped areas, and etc.</td>
<td>100</td>
</tr>
<tr>
<td>Industrial</td>
<td>Land associated with light or heavy manufacturing plants</td>
<td>90</td>
</tr>
<tr>
<td>Multifamily housing</td>
<td>Land occupied by apartment and townhouse complexes where net density generally exceeds eight units per acre.</td>
<td>86</td>
</tr>
<tr>
<td>Single-family housing</td>
<td>Single-family detached housing</td>
<td>100</td>
</tr>
<tr>
<td>Forest</td>
<td>All forested land dominated by trees that lose their leaves at the end of the frost-free season, except those characteristic of wetlands</td>
<td>76</td>
</tr>
</tbody>
</table>
Figure 4.5. Samples of five land-use/land-cover classes from the panchromatic band of the IKONOS image of Atlanta. (a) commercial land, (b) industrial land, (c) multifamily housing, (d) single-family housing, and (e) forest
4.3.2 Wavelet Transform and Measures

In this study, textural samples were processed using the discrete wavelet analysis up to two levels with Daubechies wavelets of order 2 and 8 (henceforth referred as db2 and db8 respectively), which have a support (number of non-zero neurons) of 4 and 16, respectively (Daubechies 1992). Daubechies wavelet family is one of the most popular wavelet families, which includes the simple Harr wavelet. Daubechies wavelets are compactly supported and orthonormal, which make them particularly suitable for studying variations at local scales. Figure 4.6 shows the shape of the two wavelets used in this study.

Various measures have been used in the literature (Zhu and Yang 1998, Myint et al. 2004). Four measures, mean, standard deviation, energy, and entropy, were extracted from the decomposed subimages through equations (6)–(10):

\[
\text{mean} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |P(i, j)| \tag{6}
\]

\[
\text{stand deviation} = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (|P(i, j)| - \text{mean})^2} \tag{7}
\]

\[
\text{energy} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Q(i, j)^2 \tag{8}
\]

\[
\text{entropy} = -\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Q(i, j) \log Q(i, j) \quad (Q(i, j) \neq 0) \tag{9}
\]

\[
Q(i, j) = \frac{|P(i, j)|^2}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P(i, j)^2}} \tag{10}
\]

where \(P(i, j)\) is the value of the decomposed image at pixel \((i, j)\), \(M\) and \(N\) are the number of rows and columns, and \(Q(i, j)\) is normalized from \(P(i, j)\).

4.3.3 Configuring the Neural Network

A three-layer MLP neural network was used to classify each of the four wavelet measures. For each land-cover class, the number of inputs was equal to the number of decomposed images,
Figure 4.6 Daubechies wavelets of order 2 (Left) and order 8 (Right), generated with Matlab
i.e., 4 inputs of wavelet measures (computed from 4 subimages) for level-1 decomposition and 8 inputs for level-2 decomposition. Inputs were scaled to the range of $-1$ to $+1$. There is no widely accepted guideline for determining the number of neurons in the hidden layer. Kavzoglu and Mather (2003) compared several empirical rules of thumb for deciding the number of hidden neurons. They concluded that most of them produced similar results, which was also verified by the preliminary testing in this study.

All the results reported in the following section were based on neural networks whose number of neurons in the hidden layer was twice the number of the inputs. The output layer had five neurons, with each neuron corresponding to one class. In the training process, an output of 1 was chosen to represent the correct class and 0 to represent all other classes. The initialization of the weighting coefficients may have a significant effect on the training process since the training may get stuck in a local minimum and never reach the global minimum.

Many researchers trained neural networks several times with different initiating weighting coefficients to take into account this effect (Walder and Maclaren 2000, Kavzoglu and Mather 2003, Tedesco et al. 2004). In this study, 10 neural networks were trained for each dataset with different random initial coefficients using algorithms described by Nguyen and Widrow (1990), which helped to accelerate the convergence of the networks and reduced training time. The validation set was used to prevent the networks from being over-trained. The three networks with the highest performances on validation data set were found to produce similar results on the corresponding test dataset. All the results reported in the following section were based on the network with the highest classification accuracy on the validation set. The networks were trained with the Levenberg-Marquardt algorithm (Hagan and Menhaj 1994, Demuth and Beale 2004), which is implemented in Matlab™.

For the purpose of comparison, this study also considered the minimum-distance classifier (Zhu and Yang 1998, Zhao 2001) and linear discriminant analysis (Myint et al. 2004), which
were used in previous studies for classifying wavelet measures.

4.4 Results and Discussion

4.4.1 First Experiment

Figure 4.7 shows each land-cover class and decomposed subimages (at the first decomposition level) the means and standard deviations of four measures computed from the training samples using window size of 64m×64m. Generally speaking, commercial and industrial land uses showed more spread and variation than single-family housing and forest. This means that classification accuracy for commercial and industrial land uses would likely be lower than those for single-family and forest. Not a single wavelet measure based on a specific decomposed subimage seems to provide good discriminating power among all five classes. Separation among some classes is good for some measures computed from one decomposed subimage. For example, standard deviation and entropy seem suitable for the separation of commercial land use from other land uses using the approximate subimage.

Table 4.2 compares the overall classification accuracy of the three classifiers (ANN, discriminant analysis, and minimum distance) using a window size of 64m×64m. Some observations can be made. First, the neural network approach and discriminant analysis performed much better than the minimum distance classifier regardless of the measure, decomposition level, and wavelet type. It is clear that the minimum distance classifier should not be used in the future for classifying land covers from textural measures. As a result, the following discussion does not include the minimum distance classifier.

Second, the difference in the overall classification accuracy between the neural network classifier and discriminant analysis was inconsistent and insignificant. At the first decomposition level, discriminant analysis slightly outperformed the neural network classifier for standard deviation and entropy using the db2 wavelet and for mean using the db8 wavelet. In all other cases, the neural network classifier performed slightly better than discriminant analysis. When
Figure 4.7. Mean and standard deviation plots of measures of four subimages at the first decomposition level using a window size of 64 by 64 and the db2 wavelet. The center of each bar is the mean and the endpoints represent one standard deviation from the mean. From left to right, the plot corresponds to the approximate, horizontally, vertically, and diagonally detailed images respectively.
Table 4.2. Overall classification accuracy (%) of textural samples with window size of 64 by 64. (NN: neural network classifier; DA: discriminant analysis classifier; MD: minimum-distance classifier; Level #: decomposition level; NS: number of subimages; std: standard deviation)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Level#</th>
<th>NS</th>
<th>db2 mean</th>
<th>db2 std</th>
<th>db2 energy</th>
<th>db2 entropy</th>
<th>db8 mean</th>
<th>db8 std</th>
<th>db8 energy</th>
<th>db8 entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN</td>
<td>1</td>
<td>4</td>
<td>70.86</td>
<td>66.23</td>
<td>58.94</td>
<td>60.26</td>
<td>66.89</td>
<td>69.54</td>
<td>63.58</td>
<td>65.56</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>76.82</td>
<td>66.23</td>
<td>60.26</td>
<td>64.24</td>
<td>75.50</td>
<td>69.54</td>
<td>64.24</td>
<td>74.83</td>
</tr>
<tr>
<td>DA</td>
<td>1</td>
<td>4</td>
<td>65.56</td>
<td>66.23</td>
<td>56.95</td>
<td>64.24</td>
<td>67.55</td>
<td>64.24</td>
<td>56.29</td>
<td>58.28</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>72.85</td>
<td>63.58</td>
<td>52.32</td>
<td>62.91</td>
<td>69.54</td>
<td>63.58</td>
<td>60.26</td>
<td>64.24</td>
</tr>
<tr>
<td>MD</td>
<td>1</td>
<td>4</td>
<td>37.75</td>
<td>59.60</td>
<td>39.07</td>
<td>52.98</td>
<td>38.41</td>
<td>60.26</td>
<td>39.74</td>
<td>52.32</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>39.07</td>
<td>60.26</td>
<td>43.05</td>
<td>58.94</td>
<td>39.07</td>
<td>59.60</td>
<td>49.01</td>
<td>56.29</td>
</tr>
</tbody>
</table>
Table 4.3 Kappa Z statistics of the neural network classifier over linear discriminant analysis with window size of 64 by 64 applied to textural samples

<table>
<thead>
<tr>
<th>Level#</th>
<th>NS</th>
<th>db2 mean</th>
<th>db2 std</th>
<th>db2 energy</th>
<th>db2 entropy</th>
<th>db8 mean</th>
<th>db8 std</th>
<th>db8 energy</th>
<th>db8 entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>1.03</td>
<td>-0.04</td>
<td>0.33</td>
<td>-0.82</td>
<td>-0.16</td>
<td>1.04</td>
<td>1.42</td>
<td>1.40</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>0.83</td>
<td>0.49</td>
<td>1.50</td>
<td>0.24</td>
<td>1.21</td>
<td>1.17</td>
<td>0.76</td>
<td>2.13</td>
</tr>
</tbody>
</table>
the differences of the overall classification accuracy between the two classifiers were quantified as the Kappa Z statistic (Congalton and Mead 1983), shown in Table 4.3, all the Z statistics, except the one associated with level 2 and db8 wavelet (Z=2.13), were less than 1.96 in absolute value and none of the differences was significant at the 95% confidence level.

Third, it is expected that the more decomposition levels used, the more information obtained, and the higher the classification accuracy. However, it is important to point out that a sufficient number of pixels in the decomposed images must be present after decomposition to ensure a more reliable representation. In this study, two decomposition levels were used, as further decomposition would lead to very small subimages, making textural patterns difficult to identify (For window size of 64m×64m, a third decomposition level results in subimages of 8×8 sq. meters). The results (Table 4.2) show that the neural network classifier consistently produced higher or the same classification accuracy at two decomposition levels than at the first decomposition level only, irrespective of the measure and the wavelet used. This does not hold true for discriminant analysis, which sometimes yielded slightly lower overall classification accuracy at the two decomposition levels than at the first decomposition level for some combinations of measures and wavelets (e.g., the standard deviation with both wavelets, the energy and entropy with the db2 wavelet).

Fourth, when the type of wavelet is compared, with the exception of the mean measure, the neural network approach produced lower accuracy with the db2 wavelet than with the db8 wavelet for all measures regardless of the level of decomposition. The effect of the wavelet type on discriminant analysis, however, depends on the specific measure and decomposition level. The results show that the energy measure was always associated with the lowest overall classification accuracy, whereas the mean measure achieved the highest or the second highest accuracy in all cases.

With the mean measure and the db2 wavelet at two decomposition levels, the neural network
classifier and discriminant analysis achieved the highest overall accuracy in the study. Table 4.4 and Table 4.5 present the classification accuracy of each individual class at this combination of settings respectively. Interestingly, although commercial land use exhibited the largest variation in the measure, it attained the highest producer’s and user’s accuracy among the five classes by the neural network approach. With discriminant analysis, however, forest had the highest producer’s accuracy, followed by single-family housing. This difference by the two approaches could be attributed to the prerequisites of the methods and the distribution of the data. As shown in Figure 4.7, forest and single-family housing showed the least variability in the mean measure, therefore satisfying the statistical requirements of linear discriminant analysis more than other classes and achieving higher producer’s accuracy, i.e., lower omission error. The neural network classifier, on the other hand, has no requirement on the data distribution and seems to be able to capture the complexity of the commercial data, thus yielding the best results among all five classes.

Regarding the effects of window size on the classification, Figure 4.8 and Figure 4.9 compare the overall accuracy of the three window sizes, 64m×64m, 48m×48m, and 32m×32m. Smaller window sizes led to lower accuracy, with very minor exceptions. For the neural network classifier, two exceptions were the energy measure with the db2 wavelet at two decomposition levels for window size 48m×48m, and the mean measure with the db8 wavelet at the first decomposition for window size 48m×48m. For discriminant analysis, the three minor exceptions were the energy and entropy measures with the db2 wavelets at two decomposition levels and the entropy measure with the db8 wavelets at one decomposition level. Without exception, window size of 32m×32m was always associated with the lowest classification accuracy. Hence, in order to be detectable, textural samples with a window size of 64m×64m or larger for urban applications are preferable.
Table 4.4. The error matrix by the neural network classifier with the mean measure and the db2 wavelet at two decomposition levels applied to textural samples. (C: commercial, I: industrial, MF: multifamily housing, SF: single-family housing, and F: Forest)

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>C</th>
<th>I</th>
<th>MF</th>
<th>SF</th>
<th>F</th>
<th>User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td>31</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>86%</td>
</tr>
<tr>
<td>I</td>
<td></td>
<td>1</td>
<td>18</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>75%</td>
</tr>
<tr>
<td>MF</td>
<td></td>
<td>1</td>
<td>3</td>
<td>19</td>
<td>6</td>
<td>0</td>
<td>66%</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>24</td>
<td>1</td>
<td>75%</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>80%</td>
</tr>
<tr>
<td><strong>Producer’s Accuracy</strong></td>
<td></td>
<td><strong>94%</strong></td>
<td><strong>60%</strong></td>
<td><strong>66%</strong></td>
<td><strong>73%</strong></td>
<td><strong>92%</strong></td>
<td><strong>77%</strong> (overall)</td>
</tr>
</tbody>
</table>
Table 4.5. The error matrix by discriminant analysis with the mean measure and the db2 wavelet at two decomposition levels applied to textural samples. (C: commercial, I: industrial, MF: multifamily housing, SF: single-family housing, and F: Forest)

<table>
<thead>
<tr>
<th>Classified</th>
<th>C</th>
<th>I</th>
<th>MF</th>
<th>SF</th>
<th>F</th>
<th>User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>23</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96%</td>
</tr>
<tr>
<td>I</td>
<td>3</td>
<td>19</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>70%</td>
</tr>
<tr>
<td>MF</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td>3</td>
<td>0</td>
<td>62%</td>
</tr>
<tr>
<td>SF</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>26</td>
<td>0</td>
<td>63%</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>26</td>
<td>79%</td>
</tr>
<tr>
<td><strong>Producer’s Accuracy</strong></td>
<td>70%</td>
<td>63%</td>
<td>55%</td>
<td>79%</td>
<td>100%</td>
<td><strong>73% (overall)</strong></td>
</tr>
</tbody>
</table>
Figure 4.8. The overall classification accuracy of three window sizes: 64×64, 48×48, and 32×32 with the db2 wavelet (top) and the db8 wavelet (bottom) by the neural network approach. Left: one decomposition level only. Right: two decomposition levels.
Figure 4.9. The overall classification accuracy of three window sizes: 64×64, 48×48, and 32×32 with the db2 wavelet (top) and the db8 wavelet (bottom) by discriminant analysis. Left: one decomposition level. Right: two decomposition levels.
Figure 4.10. Top: a mosaic image composed of 5 land-use classes (commercial, industrial, multifamily, single-family, and forest), each of which is $400 \times 400$. Middle: classified map by the neural network approach using the \textit{mean} measure from 2 levels of decomposed images by the db2 wavelet. Bottom: classified map by discriminant analysis using the same parameters and wavelet type.
Figure 4.11. Top: User's accuracy of the mosaic image. Bottom: Producer's accuracy of the mosaic image. The overall accuracy for neural network and discriminant analysis were 84.9% and 78.8%, respectively.
4.4.2 Second Experiment

In the first experiment, it was found that with a window size of 64m×64m, the mean measure vector computed from decomposed images at two levels with the db2 wavelet provided the highest classification accuracy of the textural samples for both the neural network approach and discriminant analysis. Therefore, in the second experiment, the same parameters were used to compute the measure vector. The neural network trained with the measure vector in the first experiment was used in the classification of the mosaic image.

Figure 4.10 show the classification results of the mosaic image by the neural network approach and discriminant analysis. The overall accuracy by the neural network approach was 84.9% and 78.0% by discriminant analysis. In Figure 4.11, it shows that the neural network approach generally yielded higher producer's accuracy and user's accuracy for each individual class with two exceptions. One exception was the producer's accuracy of the forest, where discriminant analysis yielded slightly higher value than the neural network approach. The other exception was the user's accuracy of the commercial land use. Some errors arose along the class boundaries, which could be attributed to the edge effect (Ferro and Warner 2002). Some parts of industrial land use were misclassified as the commercial land use due to their similar patterns on the image. The lower user’s accuracy of the industrial land use and the multifamily housing were largely caused by the confusion between each other. Some parts of the single-family housing were misclassified as multifamily housing mainly because the window size was not large enough to cover the repeated patterns shown on the single-family housing.

4.5 Conclusions

In this Chapter, wavelet transforms were used to extract measures from high-resolution urban land-use/land-cover classes and neural networks were used to carry out the classification of wavelet measures. The performance of neural networks as a textural classifier was compared with that of discriminant analysis and minimum distance classifier, which had been used in
previous literature. Results from this study show that the minimum distance classifier gave unsatisfactory results and should not be used with respect to classifying wavelet measures. Neural networks generally yielded slightly better results than discriminant analysis, although the difference was not statistically significant. More decomposition levels and wavelets of longer support generally provided better results for the neural network classifier. The energy measure was usually associated with the lowest classification accuracy, whereas the mean measure yielded the highest or second highest accuracy in all cases. The forest class with smaller variability in wavelet measures had higher producer’s accuracy with discriminant analysis, whereas commercial class was better detected by neural networks. This performance difference in the two classifiers could be attributed to the assumptions of the methods and the distribution of the data. As expected, neural networks seemed to work better when the distribution of the data is not normal, and vice versa. The effect of the window size was apparent. Larger windows in most cases were associated with higher accuracy than smaller windows. For IKONOS image, a window size of at least 64mx64 m was needed to capture the texture features sufficiently for accurate classification in urban environment.
Chapter 5
Textural Discrimination of Urban and Suburban Neighborhoods of Various Socioeconomic Statuses

5.1 Introduction

In Chapter 4, wavelet analysis and artificial neural networks were combined together to classify textural samples of urban land uses. The land use classes considered in Chapter 4 belonged to the USGS Level III land-use/land-cover classes (Anderson et al. 1976). In this Chapter, experiments were carried out to test whether textural measures could be used to discriminate urban/suburban neighborhoods of different social-economic status. This went beyond the USGS Level III schemes and was an attempt to link spatial patterns on remotely sensed imagery to social-economic status on the ground (Liverman et al. 1998).

The identification of urban neighborhoods of various social-economic conditions is important to many governmental agencies and private companies. For the past half a century, urban sprawl has been a major trend of the urbanization process across the United States (Gillham 2002). The massive exodus of the middle and upper classes to the suburbs left behind run-down neighborhoods at the centers of cities. Many cities have taken steps to revitalize urban neighborhoods through such measures as renovation and gentrification. As a result, neighborhoods of different social-economic statuses co-exist in urban areas, ranging from run-down neighborhoods through renovated and gentrified neighborhoods to well-managed middle- and upper-class neighborhoods. Distinct textural patterns on remotely sensed imagery can be observed for urban neighborhoods of different social neighborhoods. These differences in spatial patterns are due to many factors. For example, low-income neighborhoods tend to have dense housing patterns, whereas middle- and upper-income neighborhoods are less compacted. Moreover, single-family housing generally has more tree cover than multifamily apartment housing. Multifamily housing, on the other hand, exhibits more regular linear or gridded
patterns. These observable differences among different neighborhoods provide the basis for the successful discrimination using textural measures. Through the study of textures, the identification offers a way to link remote sensing analysis to the social-economic aspects of human activities, i.e., to "socialize the texture", following the phrase "socializing the pixel" (Geoghegan et al. 1998).

The objectives of this study were to test the discriminative abilities of textural measures applied to urban/suburban neighborhoods using high-resolution imagery. Tests were carried out on different spectral bands and band combination, difference window sizes and different image resolutions.

5.2 Research Design

5.2.1 Study Area

Six urban/suburban residential neighborhoods were selected from East Baton Rouge parish, Louisiana. They were Melrose Place East, Spanish Town, St. Gerard Catholic Church, Jefferson Terrace, Garden District, and Woodgate-Woodstone. They represent the full spectrum of Baton Rouge neighborhoods in terms of median household income in 1999 according to Census 2000. Their locations are shown in Figure 5.1.

Melrose Place East is bordered by Florida Blvd on the south, N. Ardenwood Dr on the west, N. Donmore Ave on the east, and the railroad track on the north. It consists mostly of apartment complexes. It is a low-income neighborhood with median household income below $15,300 in 1999 (Census 2000). Spanish town is a historic community of downtown Baton Rouge. Dating back to 1805, Spanish Town holds the title of the oldest neighborhood in Baton Rouge and was added to the National Register of Historical Places in 1978 (http://www.nationalregisterofhistoricplaces.com/). The neighborhood has largely been gentrified. The majority of the historical homes have been renovated. The median household income of Spanish Town in 1999 was $20,511. The neighborhood around St. Gerard Catholic Church is considered old North Baton Rouge.
Figure 5.1. Locations of six urban/suburban neighborhoods in East Baton Rouge parish, Louisiana.
(Clarence 2004) with a median household income between $26,000 and $30,788 in 1999. The Jefferson Terrace neighborhood is located to the east of I-10, to the west of Airline Hwy, to the south of Bluebonnet Rd, and to the north of Siegen Ln. The median household income of the Jefferson Terrace neighborhood in 1999 was between $40,000 and $45,000. The Garden District is an old established upscale historical area and is bordered by Government Street to the north, LSU to the south, I-110 to the west (Clarence 2004). It has a median household income around $59,000 in 1999. The Woodgate-Woodstone community is located to the east of Highland Rd and to the south of Lee Dr. It is a high-income neighborhood with a median household income over $90,000 in 1999.

Based on their median household income in 1999, Melrose Place East is a very low-income neighborhood; Spanish Town a low-income neighborhood St. Gerard Catholic Church a lower middle income neighborhood; Jefferson Terrace a middle income neighborhood; Garden District a upper middle income neighborhood; and Woodgate-Woodstone a high-income neighborhood.

5.2.2 Data and Sampling Method

The U.S. Geological Survey (USGS) acquired high resolution color orthoimagery for 133 most populated metropolitan areas of the United States, including Baton Rouge metropolitan area. A number of one-foot resolution color aerial images were taken on March 22, 2002 in the Baton Rouge area by the USGS. Aerial photography was used for the orthoimagery captured on true color stable-base aerial photography film flown at 7,500 feet above average ground elevation. After the photograph was taken, the print was scanned. The approximate spectrum would be from 380 to 780nm, though this could be affected by the camera lens, film, atmospherics, and the scanning process. Each image was scanned and stored in RGB as with other standard image file.

The images had been geo-referenced and were in Universal Transverse Mercator (UTM)
projection. The images covered all the above six neighborhoods. One subset was taken for each neighborhood from the image where it was located. Each subset was entirely within a neighborhood to avoid boundary issues. The dimensions of each subset and their median household income in 1999 are listed in Table 5.1. The six subsets are shown in Figures 5.2, 5.3 and 5.4.

Each subset was divided into two small subsets (left and right). The left subsets were used for training purposes and the right subsets were used for accuracy assessment. For each training and validation subset, one hundred textural samples were randomly collected. Training textural samples might overlap with each other and validation textural samples might overlap with each other. But training textural samples and validation samples never overlapped because they were collected from each separate subset.

5.2.3 Window Size

Three window sizes were considered in this study: 45m×45m, 65m×65m and 85m×85m. Window size was expected to have a great effect on the classification accuracy. Small window sizes do not yield stable textural measures, whereas large window sizes can capture enough variations in textural patterns to yield stable wavelet measures. By examining the neighborhoods under study, a window of size 65m×65m was large enough to accommodate the spatial patterns in all neighborhoods. In comparison, the other two window sizes were also considered.

5.2.4 Band Combination

Different spectral bands have different textural patterns because of the difference in spectral response patterns. Myint et al. (2004) reported a substantial increase in overall accuracies when multiple bands were used for the discrimination of urban spatial features. The effects of multiple bands on the classification accuracy of urban neighborhoods were examined in this Chapter.
Table 5.1. Dimensions of subsets of six urban neighborhoods at 1-foot resolution and their median household income in 1999

<table>
<thead>
<tr>
<th>Neighborhood</th>
<th>Household Income class</th>
<th>Median Household Income ($) in 1999</th>
<th>Height(m) × Width(m)</th>
<th>Dimensions (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melrose Place East</td>
<td>Very low income</td>
<td>&lt;15,300</td>
<td>624.8 × 770.7</td>
<td>2083 × 2569</td>
</tr>
<tr>
<td>Spanish Town</td>
<td>Low income</td>
<td>20,511</td>
<td>328.2 × 334.8</td>
<td>1094 × 1116</td>
</tr>
<tr>
<td>St. Gerard Catholic Church</td>
<td>Lower middle income</td>
<td>26,000~30,788</td>
<td>581.7 × 563.1</td>
<td>1939 × 1877</td>
</tr>
<tr>
<td>Jefferson Terrace</td>
<td>Middle</td>
<td>40,000~45,000</td>
<td>448.5 × 563.1</td>
<td>1495 × 1429</td>
</tr>
<tr>
<td>Garden District</td>
<td>Upper middle income</td>
<td>59,000</td>
<td>394.5 × 608.7</td>
<td>1315 × 2029</td>
</tr>
<tr>
<td>Woodgate-Woodstone</td>
<td>High income</td>
<td>&gt;90,000</td>
<td>446.7 × 488.1</td>
<td>1489 × 1627</td>
</tr>
</tbody>
</table>
Figure 5.2. Image subsets of urban neighborhoods. Top: Melrose Place East; Bottom: Spanish Town.
Figure 5.3. Image subsets of urban neighborhoods. Top: St. Gerard Catholic Church; Bottom: Jefferson Terrace.
Figure 5.4. Image subsets of urban neighborhoods. Top: Garden District; Bottom: Woodgate-Woodstone.
5.2.5 Scale

Geographic phenomena operate at different scales (Lam et al. 2004). Spatial resolution is one of the most crucial issues in urban remote sensing (Welch 1982, Jensen and Cowen 1999). As the resolution becomes finer, texture information becomes more dominant, hence critical and indispensable for the interpretation of remotely sensed images. There is a range of resolutions at which texture patterns are discerned. Low resolution tends to smooth an image and to form a continuous gray-tone. High resolution tends to highlight the details and to form varied patterns. A forest is a continuous tone or color in a Landsat MSS or TM image. The tree patterns begin to appear in an IKONOS panchromatic image. As the resolution becomes much finer, leaf patterns may become dominant. Finding a range of characteristic scales suitable for the study of a particular geographic issue is a prime research area.

To examine the effects of image resolution on the spatial patterns, subsets at three resolutions were considered. The original image subsets were of 1-foot resolution. The subsets were down-sampled to generate subsets at 0.9m resolution and 2.7m resolution. For a pixel at 0.9m resolution or 2.7m resolution, the down-sampling averaged all the small pixels at the 1-foot resolution that make up the large pixel. As a result, three sets of six neighborhood image subsets at three resolution levels were subjected to analysis.

5.2.6 Textural Measures and Classification Method

As in Chapter 4, wavelet analysis was used to extract textural measures from decomposed subimages. A number of textural measures could be derived from the decomposed images. The means and standard deviations of the decomposed subimages at the first decomposition level were calculated in this study. Daubechies wavelet of order 8 was used as it generally yielded better results than Daubechies wavelet of order 2, as found in Chapter 4.

Although neural networks were found to yield slightly better classification accuracy in
Chapter 4 than discriminant analysis, the application of neural networks does require experimenting of various parameters. To minimize the effects of various parameters in neural networks on the discriminating capabilities, linear discriminant analysis, instead of neural networks, was used as the classifier.

5.3 Results

5.3.1 Test of Normality and Equal Covariance Matrices

One assumption required by linear discriminant analysis was the multivariate normal distribution of the input data among all classes ((Klecka 1980, Tatsuok and Lohnes 1988). The Kolmogorov-Smirnov test was conducted in SPSS™ to test the normality of the distribution of each wavelet measure from each subimage. At the 1-foot resolution level, regardless of the spectral band, the p-value from the Kolmogorov-Smirnov was less than 0.01 for some subimages, particularly mean values from the diagonally detailed subimage and standard deviation values from the approximate sumbimage and horizontally and vertically detailed subimages on the training dataset. This led to the rejection of the normal distribution of each corresponding wavelet measure. Since some wavelet measures were not normally distributed, the distribution of all wavelet measures was not normally distributed. The hypothesis of normality of wavelet measures was rejected.

Another major assumption about the data distribution required by linear discriminant analysis is equal covariance matrices among all classes. Box's M test was conducted to test the null hypothesis of equal covariance matrices in SPSS™. At the 1-foot resolution level, the p-value from the Box’s M test was 0.0 for all three bands based on the training dataset, leading to the rejection of the hypothesis of equal covariance matrices. (P-value and normality test)

Although the assumptions of normality and equal covariance were rejected, discriminant analysis was a robust method which could accommodate some deviations from the assumption (Tom and Miller 1984, Hubert 1994). In future studies, multinomial logistic regression may be
used as the classifier, which was found to outperform discriminant analysis slightly when the assumption of multivariate normality or equal covariance matrices is not met (Press and Wilson 1978).

### 5.3.2 Image Subsets of 1-Foot Resolution

For all decomposed subimages at the first decomposition level of each urban neighborhood, Tables 5.2, 5.3, and 5.4 list the means and standard deviations of the two measures (mean and standard deviation) computed from the training samples using window size of 85m×85m for Band 1, Band 2, and Band 3 respectively. The values are also plotted in Figures 5.5, 5.6, and 5.7, respectively. It can be observed that not a single measure of one decomposed subimage was sufficient for the discrimination of all neighborhoods. Measures from all decomposed subimages were needed to be combined together to achieve good classification results. For all three bands, generally speaking, St. Gerard Catholic Church neighborhood exhibited the largest variation in textural measures and the range of the measures overlap with the ranges of other neighborhoods to a large degree, which might result in low classification accuracy. Measures of other neighborhoods may show large variation for one textural measure at one particular decomposed subimage, but most of them separate very well for one textural measure at some decomposed images, which may contribute to higher producer’s accuracy and user’s accuracy.

#### 5.3.2.1 Band 1

Tables 5.5, 5.6, and 5.7 show the error matrixes of the classifications with neighborhood subsets using Band 1 and window sizes of 45m×45m, 65m×65m, and 85m×85m, respectively. One first observation is that the overall classification accuracy increased from 54.67% to 78.67%, and to 83.83% as the window size increased from 45m×45m to 65m×65m, and to 85m×85m. Clearly, a 45m×45m window was not large enough to capture the variations in spatial patterns present in Band 1. With a 65m×65m window, the accuracy increased substantially by
Table 5.2. List of means of the two measures (*mean* and *standard deviation*) of four subimages at the first decomposition level using Band 1, a window size of 85m by 85m and the db8 wavelet. Standard deviations of the measures are in the parentheses. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone. A: approximate subimage; H: horizontally detailed subimage; V: vertically detailed subimage; D: diagonally detailed subimage.

<table>
<thead>
<tr>
<th></th>
<th>Measure of <em>Mean</em></th>
<th>Measure of <em>standard deviation</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H</td>
</tr>
<tr>
<td>MP</td>
<td>340.13 (22.50)</td>
<td>9.31 (0.78)</td>
</tr>
<tr>
<td>ST</td>
<td>279.09 (11.50)</td>
<td>11.41 (0.28)</td>
</tr>
<tr>
<td>SG</td>
<td>205.45 (22.14)</td>
<td>10.94 (0.55)</td>
</tr>
<tr>
<td>JT</td>
<td>226.96 (16.34)</td>
<td>10.74 (0.47)</td>
</tr>
<tr>
<td>GD</td>
<td>216.63 (14.88)</td>
<td>12.04 (0.57)</td>
</tr>
<tr>
<td>WG</td>
<td>252.05 (14.93)</td>
<td>11.69 (0.60)</td>
</tr>
</tbody>
</table>
Table 5.3. List of means of the two measures (*mean* and *standard deviation*) of four subimages at the first decomposition level using Band 2, a window size of 85m by 85m and the db8 wavelet. Standard deviations of the measures are in the parentheses. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone. A: approximate subimage; H: horizontally detailed subimage; V: vertically detailed subimage; D: diagonally detailed subimage.

<table>
<thead>
<tr>
<th></th>
<th>Measure of <em>Mean</em></th>
<th>Measure of <em>standard deviation</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H</td>
</tr>
<tr>
<td>MP</td>
<td>327.30</td>
<td>10.64</td>
</tr>
<tr>
<td></td>
<td>(21.24)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>ST</td>
<td>287.41</td>
<td>14.48</td>
</tr>
<tr>
<td></td>
<td>(12.27)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>SG</td>
<td>212.68</td>
<td>13.17</td>
</tr>
<tr>
<td></td>
<td>(19.96)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>JT</td>
<td>232.08</td>
<td>14.46</td>
</tr>
<tr>
<td></td>
<td>(13.80)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>GD</td>
<td>238.43</td>
<td>16.80</td>
</tr>
<tr>
<td></td>
<td>(13.12)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>WG</td>
<td>244.50</td>
<td>13.07</td>
</tr>
<tr>
<td></td>
<td>(12.70)</td>
<td>(0.70)</td>
</tr>
</tbody>
</table>
Table 5.4. List of means of the two measures (*mean* and *standard deviation*) of four subimages at the first decomposition level using Band 3, a window size of 85m by 85m and the db8 wavelet. Standard deviations of the measures are in the parentheses. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone. A: approximate subimage; H: horizontally detailed subimage; V: vertically detailed subimage; D: diagonally detailed subimage.

<table>
<thead>
<tr>
<th></th>
<th>Measure of <em>Mean</em></th>
<th>Measure of <em>standard deviation</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H</td>
</tr>
<tr>
<td>MP</td>
<td>342.76</td>
<td>11.17</td>
</tr>
<tr>
<td></td>
<td>(21.80)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>ST</td>
<td>287.50</td>
<td>17.01</td>
</tr>
<tr>
<td></td>
<td>(13.02)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>SG</td>
<td>211.31</td>
<td>12.53</td>
</tr>
<tr>
<td></td>
<td>(20.72)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>JT</td>
<td>233.09</td>
<td>14.33</td>
</tr>
<tr>
<td></td>
<td>(13.93)</td>
<td>(0.58)</td>
</tr>
<tr>
<td>GD</td>
<td>221.12</td>
<td>15.26</td>
</tr>
<tr>
<td></td>
<td>(14.74)</td>
<td>(0.52)</td>
</tr>
<tr>
<td>WG</td>
<td>250.09</td>
<td>12.53</td>
</tr>
<tr>
<td></td>
<td>(13.36)</td>
<td>(0.42)</td>
</tr>
</tbody>
</table>


Figure 5.5. Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 1, a window size of 85m by 85m and the db8 wavelet. The center of each bar is the mean and the endpoints represent one standard deviation from the mean. From left to right, the plot corresponds to the approximate, horizontally, vertically, and diagonally detailed images respectively. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.
Figure 5.6. Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 2, a window size of 85m by 85m and the db8 wavelet. The center of each bar is the mean and the endpoints represent one standard deviation from the mean. From left to right, the plot corresponds to the approximate, horizontally, vertically, and diagonally detailed images respectively. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.
Figure 5.7. Mean and standard deviation plots of the two measures of four subimages at the first decomposition level using Band 3, a window size of 85m by 85m and the db8 wavelet. The center of each bar is the mean and the endpoints represent one standard deviation from the mean. From left to right, the plot corresponds to the approximate, horizontally, vertically, and diagonally detailed images respectively. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.
24%. An 85m×85m window reached the highest overall accuracy, i.e., 83.83%. Urban neighborhoods are composed of many elements on the ground, like houses, cars, trees, lawns and roads. Large windows help to accommodate these patterns on remotely sensed imagery.

For the window size of 85m×85m, the greatest confusion arose between Garden District (upper middle income) and St. Gerard Catholic Church neighborhood (lower middle income). This might be due to the fact that both neighborhoods were of low-density and showed similar patterns in Band 1. The highest producer’s accuracies occurred in Spanish Town and Woodgate-Woodstone, which meant that from a producer’s point of view, 100% of the samples belonging to Spanish Town on the ground were correctly classified as such, and 99% of the time an area belonging to Woodgate-Woodstone on the ground was correctly classified as Woodgate-Woodstone. The user’s accuracy achieved by both Melrose Place East and Jefferson Terrace was 100%, which meant that from a user’s point of view, 100% of the samples that were classified as either Melrose Place East or Jefferson Terrace were as such on the ground.

5.3.2.2 Band 2

Tables 5.8, 5.9, and 5.10 show the error matrixes of the classifications with neighborhood subsets using Band 2 and window sizes of 45m×45m, 65m×65m, and 85m×85m, respectively. As with Band 1, the overall classification accuracy increased as the window sized increased with Band 2. However, a 45m×45m window reached an overall accuracy of 80.83%, much higher than 54.67%, the accuracy achieved in Band 1. With a 65m×65m window, the accuracy increased 4.67%. An 85m×85m window reached the highest overall accuracy of 91.67%. It is obvious that urban neighborhoods were more easily differentiated using Band 2 (greed band) than Band 1 (red band). Band 2 was associated with more spatial information that helped to distinguish one neighborhood from another.

For the window size of 85m×85m, the largest confusion arose between Woodgate-
Table 5.5. Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 45m by 45m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>MP</td>
<td>60</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>15</td>
<td>75.00</td>
</tr>
<tr>
<td>ST</td>
<td>ST</td>
<td>18</td>
<td>80</td>
<td>9</td>
<td>19</td>
<td>39</td>
<td>15</td>
<td>44.44</td>
</tr>
<tr>
<td>SG</td>
<td>SG</td>
<td>0</td>
<td>0</td>
<td>42</td>
<td>5</td>
<td>14</td>
<td>1</td>
<td>67.74</td>
</tr>
<tr>
<td>JT</td>
<td>JT</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>51</td>
<td>8</td>
<td>0</td>
<td>72.86</td>
</tr>
<tr>
<td>GD</td>
<td>GD</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>19</td>
<td>26</td>
<td>0</td>
<td>50.98</td>
</tr>
<tr>
<td>WG</td>
<td>WG</td>
<td>20</td>
<td>9</td>
<td>42</td>
<td>6</td>
<td>11</td>
<td>69</td>
<td>43.95</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>60.00</td>
<td>80.00</td>
<td>42.00</td>
<td>51.00</td>
<td>26.00</td>
<td>69.00</td>
<td>54.67 (overall accuracy)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.6. Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 65m by 65m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MP</td>
<td>89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>3</td>
<td>99</td>
<td>1</td>
<td>2</td>
<td>32</td>
<td>6</td>
<td>69.23</td>
</tr>
<tr>
<td></td>
<td>SG</td>
<td>3</td>
<td>0</td>
<td>43</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>69.35</td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>80</td>
<td>0</td>
<td>1</td>
<td>96.39</td>
</tr>
<tr>
<td></td>
<td>GD</td>
<td>1</td>
<td>1</td>
<td>34</td>
<td>2</td>
<td>68</td>
<td>0</td>
<td>64.15</td>
</tr>
<tr>
<td></td>
<td>WG</td>
<td>3</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>93</td>
<td>79.49</td>
</tr>
<tr>
<td></td>
<td>Producer’s Accuracy(%)</td>
<td>89.00</td>
<td>99.00</td>
<td>43.00</td>
<td>80.00</td>
<td>68.00</td>
<td>93.00</td>
<td>78.67 (overall accuracy)</td>
</tr>
</tbody>
</table>
Table 5.7. Error matrix for neighborhood discrimination using Band 1 of the 1-foot aerial image with 85m by 85m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>93 0 0 0 0 0 100.00</td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>3 100 1 2 28 1 74.07</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>1 0 57 15 0 0 78.08</td>
<td></td>
</tr>
<tr>
<td>JT</td>
<td>0 0 0 82 0 0 100.00</td>
<td></td>
</tr>
<tr>
<td>GD</td>
<td>1 0 32 1 72 0 67.92</td>
<td></td>
</tr>
<tr>
<td>WG</td>
<td>2 0 10 0 0 99 89.19</td>
<td></td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>93.00 100.00 57.00 82.00 72.00 99.00 83.83 (overall accuracy)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.8. Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 45m by 45m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MP</td>
<td>ST</td>
</tr>
<tr>
<td>MP</td>
<td>86</td>
<td>0</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>SG</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>WG</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>86.00</td>
<td>94.00</td>
</tr>
</tbody>
</table>
Table 5.9. Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 65m by 65m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>94.74</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>100</td>
<td>2</td>
<td>5</td>
<td>15</td>
<td>2</td>
<td>80.65</td>
</tr>
<tr>
<td>SG</td>
<td>5</td>
<td>0</td>
<td>70</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>85.37</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>85</td>
<td>0</td>
<td>92</td>
<td>92.39</td>
</tr>
<tr>
<td>WG</td>
<td>5</td>
<td>0</td>
<td>28</td>
<td>6</td>
<td>0</td>
<td>92</td>
<td>70.23</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>90.00</td>
<td>100.00</td>
<td>70.00</td>
<td>76.00</td>
<td>85.00</td>
<td>92.00</td>
<td>85.50 (overall accuracy)</td>
</tr>
</tbody>
</table>
Table 5.10. Error matrix for neighborhood discrimination using Band 2 of the 1-foot aerial image with 85m by 85m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MP</td>
<td>ST</td>
</tr>
<tr>
<td>MP</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>SG</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WG</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>94.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Woodstone and St. Gerard Catholic Church neighborhood. The highest producer’s accuracies belonged to Spanish Town and Woodgate-Woodstone, 100% and 99% respectively, as was the case in Band 1. The user’s accuracy achieved by both Melrose Place East and Jefferson Terrace was 100%, which meant that 100% of the time, an area that was categorized as either Melrose Place East or Jefferson Terrace was as such on the ground. This was also the same in Band 1.

5.3.2.3 Band 3

Tables 5.11, 5.12, and 5.13 show the error matrixes of the classifications with neighborhood subsets using Band 3 and window sizes of 45m×45m, 65m×65m, and 85m×85m, respectively. As with Band 1 and Band 2, the overall classification accuracy increased as the window size increased with Band 3. A 45m×45m window reached an overall accuracy of 78.5%, a little bit lower than that with Band 2. The 65m×65m and 85m×85m windows yielded accuracies slightly higher than those with Band 2.

For the window size of 85m×85m, the largest confusion arose between Woodgate-Woodstone and St. Gerard Catholic Church neighborhood, as was the case with Band 2. The highest producer’s accuracies belonged to Woodgate-Woodstone and Spanish Town, 100% and 99% respectively. The user’s accuracy achieved by Melrose Place East, Jefferson Terrace, and St. Gerard Catholic Church was 100%, which meant that 100% of the time, an area that was categorized as one of them was as such on the ground. This was also the same in Band 1.

5.3.2.4 Band Combination

Figure 5.8 plots the overall accuracies of the three window sizes for each band. It was evident that Band 2 and Band 3 yielded much higher accuracy than Band 1 regardless of window size.

It is expected that multiple bands should help to improve the classification accuracy in urban neighborhood discrimination. Since Band 2 and Band 3 performed much better than Band 1, the
Table 5.11. Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 45m by 45m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>84.09</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>98</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>86.73</td>
</tr>
<tr>
<td>SG</td>
<td>6</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>85.11</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>85</td>
<td>2</td>
<td>0</td>
<td>79.44</td>
</tr>
<tr>
<td>WG</td>
<td>20</td>
<td>0</td>
<td>51</td>
<td>0</td>
<td>0</td>
<td>83</td>
<td>0</td>
<td>53.90</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>74.00</td>
<td>98.00</td>
<td>40.00</td>
<td>91.00</td>
<td>85.00</td>
<td>83.00</td>
<td>78.50 (overall accuracy)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.12. Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 65m by 65m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>94.44</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>95.19</td>
</tr>
<tr>
<td>SG</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>87.16</td>
</tr>
<tr>
<td>WG</td>
<td>15</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>60.90</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>85.00</td>
<td>99.00</td>
<td>48.00</td>
<td>93.00</td>
<td>95.00</td>
<td>95.00</td>
<td>85.83 (overall accuracy)</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.13. Error matrix for neighborhood discrimination using Band 3 of the 1-foot aerial image with 85m by 85m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>94.29</td>
</tr>
<tr>
<td>SG</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>94</td>
<td>0</td>
<td>94.95</td>
</tr>
<tr>
<td>WG</td>
<td>11</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>76.34</td>
</tr>
<tr>
<td><strong>Producer’s Accuracy(%)</strong></td>
<td><strong>89.00</strong></td>
<td><strong>99.00</strong></td>
<td><strong>78.00</strong></td>
<td><strong>98.00</strong></td>
<td><strong>94.00</strong></td>
<td><strong>100.00</strong></td>
<td><strong>93.00 (overall accuracy)</strong></td>
</tr>
</tbody>
</table>
Figure 5.8. Overall accuracies of three window sizes: 45m by 45m, 65m by 65m, and 85m by 85m using each individual band of one-foot aerial image.
two bands were combined for the classification. In combining the two bands, the dimension of the independent variables in the discriminant analysis was doubled, i.e., the means and standard deviations from the decomposed sub-images by wavelet analysis from both Band 2 and Band 3 were combined and subjected to discriminant analysis.

Tables 5.14, 5.15, and 5.16 show the error matrixes of the classifications with neighborhood subsets using the combination of Band 2 and Band 3 and window sizes of 45m×45m, 65m×65m, and 85m×85m, respectively. A considerable increase in classification accuracy was achieved by the two combined bands over any individual band regardless of window size. With a 45m×45m window, the combination yielded a 92% overall accuracy, much higher than 80.83%, which was the highest achieved by one single band (Band 2). With a 65m×65m window, the combination produced an overall accuracy of 94.5%, almost 10% higher than 85.83%, the highest achieved by one single band (Band 3). With an 85m×85m window size, the combination yielded a 96.83% overall accuracy, over 3% higher than 93%, the highest achieved by one single band (Band 3).

When the window size was 85m×85m, the producer’s accuracies and user’s accuracies for many classes were 100%. Spanish Town, Jefferson Terrace, and Garden District were successfully separated from other classes. The confusion still existed between St. Gerard Catholic Church and Woodgate-Woodstone, causing their producer’s accuracies or user’s accuracies lower than those of other classes.

5.3.3 Image Subsets of 0.9m Resolution and 2.7m Resolution

Figures 5.9 and 5.10 plot the overall accuracies at the 0.9m resolution and 2.7m resolution, respectively. At both resolution levels, larger window sizes produced higher accuracies irrespective of the band. At the 0.9m resolution, Band 1 had the poorest performance regardless of window size. However, at the 2.7m resolution level, Band 2 performed the poorest regardless of window size. In comparison, Band 3 performed the more stably.
Table 5.14. Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 45m by 45m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>86</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>SG</td>
<td>11</td>
<td>0</td>
<td>73</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td></td>
<td>83.91</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td></td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>100</td>
<td>1</td>
<td></td>
<td>97.09</td>
</tr>
<tr>
<td>WG</td>
<td>3</td>
<td>0</td>
<td>27</td>
<td>1</td>
<td>0</td>
<td>97</td>
<td></td>
<td>75.78</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>86.00</td>
<td>100.00</td>
<td>73.00</td>
<td>96.00</td>
<td>100.00</td>
<td>97.00</td>
<td></td>
<td>92.00 (overall accuracy)</td>
</tr>
</tbody>
</table>
Table 5.15. Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 65m by 65m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>Reference</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MP</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>SG</td>
<td>7</td>
<td>0</td>
<td>77</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>90.59</td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>GD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>1</td>
<td>99.00</td>
</tr>
<tr>
<td></td>
<td>WG</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>80.49</td>
</tr>
<tr>
<td></td>
<td>Producer’s Accuracy(%)</td>
<td>92.00</td>
<td>100.00</td>
<td>77.00</td>
<td>99.00</td>
<td>100.00</td>
<td>99.00</td>
<td>94.50 (overall accuracy)</td>
</tr>
</tbody>
</table>
Table 5.16. Error matrix for neighborhood discrimination using the combination of Band 2 and Band 3 of the 1-foot aerial image with 85m by 85m window size. MP: Melrose Place East; ST: Spanish Town; SG: St. Gerard Catholic Church; JT: Jefferson Terrace; GD: Garden District; WG: Woodgate-Woodstone.

<table>
<thead>
<tr>
<th>Classified</th>
<th>MP</th>
<th>ST</th>
<th>SG</th>
<th>JT</th>
<th>GD</th>
<th>WG</th>
<th>User’s Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>ST</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>SG</td>
<td>5</td>
<td>0</td>
<td>86</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>94.51</td>
</tr>
<tr>
<td>JT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>GD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>WG</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>87.72</td>
</tr>
<tr>
<td>Producer’s Accuracy(%)</td>
<td>95.00</td>
<td>100.00</td>
<td>86.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>96.83 (overall accuracy)</td>
</tr>
</tbody>
</table>
Figure 5.9. Overall accuracies of three window sizes: 45m by 45m, 65m by 65m, and 85m by 85m using each individual band of an aerial image of 0.9 meter.
Figure 5.10. Overall accuracies of three window sizes: 45m by 45m, 65m by 65m, and 85m by 85m using each individual band of an aerial image of 2.7 meter.
Table 5.17. Overall accuracies (%) for neighborhood discrimination at three window sizes at three bands and at three resolutions.

<table>
<thead>
<tr>
<th>Window size (m)</th>
<th>1-foot resolution</th>
<th>0.9m resolution</th>
<th>2.7m resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Band 1</td>
<td>Band 2</td>
<td>Band 3</td>
</tr>
<tr>
<td>45 by 45</td>
<td>54.67</td>
<td>80.83</td>
<td>78.50</td>
</tr>
<tr>
<td>65 by 65</td>
<td>78.67</td>
<td>85.50</td>
<td>85.83</td>
</tr>
<tr>
<td>85 by 65</td>
<td>83.83</td>
<td>91.67</td>
<td>93.00</td>
</tr>
</tbody>
</table>
Table 5.17 lists the overall accuracies at all three resolution levels. One observation is that the classification accuracy was higher at one-foot resolution that at 0.9m and 2.7m resolution levels in general. There were two exceptions with Band 1 at one-foot resolution. The accuracy of 54.67% with window size of 45m×45m with Band 1 at one-foot resolution was lower than those at 0.9m and 2.7m resolution. The accuracy of 83.83% with window size of 85m×85m with Band 1 at one-foot resolution was lower than 86.5% at 2.7m resolution.

For Band 3 which performed the most stably, 1-foot resolution achieved the highest accuracy regardless of window size, followed by 2.7m resolution. The 0.9m resolution performed comparably well with the 2.7m resolution.

5.4 Conclusion

In this Chapter, wavelet measures were extracted from textural samples of urban neighborhoods in East Baton Rouge parish, Louisiana and classified using linear discriminant analysis.

The effects of window size on the discrimination of textural samples were manifest. The larger the window size, the higher the classification accuracy. When the window size increased from 45m by 45m to 65m by 65m and to 85m by 85m, the overall accuracy was found to increase at least by 5% for each increase in window size, irrespective of spectral band and image resolution. This finding was consistent with the findings in Chapter 4. As discussed in Chapter 1, large window sizes usually are linked with large edge effects if traditional moving-window method is used in the classification of one entire image. If the method of reducing edge effects in Chapter 1 was used, a reasonably large window, as long as it does not exceed the minimum class polygon, should produce much better results in terms of classification accuracy.

Band selection and combination were important to achieve high accuracies. Band 3 was found to perform more consistently at different resolution levels. The combination of Band 2 and
Band 3 at the 1-foot resolution level was found to improve the classification accuracy drastically, reaching an overall accuracy of 96.83%. The spectral range of the aerial orthoimagery used in this study was mostly in the visible range. Infrared bands provide additional information beyond that in the visible bands, particularly on the identification of vegetation, which may enhance the discrimination of urban neighborhoods if used.

Image resolution was found to be an important factor in urban neighborhood discrimination based upon wavelet measures. With all three window sizes considered in this study, i.e., 45m by 45m to 65m by 65m and to 85m by 85m, the 1-foot resolution subsets were found to yield higher classification accuracy than the 0.9m resolution subsets and the 2.7m resolution subsets for both band 2 and band 3. The differences were generally over 5%.

There was consistent confusion between some neighborhoods, particularly the Woodgate-Woodstone and St. Gerard Catholic Church neighborhoods. Other neighborhoods were found to be easily distinguished when multiple bands were used. Future research may test other textural measures to examine whether confusion between the two neighborhoods could be resolved. More complex classifiers might also be applied to improve the accuracies.

In this study, six urban neighborhoods of different income levels were subjected to analysis. Future study should further investigate whether neighborhoods at the same income level exhibit similar textural patterns. If so, textural patterns hold the promise to predict the income level of a neighborhood. Otherwise, studies should be carried out to test whether neighborhoods at the same income level could be discriminated from each other successfully.
Chapter 6  
Conclusion

Urban remote sensing is one of the most fascinating and challenging realms in remote sensing. High-resolution remote sensing imagery provides new opportunities for studying urban areas through the prominent textural patterns present on the images. This dissertation investigated how spatial patterns on high-resolution images in urban areas can be utilized to enhance the classification accuracy and whether the spatial patterns can be used to detect the social-economic conditions on the ground. The identification and characterization of urban neighborhoods of different social-economic statuses holds the promises of opening up a new avenue to link remotely sensed images to the social-economic aspects of human activities.

6.1 Summary of Findings

The goals of this study were threefold: 1) to determine whether artificial neural networks in combination with wavelet analysis are effective for the characterization and classification of urban land covers; 2) to determine whether wavelet textural measures can be associated with the social-economic conditions of urban neighborhoods; and 3) to explore the effects of image resolution on discrimination of urban neighborhoods. The hypotheses were: 1) ANNs yielded higher classification accuracy than linear discriminant analysis and the minimum-distance classifier based on wavelet measures of urban land covers; 2) wavelet textural measures could be used to efficiently discriminate among urban neighborhoods of different social-economic conditions; 3) image resolution had great influences on the discrimination of urban neighborhoods; and 4) window size had great influences on the discrimination of urban neighborhoods. In addition, two technical problems related to the application of textural approach, including the edge effect and image segmentation problem, were examined.

On the issue of edge effects, the proposed new approach consistently achieved higher
accuracy than the traditional approach. The accuracy of the new approach generally increased with increasing window size and it leveled off after a threshold, whereas the accuracy of the traditional approach generally decreased with increasing window size after a threshold. The new approach was capable of classifying border pixels, while the traditional method left a strip of border pixels unprocessed. The new approach had a smoothing effect. Small isolated clusters, present in the classified images created by the traditional method, tended to be eliminated in the classified images by the new approach. The side length of the window could be an even number in the new approach because it does not use the concept of the center of the window, whereas in the traditional approach, the window size was usually odd to make sure the center pixel was the exact center.

The performance of the splitting-and-merging segmentation procedures was found to be sensitive to the selection of the parameters. This is particularly the case for complex remotely sensed images. For hierarchical splitting, the maximum widow size and the stopping criterion $X$ were critical. For agglomerative merging, the stopping criteria affected the segmentation greatly, leading to over-segmentation or under-segmentation. The segmentation was less affected by the parameters in the pixel-wise refinement state. The parameters had to be determined experimentally and varied for different images. The combination of segmentations by different textural measures proved to be helpful in identifying homogeneous regions in an image. Some clusters left out by one measure were picked up by another measure. The post-segmentation integration reflected all the clusters identified by two or more measures. Compared with other pre-segmentation weighting schemes, this method did not involve the determination of weighting coefficients subjectively.

Regarding the four hypotheses, (1) the performance of neural networks as a textural classifier was compared with that of discriminant analysis and minimum distance classifier based on wavelet measures of an IKONOS image of Atlanta, Georgia. The minimum-distance classifier
performed the worst. Neural networks were found to generally yield slightly better results than discriminant analysis but the difference was not statistically significant. The first hypothesis was shown to be invalid. More decomposition levels and wavelets of longer support generally provided better results for the neural network classifier. The performance difference in the two classifiers could be attributed to the assumptions of the methods and the distribution of the data. Neural networks seemed to work better when the distribution of the data is not normal. (2) Six urban neighborhoods of various social-economic statuses were selected from the USGS 1-foot resolution aerial images in Baton Rouge, Louisiana. Wavelet measures were extracted from textural samples of six neighborhoods and classified using linear discriminant analysis. For the selected six neighborhoods, with the largest window size of 85m by 85m considered in this study, an overall accuracy of 93.00% was achieved using band 3 and an overall accuracy of 96.83% was achieved using combination of band 2 and band 3. (3) Image resolution was found to be an important factor in urban neighborhood discrimination using wavelet measures calculated for the six neighborhoods. Based on three window sizes considered in this study, i.e., 45m by 45m, 65m by 65m and 85m by 85m, the 1-foot resolution subsets were found to yield higher classification accuracy than the 0.9m resolution subsets and the 2.7m resolution subsets for both band 2 and band 3. The differences were generally over 5%. (4) Window size was found to have great influences on the discrimination of urban neighborhoods in Baton Rouge, Louisiana. The larger the window size, the higher the classification accuracy. When the window size increased from 45m by 45m to 65m by 65m and to 85m by 85m, the overall accuracy was found to increase at least by 5% for each increase in window size, irrespective of spectral band and image resolution.

6.2 Suggestions for Future Research

On the issue of edge effects, the experiments were carried out on only the panchromatic band
of the IKONOS images in this dissertation. Future research may test the performance of the proposed approach on multispectral images and data from other sensors. Other textual measures and classifiers should be investigated further. Images in the three scenarios studied were relatively simple. More complex images with irregular boundaries can be tested further in the future.

On region-based split-and-merge segmentation, various parameters were applied to an entire image without any adaptation scheme to take the local variations into consideration. For some parts of an image, large parameter values may be needed to yield a satisfactory segmentation, whereas for other parts, small parameters are needed because of the difference in the complexity in different regions. Global parameters are hardly expected to accommodate different complexity in an image. In future study, local parameters may be used to achieve better segmentations. Textural measures played a critical role in the success of the splitting-and-merging procedures. Besides the local binary pattern/contrast (LBP/C), color, intensity measure, other textural measures may also be adopted in the segmentation framework.

On the performance of wavelet analysis and neural network approach, further exploring the combination of wavelet, decomposition level, window size, signature, and network structure will gain insights on the combined performance of wavelet analysis and neural network approach.

On the discrimination of urban neighborhoods, there was persistent confusion among some neighborhoods. Although wavelet analysis is highly efficient in capturing textural patterns, it cannot be expected to excel in all circumstances. The confusion might be mitigated to some degree if other appropriate texture measures are adopted. On the issue of scale, three resolution levels were tested. Future research may test the performance at more resolution level in an attempt to find the best scale at which urban neighborhoods could be discriminated. Images from different sensors should also be investigated to see whether some sensors are good at
capturing images which make textural patterns on urban neighborhoods more distinct from each other.
References


Geoghegan, J., L. Pritchard, Y. Ogneva-Himmelberger, R. R. Chowdhury, S. Sanderson, and B. L. Turner. 1998. "Socializeing the pixel" and "pixelizing the social" in land-sue and land-


Myint, S. W. 2001. Wavelet Analysis And Classification Of Urban Environment Using High-Resolution Multispectral Image Data, Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA.


Zhao, W. 2001. Multiscale analysis for characterization of remotely sensed images Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA, U.S.A.

Vita

Guiyun Zhou was born in 1977 in Jiangsu, China. He grew up in the countryside of northern Jiangsu. He received his Bachelor of Science degree in geography in 1999 and his Master of Science degree in cartography and geographical information science in 2002 from Beijing University. He enrolled at Louisiana State University in the fall of 2002 to work toward a degree of Doctor of Philosophy in geography.