USING REMOTE-SENSING AND GIS TECHNOLOGY FOR AUTOMATED

BUILDING EXTRACTION

A Dissertation
Presented to
The Academic Faculty

by

Liora Sahar

In partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Design Computing in the
College of Architecture

Georgia Institute of Technology
December, 2009

USING REMOTE-SENSING AND GIS TECHNOLOGY FOR AUTOMATED

BUILDING EXTRACTION

Approved by:

Dr. Steven P. French, Advisor
College of Architecture
Georgia Institute of Technology

Prof. Nickolas Faust
GTRI - Georgia Tech Research Institute
Georgia Institute of Technology

Dr. William Drummond
College of Architecture
Georgia Institute of Technology

Prof. Charles Eastman
College of Architecture
Georgia Institute of Technology

Dr. Shamkant Navathe
College of Computing
Georgia Institute of Technology

Date Approved: July 28th, 2009

DEDICATION

To my Husband, Yinon

ACKNOWLEDGEMENTS

The achievements of the past several years have not been possible without the help of numerous friends, colleagues, teachers and family members. First and foremost, I would like to acknowledge the endless support and love of my husband Yinon. There are no words to describe how grateful I am for your encouragement during the good but mainly the hard times. There is definitely a toll that my entire family had to incur throughout that time period. My three lovely daughters, Stav, Keren and Amy, who was born in the process, have all contributed in their own special way. Their love, understanding and constant reminder of what is truly important, made all the difference. I would like to thank my wonderful parents for their faith in me and their unconditional love and support all those years.

This work would not have been possible without the support, guidance and commitment of my advisor, Dr. Steven French. His contribution goes beyond this work to provide me with the knowledge and tools to pursuit my professional ambitions within the GIS arena. I would like to specially acknowledge and thank Prof. Nick Faust for the significant contribution in the fields of Remote Sensing and image processing. Your insight, constant feedback and guidance were invaluable to the accomplishment of this dissertation. I would also like to thank my committee members, Dr. William Drummond, Dr. Shamkant Navathe and Prof. Charles Eastman for their comments, patience and time. My sincere thanks go to all research scientists and students at the Center for GIS. I would like to specially thank Dr. Subrahmanyam Muthukumar for the constant support and stir of ideas as well as my dear office mate Ning Ai.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xv

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
DTM	Digital Terrain Model
DSM	Digital Surface Model
ESRI	Environmental Systems Research Inc.
GIS	Geographical Information System (Science)
ISODATA	Iterative Self Organizing Data
k-NN	k- Nearest neighbor
LIDAR	Light Detection and Ranging
NN	Nearest neighbor
RS	Remote Sensing
GLCM	Grey Level Co-occurrence Matrix
GDAL	Geospatial Data Abstraction Library

SUMMARY

Extraction of buildings from remote sensing sources is an important GIS application and has been the subject of extensive research over the last three decades. An accurate building inventory is required for applications such as GIS database maintenance and revision; impervious surfaces mapping; storm water management; hazard mitigation and risk assessment. Despite all the progress within the fields of photogrammetry and image processing, the problem of automated feature extraction is still unresolved.

A methodology for automatic building extraction that integrates remote sensing sources and GIS data was proposed. The methodology consists of a series of image processing and spatial analysis techniques. It incorporates initial simplification procedure and multiple feature analysis components. The extraction process was implemented and tested on three distinct types of buildings including commercial, residential and high-rise. Aerial imagery and GIS data from Shelby County, Tennessee were identified for the testing and validation of the results. The contribution of each component to the overall methodology was quantitatively evaluated as relates to each type of building. The automatic process was compared to manual building extraction and provided means to alleviate the manual procedure effort.

A separate module was implemented to identify the 2D shape of a building. Indices for two specific shapes were developed based on the moment theory. The indices were tested and evaluated on multiple feature segments and proved to be successful.

The research identifies the successful building extraction scenarios as well as the challenges, difficulties and drawbacks of the process. Recommendations are provided based on the testing and evaluation for future extraction projects.

Chapter 1   INTRODUCTION

1.1   General

Building footprints were shown to be very useful for a wide variety of applications. Two dimensional as well as three dimensional representations of buildings are commonly used within numerous routine civil and military operations. From establishing and managing a GIS system for a city, to urban planning and even high-tech military urban combat training, building footprints are an essential part of many daily functions within the private and public sectors. Building footprints can provide valuable information for natural hazard risk assessment, hazard mitigation and prepare for efficient emergency response. For example, building footprints and actual building shape can be used for earthquake risk assessment. The behavior of buildings under earthquake stresses is affected by multiple parameters including the symmetry of the structure. Simulation based on actual building footprints can better evaluate the damage that an area may endure as a result of an earth quake and allow advance preparation. When any natural or man-made hazard occurs, emergency response operations can greatly benefit from an updated building database that provides reliable information about possible location of individuals. Great effort is put into the development of building data sets for cities and counties all over the world. Building layers are used for urban GIS mapping, urban planning as well as resource management operations that can potentially produce revenue. For example, storm-water management requires building areas as part of impervious surface delineation.   The definition of a building may vary by application

and hence entails different building characteristics. Some applications may require general footprint information and focus on the symmetry of the shapes while other applications may need accurate corner locations and be particular with regards to attached structures (such as parking decks, balconies, garages). Gathering that information requires a significant initial effort as well as time and labor consuming update processes. Automatic extraction of buildings footprints from aerial imagery can considerably reduce the cost at all stages.

Building extraction from aerial imager poses several major difficulties that any extraction process has to overcome. Parts of the building may be obstructed from view by surrounding objects and shadows, edges of the building may be fuzzy (owing to similarity to the surrounding surfaces or sun-illumination issues), buildings vary in shapes (footprint of the roof), sizes and colors (not solid color within the roof), buildings appear different from different perspectives and much of the 3D information is omitted in a 2D image, and buildings may also contain islands of other feature with different colors such as vents and AC units. Identifying all the characteristics of buildings requires operationalizing the logic of a human operator in order to distinguish a building from its surroundings.

The method developed and demonstrated here integrates readily available remote sensing and GIS data along with image processing techniques in order to identify building footprints. By integrating existing data in the context of a knowledge base, containing data derived from advanced technologies and methods, it will be possible to

produce inventories that are more accurate and cost effective than existing approaches. Data sources would include, but not be limited to aerial photography and local tax assessor parcel data.

The proposed approach is described in figure 1:

Primary Data	Analysis	Database

Ortho-rectified images
 	
Tax-assessor parcels
    Parcel geometry
    Parcel location
    Tax data attributes
	Feature Extraction Process
	Preprocessing
	
	Initial Analysis
	
	Analysis, Validation, Elimination
	
	Output processing
		Building Inventory
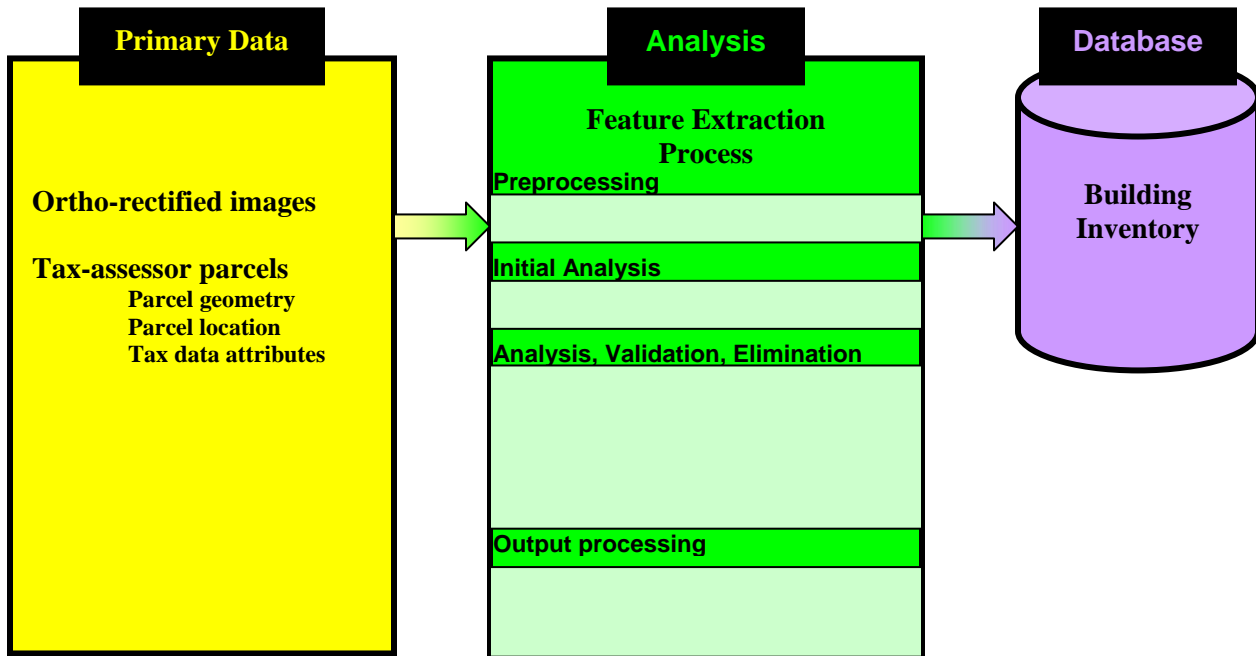
Figure 1 – Methodology of the building extraction approach

1.2   Problem Statement

This research is primarily aimed at automating the building footprint extraction process from remotely sensed sources, and as a corollary, minimizing the need for human intervention. The automated extraction process will be based on simple parameters either available or derived from the scene and parcel data, and will not require prior knowledge or expertise in photogrammetry or remote sensing. The research project will also evaluate the robustness of several extant techniques for building extraction. The study will also implement the generated techniques by integrating both vector and raster data sets in a new manner, to achieve a more complete and reliable solution.

1.3   Significance and contribution to the field

The contribution of the proposed methodology can be evaluated as contribution to the industry and academic research that demands building inventory, contribution to the automated feature extraction effort within the photogrammetry and remote sensing discipline and contribution to both image processing and photogrammetry by introducing image processing techniques rarely used within the remote sensing and photogrammetry field (such as the moment theorem) towards building extraction.   Moreover, the methodology will attempt to expand work done within image processing and can be used for processes other then building extraction.

There is evidence in the literature that supports the need for more research on automated systems for feature extraction that combines geographical information from different sources and uses GIS data as a-priori knowledge (Brenner, 2005; Baltsavias, 2004).   The methodology as presented introduces a new overall approach to building footprint extraction.   The integration of GIS and remote sensing sources as presented has not been implemented and tested as an entire approach to solving the problem. Simplification algorithms have been evaluated and tested in previous projects.   Hence, using parcel geometry and parcel attributes for simplification purposes will extend work done by Wijnant and Steenberghen (2004), Ming et al (2005) and Ohlhof, et al (2004) by evaluating the added value of using readily available parcels layers and attribute information for simplifying the task.

The automated building extraction procedure may be developed into additional inventory (roads, sidewalks etc) development tools in GIS and would enhance and benefit

a wide variety of applications. Building locations are required for day to day management of cities and counties and for more complex applications such as evaluating damage after an earthquake. All those applications can benefit from a methodology/automated procedure that can produce a large percentage of the building inventory and hence, maintain an updated inventory. Brenner (2005) anecdotally mentions a German city that acquired about 30,000 km² of features. Each building required several points, which required a huge effort. The company estimated that an update of the area will require about 70% of the initial effort. That number emphasized the concept that building extraction does not require only an initial investment, but is an on going expenditure.

The photogrammetry and remote sensing field has been attempting to develop automated and semi-automated approaches for feature extraction and in particular building extraction over the last 15-20 years. Today, we still do not have an "accepted" methodology to extract buildings from aerial imagery and therefore we normally digitize those features manually. An automated approach that can be easily replicated and takes advantage of readily available sources may contribute to that effort. The work can be viewed as a direct continuation/expansion of the work by Huertas and Nevatia (1988) that pioneered the usage of geometry and shadows for the purpose of building extraction based on edge detection, and the work of Irvin and McKeown (1989) that used shadows in different stages of the extraction process. The methodology extends many projects that concentrated on extracting specific types of buildings such as Kim et al (2004) that developed a methodology to extract large rectangular buildings. The methodology also expands the approach taken by many research projects that involve semi-automatic tools

with more considerable user intervention (especially for simplification), such as Müller and Zaum (2005) that uses seed growing mechanism and Sahar and Krupnik (1999) that initially break the image manually into regions of interest. The methodology is aimed at high-resolution (1ft) imagery that is the current standard for urban aerial imagery and more elaborated than the more heavily tested 1m resolution aerial and satellite (IKONOS) imagery.

The work of Hu (1962), Rosin (1999) and Rosin (2003) is used and implemented in the proposed methodology. Although used within the image processing discipline, the moment theorem has not been commonly and heavily applied within the photogrammetry and remote sensing field for building extraction. Evaluation of using this theory for the purpose of identifying building segments can contribute to the long effort of extracting buildings and possibly other types of features. As mentioned above, we attempt to specify index not only for rectangular shapes (Rosin, 2003), but for the "I" and "O" shapes, from the common L, T, C, I, H, O building footprints. Successful shape identification extends the work of Rosin (2003), Reiss (1991) and Schweitzer and Straach (1998) that evaluate properties of specific shapes based on moment invariants.

1.4   Organization of Dissertation

This dissertation is organized into six chapters. The introduction chapter describes the need for automated building extraction procedures. The chapter details the problems involving the extraction procedure and the added value of the proposed methodology. The introduction also defines the scope of the dissertation.

7

Chapter 2 reviews the current state of building extraction procedures within the photogrammetry and remote sensing field. The review includes a survey of extraction procedures using different types of imagery, including aerial imagery, satellite imagery, LIDAR, and RADAR. The review details the different image processing techniques used for the purpose of feature extraction and building extraction in particular as well as shape recognition techniques using the moment theorem. The chapter also includes the basis for the motivation of incorporating GIS data in the methodology as well as moving from global image processing to a local image processing approach.

Chapter 3 presents the current methodology of the project. The chapter illustrates the flow of the building extraction model and provides a general description for each analysis phase. The chapter is followed by a detailed description of the methodology implementation. The description in chapter 4 includes the tools, algorithms and techniques used to implement the image partitioning, segmentation, feature analysis and generalization of the buildings outlines. Chapter 4 contains two sections. The first section presents the implementation details and the second section presents and evaluates the results of the testing. The results evaluation section begins with general details about the testing area and the datasets. The general information is followed with a test plan for different types of buildings within the testing area. The testing evaluation includes a discussion of the success or failure and provides further analysis where required. The evaluation includes an in depth analysis of the factors that prevent successful extraction and also recognizes those scenarios that allow automatic extraction of buildings from aerial images.

Chapter 5 concludes the document and provides recap of the entire process as well as final remarks from the author regarding the contribution of the project, the recognized limitations for the approach and possible future research.

Chapter 2   LITERATURE REVIEW

The work presented in this report is mainly concerned with the possible automation of the building extraction procedure.   The research and technological advances in photogrammetry, remote sensing and computer science introduced a remarkable potential for reducing human involvement in building urban inventory. The methodology tested in this work involves different image processing techniques at different stages and emphasizes the need for data fusion during the extraction procedure.

This section begins with a review of the history of photogrammetry and major mile stones that lead to the current digital era. Section 2.2 entails a general review of the building extraction procedure and approaches taken in research for this purpose. Literature for the building extraction procedures is described in section 2.3 as well as image based classification techniques. This section deals mainly with extraction from high resolution imagery.   Section 2.4 elaborates on image processing techniques used for the extraction of buildings from aerial imagery, including classification methods and shadow extraction. Section 2.5 emphasizes the need to incorporate existing GIS data within the extraction process. Different approaches that take advantage of existing spatial information are descried. Section 2.6 provides a short review of LIDAR and laser scan technology. Although not pursued within this project, the unique advantage of this technology for feature extraction is acknowledged.   As an important part of the methodology of this project, section 2.7 explicates the inherent value of subsetting an image into smaller patches prior to extracting the building.   Section 2.8 reviews the moment theory as a tool for shape identification as relates to the methodology.

2.1 Evolution of photogrammetry and Remote Sensing leading to feature extraction

Photogrammetry has changed dramatically along side the technological advancements as developed in the past century. The evolution period can be divided into several phases (Konecny, 1985; Madani, 2001). The first phase is referred to as the "Analog Photogrammetry". This phase began around 1900 (Konecny, 2003) and inaugurated the use of aerial imagery for mapping purposes. The mapping process was based on stereo-plotters and the "stereoscopic measurement principal" (Konecny, 2003). Stereo plotters reconstruct the relative location and orientation between images at the time they are captured. Due to the different perspective of the images, a 3D model is created for the overlap area between the images. The model was mainly used to capture elevation and contour lines. The next phase began in the 50's and is called the "Analytical photogrammetry". The analytical photogrammetry introduced the first aero-triangulation implementation, DEM generation and feature extraction as a result of the breakthrough of computer-aided techniques and applications in the 60's (Madani, 2001). The third phase is the computer-aided phase that started in the early 70's and introduced a new level of efficiency to the mapping process. During this phase we see the emerging computer and graphic processing abilities (such as CAD systems) as they become more and more dominant in the photogrammetry arena.

The last and current phase is the "Digital photogrammetry". The inherent difference between that phase and the previous phases lies within the nature of the imagery. The digital era deals with pixel image coordinates and grey levels, while previously the hard

copy image was the input media. The greater power of computers and workstations, satellite imagery, photogrammetric cameras (including CCD, line scanners), on board GPS systems, scanners, RADAR technology have made an impact. Currently even personal computers are able to perform many tasks that require massive processing power and space. Advances in computer science and implementation of photogrammetric principles and techniques such as AeroTriangulation, orthophoto and DTM generation, allowed the photogrammetry and remote sensing community to move in a new innovative research path towards automation of more complex procedures. Tasks such as feature extraction are still in research as they were in the last three decades. There is a fundamental agreement that photogrammetry and remote sensing can provide an efficient and relatively easy way to collect data and maintain updated GIS systems for purposes such as resource management. As imagery improves in spatial and spectral resolution and becomes more available, we are able to extract better, more accurate information. Still today, many processes rely on a human interpreter to distinguish between different features and digitize the accurate positions of objects.

2.2   Building Extraction – General

Many studies presenting automatic or semi-automatic approaches to building detection have been published. Building extraction poses several major difficulties that any system has to overcome (as discussed in section 1.1). Marr (1982) describes the human vision as an information-processing task. This task encompasses many aspects of a human perception, such as shape, space, spatial arrangement, illumination, shading and reflectance. Since we can not fully imitate the human brain functionality, it is a challenging task to automate image vision and interpretation. Hence, operators are still

indispensable during the feature extraction phase, and many applications use a semi-automatic approach to extract point, lines, areas and complex objects (Vosselman, 1998). This approach utilizes the advantages of the human as well as the superiority of the computer for specific, repetitive tasks. Line following mechanisms, based on an initial point on the road, have been prominent for road extraction (Aviad and Carnine, 1988; Gruen and Li, 1995). Similarly, the user may specify the approximate location of an object and the computer will then perform a specific task such as a seed-growth algorithm (user defined starting point and growth according to value similarity between pixels) to locate the boundaries. Alternatively, users may locate one corner and using edge detection to locate the rest, acquiring approximate location of the corners and having the computer snap to the nearest "point-of interest" (the corner). A third approach could employ manual digitizing in one image and using epi-polar geometry between the stereo pair to locate the height and the corresponding points in the second image (Tao and Chapman, 1997).

The extraction of areas (polygons) is based on homogenous surface attributes such as grey levels (similar color for water areas, roof tops etc.). One drawback of implementing that approach on an entire image is the wide variety of grey levels within the group of man-made buildings. 3D objects such as buildings are mostly referred to in the literature as complex objects. For the extraction of those objects, the user utilizes geometric constraints to extract linear parallel lines for building edges and match those edges to overlapping images. Man made feature extraction process can take advantage of supplementary information such as DTMs (Digital Terrain Model) for both building and road extraction. That information can be used to impose geometric known characteristics

such as moderate changes in height along a road or a river, or rectangularity of buildings. These constraints can make the extraction process much more robust and reliable.

Figure 2 portraits the main approaches, techniques and sources of data for building extraction, as described in the given research review.
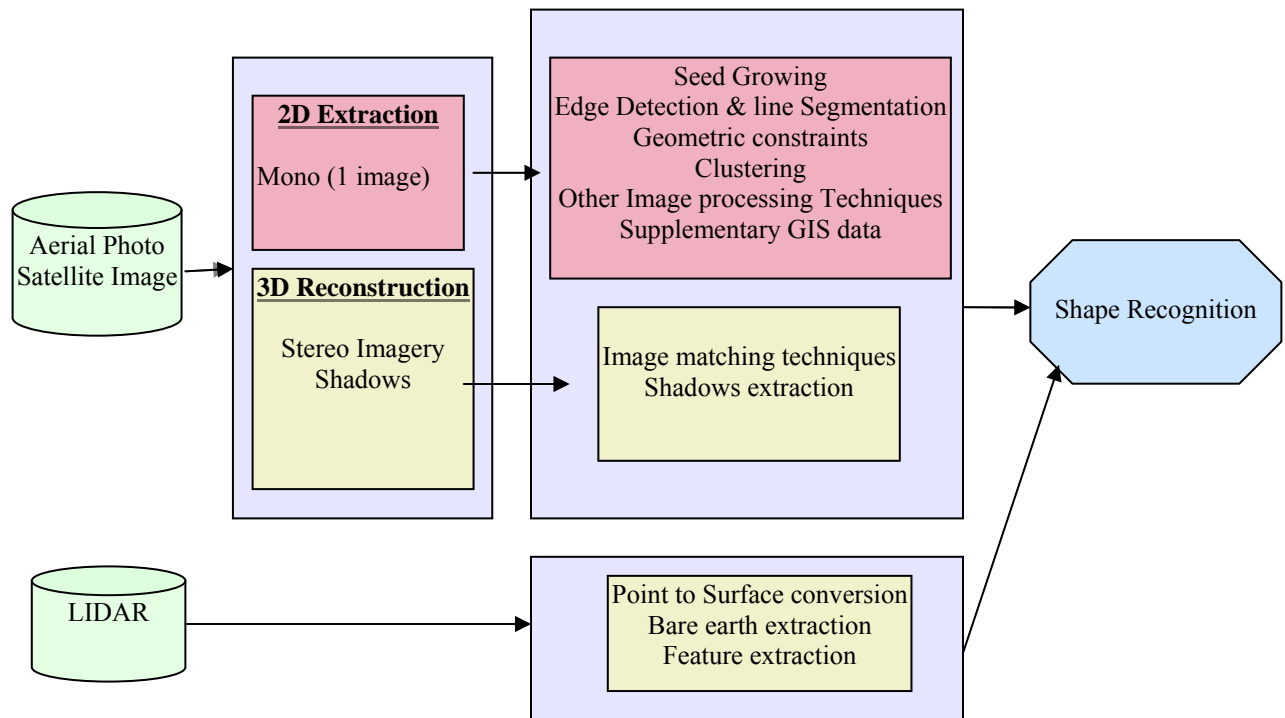
Figure 2 – Main Approaches to Building Extraction

Aerial Photographs and high resolution satellite imagery are the most common sources of data for feature extraction. In order to extract 2D characteristics of a feature, a mono image may suffice. 3D features such as buildings often require analysis of 3D cues (such as shadows) within the image in order to adequately detect and extract the outline. Any 3D reconstruction of the feature requires either stereo imagery (two or more images with enough overlap area) or sophisticated extraction and analysis of shadows cast by the features.

2.3   High Resolution Satellite Imagery

2.3.1   Data Classification Techniques from remotely sensed data

Image segmentation and classification is a common and prominent method for extracting information from images in many disciplines, using a wide variety of well established approaches as well as new, innovative techniques (Wang et al, 2001; Bin et al, 2002; Chen and Wang, 2004; Zhang et al, 2006; Zebedin et al, 2006). Segmentation is usually a first step in a process followed by subsequent analysis of the features and possible object matching algorithms. The computer science community has shown great interest in image segmentation, much of that for the purpose of image retrieval from a database (Ahu and Yuille, 1996; Shi and Malik, 2000). Smeulders et al (2000) discusses the difficulty of reaching a "strong" segmentation. He then continues to discuss "weak" segmentations that result in homogeneous regions within the image that do not necessarily cover entire objects. Compromising for a "weak segmentation" gives a rise to numerous problems in the following steps of the process and to the overall success of the image interpretation.

Classification of data from aerial and satellite imagery is a well known approach within the photogrammetry and remote sensing communities. Remote sensed data enabled a replacement of in-situ measurement for disciplines such as forest management. Information that is otherwise very hard to obtain, is available through images for planners, ecology modeling, and many other disciplines (Jensen, 2005). The remote sensing data collection records the amount of radiation reflected by the object, thus creating a "signature" of the object. The signature of the objects holds much information

about the characteristics of the objects. We are able to classify different types of objects and even sub-object using classification methods.

Remotely sensed data is usually classified using methods that can be categorized as: Supervised and Un-Supervised classification. A Classification is mainly an automatic or semi-automatic way to identify the signature of each class and their location in the image. The final goal of the classification is usually to yield land cover/ land use classes for the area of interest.

Supervised classification – supervised classification assumes prior knowledge through personal experience, interpretation of aerial images and map analysis (Jensen, 2005; Hodgson et al, 2003). The analyst manually trains the system by locating areas in the images that comprise of the classes of interest. The user should strive to define distinct classes with the least amount of overlap between them (in the spectral space) in order to allow better classification. To that goal, the user may take advantage of spectral plots that easily portray the degree of correlation between the classes.

For each class, the system calculates statistical measures (Standard Deviation, Mean, covariance matrices etc). During the classification process, every pixel is assigned a class according to the highest likelihood of being a member of the class. Once the classification is complete, a rigorous error evaluation takes place and statistics are available to the user.

There are different types of algorithms that can be used for the classification. Those can be divided into parametric and non-parametric classifiers. The parametric assumes normal (Gaussian) distribution for the observations (Schowengerdt, 1997). The most

common supervised classification method is the maximum likelihood algorithm (Ozesmi and Bauer, 2002; McIver and Friedl, 2002). This classifier calculates a probability density function based on statistical measures of each class. By placing the brightness value of the pixel in the probability function, we obtain the probability of the pixel being a member in a class. The pixel will be assigned the highest probability class.

The most common non-parametric classification algorithms and techniques are (Jensen, 2005):

Parallelepiped – Parallelepiped classification is a fairly simple to implement and efficient algorithm. For each band and class, the algorithm calculates the mean and standard deviation values. The result is an n-dimensional vector with all the mean values of the trained data for each class in each band. The boundaries for each parallelepiped are defied based on 1 standard deviation values. A pixel is evaluated according to the high and low standard deviation values (greater then the lower boundary, less then the high boundary) and if not suitable to any class, it will be assigned to an unclassified class. A problem may occur when parallelepipes overlap. In such cases, the pixel is usually assigned to the first found class or a criteria rule such as minimum distance can be used to make the decision.

Minimum distance(MD) – Similar to the Parallelepiped algorithm, the minimum distance algorithm first calculates the mean values for each class in each band. The result is a mean value vector for all the trained data. During the classification, the algorithm performs distance calculation between each pixel and the mean vectors. The minimum

distance determines the class assignment. The user is given the option of defining a maximum distance that beyond, the pixel will not be classified.

Nearest neighbor (NN) – The simplest non-parametric decision rule that weighs nearby evidence more heavily, thus classifies a pixel to the nearest class (Cover and Hart, 1967). The calculated distance between the pixel and every class is Euclidean distance. Other, less simple algorithms such as the k-nearest neighbor search for the closest k number of training pixels in the feature space to determine the class.

Artificial Neural Networks (ANN) – Neural networks have been increasingly applied within numerous applications over the past decade (StatSoft, 2003; Makhfi, 2007). The ability of ANNs to learn and reach a decision, like a human, has captured the interest of researchers from many fields. ANNs have been used to make prediction such as future stock performance (Makhfi, 2007), data modeling, function regression, pattern recognition (California Scientific, 2007) and more.

The concept of neural networks was first introduced by McCulloch and Pitts (1943), but required the advancement of computer technology to be successfully developed and applied. ANNs simulate decision making processes as achieved by inter-connecting neurons in the human brain (Jensen et al, 1999). The decision making process is based on initial training of the network. Input and desired output examples are provided to the network. During the learning process, the weights of the different connections are adjusted to achieve the specified outputs.

There are two main advantages that made ANNs appeal to the remote sensing community: ANNs do not require a normal distribution (hence, it is not necessarily a

18

parametric classifier) and they can simulate non-linear patterns (Jarvis and Stuart, 1996; Jensen et al, 1999; Jensen 2005). ANN has been implemented in remote sensing software such as ENVI, but there is no clear consensus about the superiority of ANNs over traditional classifiers. Several projects demonstrated better classification results for ANNs (Ji, 2000; Bischof et al, 1992; Jensen et al, 1999) while others show no significant advantage to ANNs or expressed more caution (Hepner at al, 1990; Jarvis and Stuart, 1996). Moreover, since there is no clear explanation to the rules as created by the neural network, it is being assessed as a "black box" (Qui and Jensen, 2004). Hence, users are reluctant to use those systems for real world scenarios. Another major disadvantage of ANNs is the training process that requires the users to be very knowledgeable about both neural networks and the area of interest for the classification. This is a major disadvantage due to the relative simplicity of running any other traditional classification.

Unsupervised classification - Unlike Supervised classification, Un-supervised classification does not require prior knowledge about the area of interest, thus, no training is required. The system searches for natural grouping/clusters of the pixels (Jensen, 2005). The most commonly used classification methods are the ISODATA and K-means.

ISODATA **-** ISODATA stands for Iterative Self-Organizing Data Analysis Techniques (Ball and Hall, 1967). No prior training is required, but the algorithm needs a starting point and thresholds (Fromm and Northouse, 1967) for split, merge and stop criteria. The algorithm assumes Gaussian distribution of the pixels in each class.
The criteria parameters include (Jensen, 2005) the maximum number of clusters; the maximum percentage of pixels allowed being unchanged between iterations - when the system reaches that number, the process stops; Maximum number of iterations - each

iteration involves recalculation of the class mean and re-reclassification of pixels; Minimum number of pixels in a class; maximum standard deviation for a class; minimum distance between cluster means; Split separation — if not 0, this value will be used to decide on the location of the new class when splitting large classes rather then using the standard deviation.

The algorithm begin by calculating initial mean vector for the classes, and iteratively move pixels between classes, merge classes or split classes based on the input parameters. The ISODATA algorithm is considered slow (Jensen, 2005).

K-means — The goal of the k-means algorithm is to divide pixels between clusters in a way that the sum of squares within each cluster is minimized (Hartigan and Wong, 1979). The mean position of all pixels within a class defined the center of the class. Pixels move between classes based on the Euclidean distance to the center of the class.

Sometimes, the boundaries of phenomena may not be distinct, but rather fuzzy. In the image space, a pixel may contain more then one land cover class ("mixed pixel"). In order to handle this case, a fuzzy classification algorithm may be used instead of the hard classification methods described above (Laha et al, 2006). A fuzzy algorithm is based on replacing the hard boundary between the classes with more gradual transition between the classes. Those methods assign to each value several probabilities according to the set of classes it might belong to (Jensen, 2005).

In the past decade we have witnessed the development of *object-oriented classification* methods. Unlike the per-pixel classification algorithms, Object Oriented classification techniques aim to extract homogeneous regions within the image that bear

meaningful information. The image is usually divided into sub-areas based on spectral and spatial characteristics and then each region is assigned to a class (Wang et al, 2004). The combination of spectral and spatial information is useful for land cover classification since, often, the same class encompasses several similar spectral signatures or different classes share spectral signatures.

The classification techniques illustrated above have been widely employed by the photogrammetry and remote sensing communities for land cover classification. The classification processes were based on one specific technique (Samaniego et al, 2008; Davis and Wang, 2002) or a fusion of several algorithms. Zebedin et al (2006) illustrate an approach to automatically generate land cover/land use maps from aerial imagery. The images include both high resolution series of panchromatic overlapped images and low resolution multispectral images. The methodology encompasses different classifications – maximum likelihood, neural network, decision tree and support vector machine. Substantial effort is devoted to image matching DTM and DSM (Digital Surface Model) generation and AT (Aerial Triangulation). Their result is a raster land cover classification map that showed a high accuracy of vegetation detection. Li, Wang and Ding (2006) propose a feature extraction method that can be used for urban area mapping based on a potential function clustering method. This clustering method segments the image by selecting peaks within the image histogram. The claim is made that, within an urban image, a grey level histogram peak can be used to segment the entire image. Once the segmentation is complete, the buildings can be selected manually. Every candidate is extracted using seed region growing followed by edge detection, dominant line detection and outline mapping. The method was tested on a region within

a Quickbird image (0.61m) that, based on the example provided, includes mostly regular buildings. The results shown were visual and the accuracy was depended on the selected grid size for the building outline mapping.

Segmentation is an early step within the building extraction process in this project. By analyzing the histogram of a localized image, regions in the image are segmented and analyzed to identify buildings. This approach can be more closely related to the object-oriented segmentation approaches as discussed above.

2.3.2   Building Extraction from Satellite Imagery

Fraser et al (2002) investigated to ability to extract buildings manually from IKONOS imagery in order to construct 3D models. One of their conclusions was that about 15% of the buildings could not be identified in the imagery. They reported possible sub-meter accuracy for stereo input images under certain conditions and data configuration. Xiao, Lim, Tan and Tay (2004) use high-resolution IKONOS satellite stereo-pairs to extract roads and 3D models of buildings. The building extraction relies on existing roads and previously extracted vegetated areas. The building extraction process is semi-automatic, based on edge detection and "thinning" and allows the user to select between several potential rooftop alternatives and adjusts corners and edges. For small buildings, the user may predefine rooftops to be recognized using Neural Networks. Heights are eventually computed using the stereo-pair images. Sohn, Park, Kim and Heo (2005) propose a building extraction method based on high resolution IKONOS multispectral stereo pair images. The algorithm is based on an image processing

technique BDT (Background Discriminant Transformation) on multi-spectral images. This technique is scale invariant, reduces the variability in the background, and enhances the non-background. Similar to the principal component, several bands are created with maximum to minimum differences between the background and the non-background (here, the potential feature). The result of the first step allows the classification and clustering of buildings. Once the buildings are enhanced using the previous step, they are clustered by the ISODATA algorithm (See section 2.3.1). In the next step, using color indexing and distance measurements, matched buildings between different images are located. Matching the buildings between the stereo-pair enables the generation of a 3D model. This article highlights the growing need for extracting 3D characteristics of objects. This is definitely an open problem that needs to be tackled, although the results reported emphasize the need for future research that could utilize information such as shadows, since matches failed mostly owing to buildings obscured by shadows. Wei, Zhao and Song (2004) use image processing techniques in order to extract buildings from high-resolution satellite imagery (using Quick Bird panchromatic images). Their application is based on unsupervised clustering using histogram analysis and shadows in order to detect and locate buildings. Edge detection and subsequent Hough transformations are used to extract the dominant lines of the buildings and construct the building footprint.

Table 1 - Sample of available aerial and satellite imagery

Sensor Type	Spatial Resolution	Radiometric Resolution	Temporal Resolution
*Panchromatic	0.5 ft	2^{12} (0 - 4096)	Varies
*Color film	0.5 ft	2^{12} (0 - 4096)	Varies
*Color IR film	0.5 ft	2^{12} (0 - 4096)	Varies
*Panchromatic	1 ft	2^{12} (0 - 4096)	Varies
*Color film	1 ft	2^{12} (0 - 4096)	Varies
*Color IR film	1 ft	2^{12} (0 - 4096)	Varies
QuickBird-2 Pan	0.6 m	2^{11} (0 - 2048)	3 days
QuickBird-2 MSS	2.4 m	2^{11} (0 - 2048)	3 days
IKONOS-2 Pan	1 m	2^{11} (0 - 2048)	3 days
IKONOS-2	2.4 m	2^{11} (0 - 2048)	3 days
IKONOS-2 MSS	4 m	2^{11} (0 - 2048)	3 days
SPOT-5 Pan	2.5 m	2^{8} (0 - 255)	3 days
SPOT-4 MSS	20 m	2^{8} (0 - 255)	3 days
SPOT-1,2,3 Pan	10 m	2^{8} (0 - 255)	3 days
SPOT-1,2,3 MSS	20 m	2^{8} (0 - 255)	3 days
Landsat TM 7 Pan	15 m	2^{8} (0 - 255)	16 days
Landsat TM 7	30 m	2^{8} (0 - 255)	16 days

* - spatial resolution depends on sensor altitude

For testing the methodology presented in this research, we will use 1ft resolution color aerial imagery. That detailed resolution allows an accurate detection of the building outline. At the same time, this resolution presents problems such as the inability to use simple clustering functions (such as ISODATA) that are common for coarser resolutions (See SPOT and Landsat in table 1).

2.4   Image processing techniques

According to pre-defined models of buildings, Huertas and Nevatia (1988) impose geometric constraints during the detection stage, and shadows are used to verify the outline of the feature as well as to estimate the height.   Shadows have been used in different stages of extraction by Irvin and McKeown (1989), as part of shape prediction, grouping, verification and height estimation.   Constraints are the core of the building recognition proposed by Kolhe, Plumer and Cremers (2000).   Their approach is based on constraints and logic programming, and employs hierarchical building models with 2D and 3D representations.   Sahar and Krupnik (1999) used methods that combine edge data, stereo-analysis and shadows in order to extract the 3D shape of the building.   3D information was found to be a significant source of information for the building detection process.   Sung Chun, et al (2006) developed a system aimed at 3D extraction of the building outline.   The user selects a point within the building on one image and the system locates the matching feature on other images and provides the best 3D hypothesis for the building.   The experiment results show an efficient semiautomatic approach for 3D complex buildings.   Weidner and Förstner (1995) generate a high-resolution DTM, and with topographic data available for the scene, they extract and reconstruct the buildings.   Shi and Shibasaki (1996) use stereo imagery and line-based matching to overcome the absence of shadow information in an image in order to estimate ground elevation and evaluate the 3D lines of the building.   Avrahami et al (2008) describe a method to extract rooftops based on the assumption that rooftops are combination of polygons. The approach consists of both manual and automatic steps, where the user has to select a parameterized model for the extraction and point to the location within the left

25

image. Based on that input and the right image, the roof is constructed. Multiple building extraction procedures were developed to identify buildings from a single image (mono). Kim et al (2004) developed a methodology to extract large rectangular buildings from 1m resolution imagery using a single image and line detection techniques. The user's interaction includes clicking on a point inside the building, and initiates a seed location for the process to identify the boundaries. Their results conclude that short sides of the buildings are not well extracted and there were several cases of false line extraction. Although some buildings were not oriented correctly, they believe the overall approach of pointing the building as a cue is a promising approach. Müller and Zaum (2005) present an approach based only on an aerial image and image processing techniques. Their approach initially uses a seeded region growing mechanism to segment the image. Segments are then evaluated using geometric (area, roundeness etc) photomentic (Hue angle, mean hue angle etc) and structural (Shadows, relationship to other extracted buildings etc) parameters. Jin and Davis (2004) integrate different algorithms based on spectral, structural and contextual (position and size of adjacent buildings) information in the image. Their approach is aimed at 1m IKONOS imagery and they develop a DMP (Differential Morphological Profile) to generate the structural information as well as extract shadow regions. The evaluation report shows that 72.7% of the building areas are extracted with a quality percentage 58.8%. Over 70% is considered a good success rate, while the 58% quality shows need for improvement. Some gabled roofs and dark parking lots were not detected or falsely classified as buildings. Lee et al (2003) used supervised and unsupervised classification techniques on both the multispectral and the panchromatic images of IKONOS satellite. Their building

squaring techniques were based on Hough transformation and their overall extraction rate is 64.4%. Most misclassification occurred owing to shadow inclusions and false road classification. Tang et al (2004) tackled the problem of extracting high rise buildings. High-rise buildings usually have complex shapes and have not received enough attention within the feature extraction research. The extraction procedure was based on identification of vertical lines in a mono image using fuzzy Hough–transform, complemented by photogrammetric principles that provide clues about the location of vertical lines in the image. Locally, the authors analyze the texture of the windows on the wall of the building. Roof extraction was accomplished using edge detection and line segmentation. The approach was tested on a 0.2 meter scene from an aerial image and reconstruction results were presented for four high rise buildings in the image.

The methodology of this research is geared towards minimizing user's intervention; therefore, techniques such as selection of a starting point for seed growing, is not used. Shadows are extracted as means of verifying the location of the building and eliminating non-building (2D) features. The process also evaluates the geometry of the feature as well as the validity of the building candidate based on geometric characteristics such as minimum size and width.

2.5   Supplementing with existing spatial data

Existing GIS data provides prior knowledge about the area and can add considerable value to the extraction procedure. Baltsavias (2004) provides an overview of the status of feature extraction in research. The author acknowledges the need for more projects that exploit a-priori information. He points out to the fact that only few use a-priori knowledge in the form of GIS data, maps and geodatabases. Duan, Prinet and Lu (2004)

used GIS data about the position and shape of buildings. Their application uses a fuzzy segmentation technique for roof extraction without allocating the actual footprint of the building. Segmentation of the building area is accomplished using seed growing mechanism. The seed point is selected by calculating centroid points for buildings in a GIS building layer. The building vector layer is converted to a raster image. The segmented regions and the original building raster image are compared through out the process to minimize differences in area between the polygons. Koc and Turker (2005) developed similar applications to update an existing vector building layer. They utilized image Supervised classification, NDVI (Normalized Difference Vegetation Index – indicator of vegetation presence in remote sensing images), DSM (Digital Surface Model) extracted from a stereo pair and object extraction techniques. Khoshelham (2004), demonstrate the fusion of images data, height and plan 2D data (XY coordinates of the roof) to reconstruct the 3D model of the building. The images used in that project include the near IR channel and hence allow simple identification of vegetation for the first step of vegetated-area elimination.

Available GIS data such as tax assessor attribute data is used in this research as a cue for existing buildings, as well as a way to eliminate different areas within the image. For example, vacant parcels are not searched. In vegetated area, NDVI can be used as a first step of pre-processing before running the proposed application.

2.6   LIDAR and Laser Scan based methods

In recent years, LIDAR points have become a valuable source of data for feature extraction. The density of the points allow the generation of detailed DTMs as well as extraction of buildings, roads and other features (Gamba and Houshmand 2000,

Rottensteiner, Trinder and Code 2005, Nardinocchi, Scaioni and Forlani 2001).

Rottensteiner, Trinder and Code (2005), describe their effort in automatic extraction of buildings and roads using LIDAR data. Current LIDAR resolution, while high (about 2m), actually causes ambiguity between buildings and vegetation features such as trees for the extraction process. Overcoming that ambiguity requires the use of other sources of information such as intensity and/or RGB orthophoto that allows the classification and subsequent elimination of vegetation using NDVI or pseudo-NDVI. In the project described in the article, a 2 phase algorithm was developed in which the buildings are first detected and then reconstructed. The detection is based on a pixel-based classification. The authors defined 4 basic classes - buildings, trees, grass and bare soil. Each pixel receives a probability score of being in each class (sum of all probabilities for each pixel is 1). Other measures are used to eliminate different classes. For instance, differences in height between the DSM and DTM can distinguish between building/tree object groups and flat grass/bare earth object groups. Vegetation indices can distinguish between tree/grass object groups and building/soil groups. By imposing these tests, a binary image is created where buildings are high-lighted. The result is refined by eliminating areas that are suspected as trees. The method failed mostly in residential small buildings and using NDVI seemed to increase the accuracy and extraction results. (95% of all buildings greater than 40 sq-m; buildings smaller a 30 sq-m were not detected at all). Finally, the reconstruction of the building and extraction of the boundary is created with the aid of GIS external data (such as streets to separate buildings) and geometric characteristics for buildings. Sohn and Dowman (2003) presented their approach to building extraction that includes a recursive extraction procedure to evaluate

the existence of buildings within both the LIDAR cloud and IKONOS image. The building extraction was achieved using hierarchical segmentation of LIDAR DTM complemented by color imagery.

Building detection using LIDAR point data is currently beyond the scope of the research, mainly since LIDAR data are not widely available for every urban area. Future research can investigate the integration of the proposed approach with available LIDAR data.

2.7 Image subsetting approaches

Image partitioning has been tackled within the computer science community mainly for the purpose of content-based image retrieval and matching. In order to effectively retrieve information automatically from images, global image analysis is rarely sufficient. Numerous algorithms have been developed to partition the image into regions that allow easier interpretation of the objects (Berretti et al, 2000; Carson et al, 2002; Yixin and Wang, 2002; Jia and Wang, 2003; Chen and Wang, 2004). The regions are created either by subdividing the image into pre-defined region sizes or number of regions, or locating homogenous regions in the image. The inherent problem of pre-defined subdivision of the image is the cutting of objects between regions.

Jiang, et al (2008) provide a semiautomatic methodology that first partition the image into homogeneous regions via image segmentation. The segmented regions are then processed to improve the segmentation result. The building extraction is performed within the homogeneous regions through an interactive selection of points for seed growing and then a region merge to union the buildings and remove overlaps. Finally, edge detection is performed. The method was tested on RGB image with 20cm resolution on an area that spans 867×617 pixels that consists mostly of regular buildings. The result was compared with ISODATA classification result. Zhengjun, et al (2008) present a region based and feature based methodologies for building extraction. The region based method relies on localized region extraction. In order to extract homogeneous building regions, the user manually identifies an area that contains buildings and background. This area is used to launch a mean shift segmentation method

31

that results in the areas of attraction. A seed growing technique based on manually selecting the area of the building is applied on the segmented result. Boundary fitting and shape recognition follow that step. The feature extraction methodology is based on edge detection (Canny), line and corners detection followed by locating the boundary of the building and shape reconstruction. Constraints on the length and the right angle search process for the corners are imposed. The methodology was tested on QuickBird imagery and shows success rate of 75% (total of 35 buildings were tested) for extracting major rectangular and regular building rooftops.

Sohn and Dowman (2003) simplified the building extraction procedure in an urban scene by localizing the search to the building area. Their approach was based on LIDAR points and a technique that automatically distinguished between on and off terrain points. In order to differentiate between those types of points, the technique located homogenous slope sub regions in the LIDAR point cloud. This localization approach failed for most residential buildings. The low density of LIDAR points and surrounding trees seemed to be the leading reasons to the failure. As regions as sub areas in the image are discussed within the computer science discipline, parcels as sub areas are commonly discussed within the GIS community. The next several articles have a title that incorporates the term "parcel" for feature extraction. Wijnant and Steenberghen (2004) used parcels as a measure to summarize land cover classes for land use rather then to reduce signature confusion. Their methodology includes initial per-pixel classification of the 1m image to acquire land cover classes and then summation of the classes per parcel to distinguish different land use parcels by the type of classes they contain. Ming et al (2005), use the term parcel as related to a region or a field that is acquired by initially breaking the image

into homogeneous segments based on computational rules. A significant aspect of their project is developing a mechanism to partition the image based on spectral, texture and relationships between objects at different scales. Partitioning of the image occurs from rough classification on a large scale to more focused processing on the more elaborated and smaller regions. One of the major conclusions of the paper is the necessity to break the image into smaller areas. Ohlhof, et al (2004), also refer to parcels as regions that have statistical parameters that differ from their neighbors, are homogeneous and have distinct boundaries. Their projects involved development of an algorithm to extract those regions from the image as well as region growing techniques based on geometric and radiometric characteristics of the features.

An easy way to partition the image prior to the building extraction process can tremendously simplify the entire procedure. In this research I suggest the use of a graphic parcel layer to restrict the search to the close proximity of the building. Each parcel is cut through the image to create a smaller search region that contains one or more buildings.

2.8   Shape identification techniques and measures

The different candidates for the features are segmented in the image space. A significant problem is first to be able to describe automatically the properties of the segments. Delineating segment properties allow the evaluation of the probability of a segment to be part of a valid feature. Prokop and Reeves (1992) provide a significant survey of techniques based on moments for the recognition and representation of objects. The survey includes different types of moment invariants, the image ellipse and a variety of applications that take advantage of the "moment theorem". Hu (1962) introduced the

use of image moments for pattern recognition. In this ground breaking paper, Hu showed the ability to determine the shape of a pattern using moment invariants. Hu introduced seven different moments that define the area, COM (Center Of Mass), orientation, elongation in both axes (the standard deviation ellipse). Higher order moments can be defined to be invariant to scale, skewness and rotation. Hu showed that using moments can distinguish between letters of the alphabet. Using the first two invariant moments, Hu defined X and Y coordinates and plotted them on a simple 2D Cartesian coordinate system. The result plot demonstrated the distinct separation of alphabet characters over a graph.   A new letter can be identified by the minimum distance to an existing point on the graph.   Alt (1962) used moments to show that a limited number of moments are sufficient to describe and distinguish between patterns such as letters and numbers. Reiss (1991) claims to discover mistakes in Hu's theorem and revised Hu's moment to be invariant under general linear transformations as well as changes in illumination. Reiss also laid a foundation to evaluation of specific shape quality by calculating invariants specific to square and equilateral triangle shapes.   Schweitzer and Straach (1998) expanded the moment invariant theorem to describe automatically the properties of specific shapes based on the identification of the invariant properties of each shape. Their prototype shapes included square, rectangle, right triangle and triangle.   Rosin (1999) developed three rectangularity measures, based on MBR (Minimum Bounding Rectangle) and image moments. Rosin showed how using simple algebraic calculation of moments can identify the rectangularity and orientation of the shape in the image. Rosin (2003) expanded his methodology to evaluate ellipticity, rectangularity, and triangularity. Rosin includes measure of circularity, compactness, convexity and eccentricity and

shows limited results for automatic identification of all shapes. The best achieved results were shown for a combination of measuring ellipticity using DFT (Discrete Fourier Transform) and triangularity via moment invariants. Elad et al (2004) propose the use of complex moments for reconstruction of planar polygon.   This "shape-from-moment" approach extended previous work by calculating moments that are contaminated by noise (which is the expected scenario when working with real world data).   Guienko and Doytsher (2003) use moments to evaluate the geometric parameters of features such as buildings and parcels and after training the system, classification rules are used to extract different features within the urban environment.

The "moment theorem" is used in the proposed methodology to evaluate geometric characteristics of each building candidate.   Once an area is segmented as a possible building, the raster image is used to calculate the different moments of the segment and as a result evaluate characteristics such as the rectangularity of the feature.

Chapter 3   METHODOLOGY

The different research projects described above demonstrate several aspects of the extraction process that should be taken into consideration: Buildings are complex features and no single geometry can describe them; Edge detection and line segmentation cannot be used as a stand alone solution for the extraction process; The image should be partitioned in a way that will simplify the extraction procedure; Ancillary information, such as GIS data, should be used to complement the solution; Extraction of 3D features should integrate depth cues such as shadows or stereo imagery; Defining segment properties and shape identification measures are essential for the elimination and validation of the extracted segments and hence should be incorporated in the process.

Based on the above conclusions and initial experiment results (Appendix I), a methodology based on integrating remote sensing sources and available GIS data is proposed. The remote sensing data sets include high-resolution (1ft) ortho-rectified aerial imagery. Available GIS data sets are used in different stages of the analysis, for initial simplification and filtering and later as elimination, validation and supporting a-priori information for the process. Since buildings may appear in an image (also between images and different resolutions) in a wide variety of shapes and spectral signatures, the proposed procedure refrains from supervised training of the system. Moreover, even though we evaluate the rectangularity of the segments, a rectangle shape is not enforced on the extracted segment.

The suggested integration of several GIS sources, their attributes, topological relationships between raster extracted segments and existing vector data, as well as simplification first step based on tax assessor information, form a new, innovative approach for automatic building extraction. Fusion of available data sources with spatial analysis and image processing techniques will provide a robust approach towards solving the problem. The research presents an approach that includes several new approaches to alleviate the complexity of the problem, such as, cutting to parcel size to simplify the image, adding assessor attribute data and including user-one-time-intervention for shadow direction definition.

All the stages, the suggested input information and the techniques are elaborated later in this section with initial experiments results. The steps are depicted in figure 3 below.

(1)	**Parcel-based image partitioning. Eliminate vacant parcels. Identify parcel type.**
(2)	**Histogram Analysis**
(3)	**Feature segmentation in the image space.**
(4)	**Parcel-attribute based elimination.**
(5)	**Use of shadows for building segment verification and elimination**
(6)	**Geometry-based Elimination of low-probability building segments**
(7)	**Finding the footprint: Raster To Vector**
(8)	**Footprint Generalization**

Figure 3-Proposed Feature Extraction Methodology

The extraction was based on prior GIS data that restrict the searching area within the scene, and image processing techniques for the extraction. The following sections describe the various steps depicted in the methodology flow of figure 3.

38

Parcel-based image partitioning - This first stage can be explained as a simple "cookie-cutting" of the original image into parcel-sized images (referred to as "patches"). Tax assessor data sets that include parcel information are easily available for most areas within the USA and many countries around the world. Moreover, each parcel includes attributes that classify the use of the building within the parcel, whether industrial, residential etc, and specify ancillary information such as built square-footage for taxation purposes. The parcel attribute information can be used to:

- Eliminate vacant parcel (Where improvement value is 0)
- Distinguish between types of parcels, hence types of buildings. The extraction methodology assigns different characteristics to Single Family residential houses and commercial or other types of buildings. Landuse information about the parcel can be used to retrieve the type of building and assist during the extraction. Buildings within the same type are not necessarily within a continuous geographic area. The process, regardless of the location, retrieves the parcel based on the defined type.
- Calculate expected building footprint size – using the square footage and number of stories of the built features (if available).

Zoning ordinance information about each parcel may be used to isolate the area of the parcel, most likely to contain the building. Zoning information provides setbacks for the parcels which define the area allowable for building. Final results of the process may be further evaluated using the zoning information to estimate the added value of zoning for future use within the building extraction process.

39

3.1   Histogram Analysis

The histogram analysis is based on locating peaks within the histogram. Peaks in the histogram represent dominant features within the parcel. Since the parcel is a finite 2D space with limited number of features, the building spans over a significant number of pixels – essentially a peak within all 3 bands (see figure 4).   Investigation of some preliminary images shows consistent peak generation histogram for the building roof. Within a large sample of buildings in the test images, it was noted that about 50% of the parcels contain a building that creates the majority peak (a highest peak in the histogram). In other cases (see figure 5), the building generates a peak, but not the majority one, and in some cases (figure 6), there is more than one building in the parcel. Within the scope of the project, it will be evaluated whether more then 2 peaks need to be identified in the analytical process.
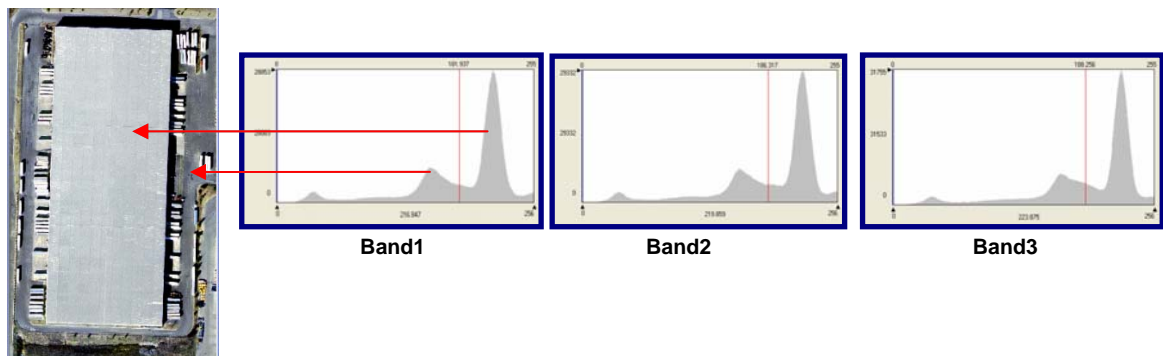
Figure 4 - Bands 1/2/3 for the image on the left. The high sine wave represents the building. Values span (dark)0-255 (light).
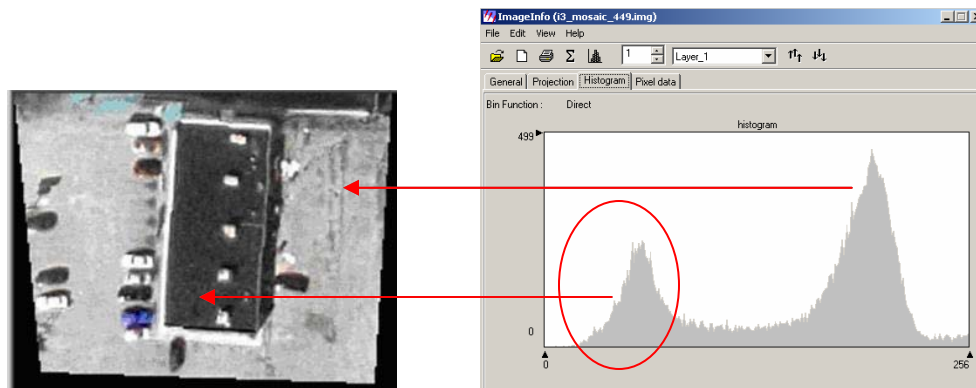
Figure 5 - A parcel with a building peak that is not the majority. Zero values are ignored.
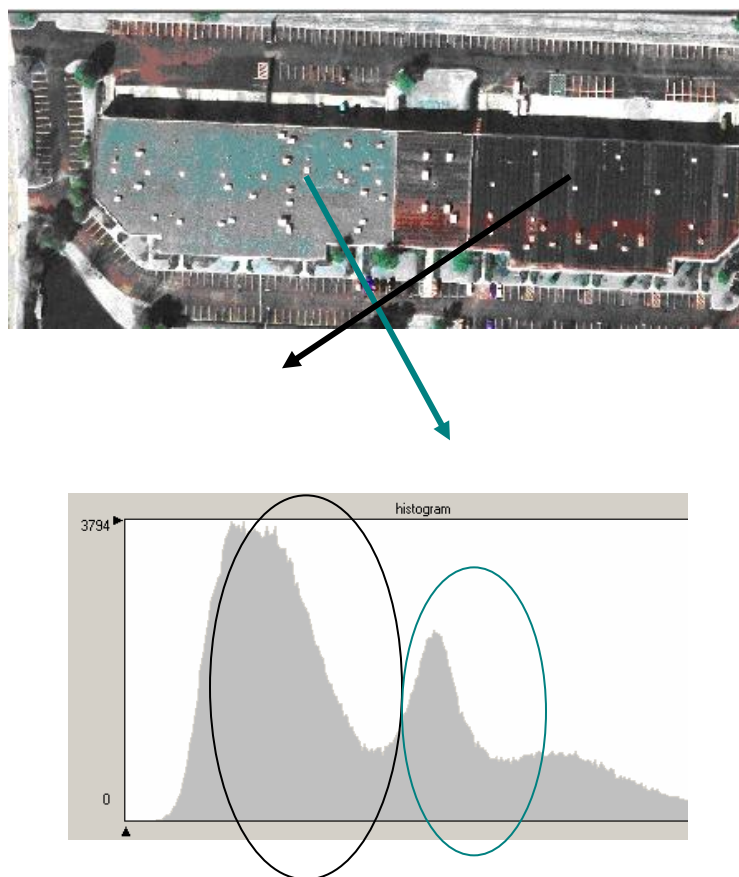
Figure 6 - A parcel with two buildings with different roof signatures. Note that the higher

peak also includes the parking lot area within the parcel.

3.2   Feature Segmentation

According to the located peaks within the histogram, features corresponding to the peaks are segmented in the image space. The segments are then processed further for smoothing and eliminating holes if possible. Currently, results evaluate the usage of only the first and possible second highest peaks for the parcel. A result of the initial segmentation is shown in figure 7.
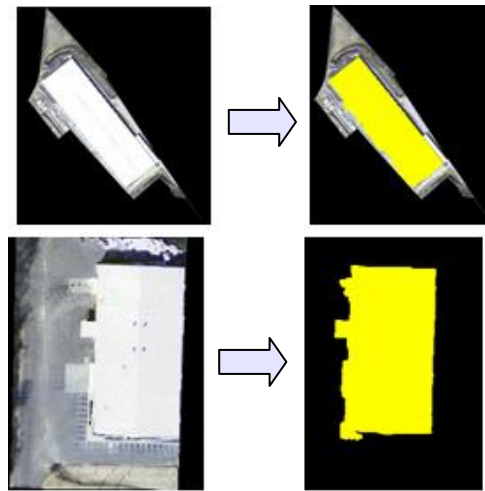
Figure 7 - Preliminary results - Original building on the left; segmented feature of only the majority peak on the right

When different peaks are classified into the image patches, several features are segmented, as shown in figure 8, and the system has to evaluate automatically which segment/s is part of the building.
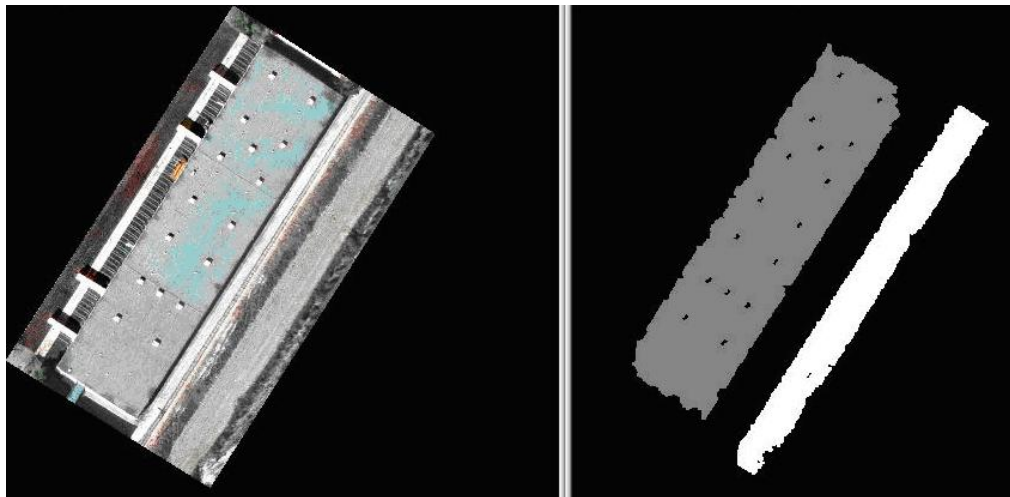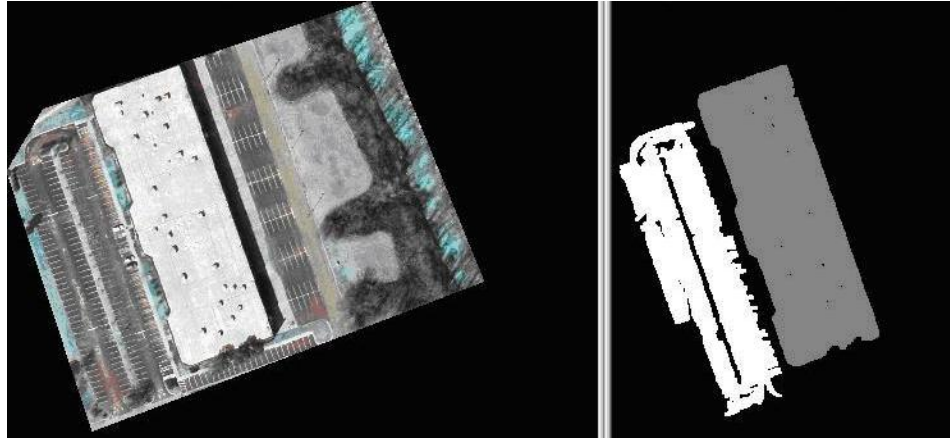
Figure 8 - Segmentation result of 2 peaks within histogram.

Stages 4-6 are used as a mechanism to evaluate whether a segment is part of a building. Segments that are definitely not buildings, based on the different measures, will be eliminated. For the rest, a probability score is attached to each segment at the end as an evaluation measure for the user.

3.3 Parcel-attribute based elimination

Size limitation may be used to eliminate segments that do not comply with the parcel built features. The size limitation is either evaluated using the SQUARE_FT and NUM_STORIES retrieved from the attribute table, or calculated as a minimum percentage of the parcel size.

3.4 Shadows

Different research projects concluded the importance of using and extracting depth cues for detection of 3D features such as buildings. Shadows have been proven to be a way to identify the existence of buildings and to distinguish between buildings from flat, ground surface features. For the purpose of this research the user is required to provide simple indication of the sun illumination. A brief look at the image, by any parametric user, can derive that information. A GUI with directions and attached radio buttons can allow the user to define sun illumination aspect as a parameter for the application:
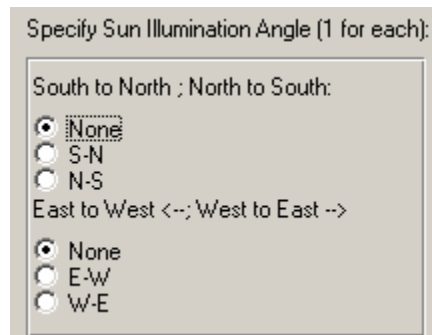
Figure 9 – User GUI for inserting Sun-illumination direction

For Example, in the image shown in figure 10, the user will indicate the 6st (West to East) and 2nd (South and North) radio buttons.

Figure 10 - Sun illumination orientation S->N and W-E

Knowing the general orientation of the sun can help distinguish between segments that have the shadow on the correct position related to the segment and are buildings, to non-buildings as can be easily detected in figure 11 below.
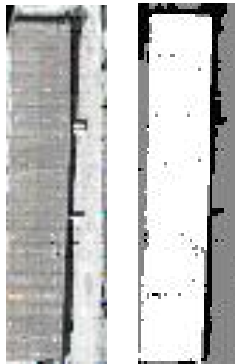
Figure 11 - White and Grey segments (right image) share a shadow.   The known orientation of the shadow can easily eliminate the grey segment from being a candidate building.

3.5 **Geometry based elimination of low-probability building segments**

In this step, the different Hu moments (Hu, 1962) will be evaluated and used for elimination of shapes as well as other recognized geometric characteristics. Currently in the literature, measures of circularity (not relevant for that procedure) rectangularity and ellipticity (Rosin, 2003) have been developed and shown to be successful for different applications. In order to evaluate the probability of the geometry to be a building, we will calculate the following:

- Rectangularity and ellipticity of the segment.

- Solidity – Solidity is defined as the (BoundingPolygonArea)/ Area. That index measures the amount of holes within the segment

- Convexity – Convexity is defined as (ConvexPolygonPerimeter)/Perimeter. That index should give 1 for a complete convex shape such as a rectangle or a square.

- Compactness – Area/(AreaOfShapeWithSamePerimeter). In the literature, the compactness is usually defined for the image ellipse (a circle is given a perfect 1 as the most compact shape). For this research we propose using a square as the most compact shape, and will need to define the compactness index accordingly ($16*Area/Perimeter^2$).

Possible utilization of islands within the segment will be evaluated as an indication for elimination of the segment (holes that are bigger then expected will be an indicator for non-building). Simple constraints such as a minimum width of a building will be used (for a space to be livable or used as a working area, minimum width is required).

46

According to the calculated indices (each index will be given a percentage, all summed to 1), a probability (confidence measure) is attached to each segment.

I will attempt to use the measures defined in the literature as well as the moment theorem to define a specific index for one of the known shapes of the buildings ("O" and "I"), from the common L, T, C, I, H, O building footprints.

3.6 Locating the footprint of the building

This step encompasses the conversion of raster segments into vector polygons as depicted in the figure below.
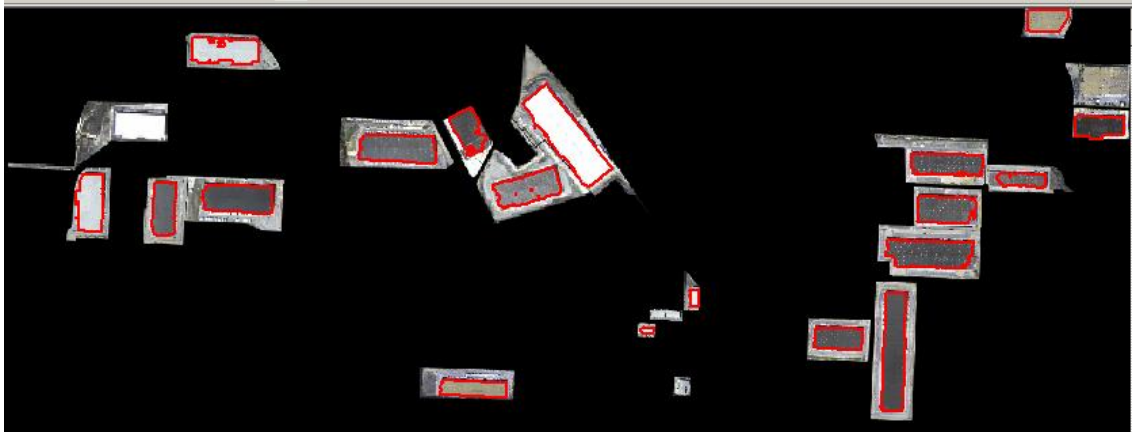
Figure 12 - A polygon shape file created for several parcels

The result of the entire process is a vector file that contains the polygon geometry as well as the probability of the polygon to be a building.

3.7   Generalization

Each polygon needs to be "cleaned" by generalization and orthogonalization of the corners.   Since generalization techniques have been implemented in various GIS applications, this step will not be implemented, and achieved using external software.

For practical use within a GIS production environment, the process has to be automated as much as possible. Once all the parcel-sized images are created (a separate process already developed), a batch process for each image that includes several models and c/c++ programs is invoked. Most implementation is transparent to the user who only has to enter the image/parcel layer and sun illumination direction. The implementation is currently done within the ERDAS-IMAGINE environment.

Chapter 4   IMPLEMENTATION AND EVALUATION

Following the methodology section, this chapter provides implementation details as well as result discussion and evaluation.   The section begins with elaborated implementation details and examples correlated with the process flow depicted in Figure 3.   The section entails the different assumptions, models, parameters and the implementation venue. The shape recognition procedure based on raster image moments concludes the section.   The implementation section is followed by results evaluation. The Memphis test area, aerial photography and GIS parcel datasets as used in this project, are presented.   Results for the commercial, residential and high-rise building are presented and discussed.   The feasibility of the methodology for each land use type is evaluated as well as the reasons for failure in extraction and potential drawbacks.

4.1   Methodology Implementation Documentation

4.1.1   Image Subsetting

The image subsetting procedure is a simple "cookie cutting" procedure that requires an image and a polygon layer as input. In this project, the input polygon layer is a parcels layer. For each parcel within the image extent, a new, smaller, parcel-sized image is created that follows the parcel polygon boundary.   Since the result image has a rectangular shape, area outside the parcel receives the value 0 (black background). Example of the image subsetting is shown in Figure 13.

(a)　　　　　　　　　　　　　(b)
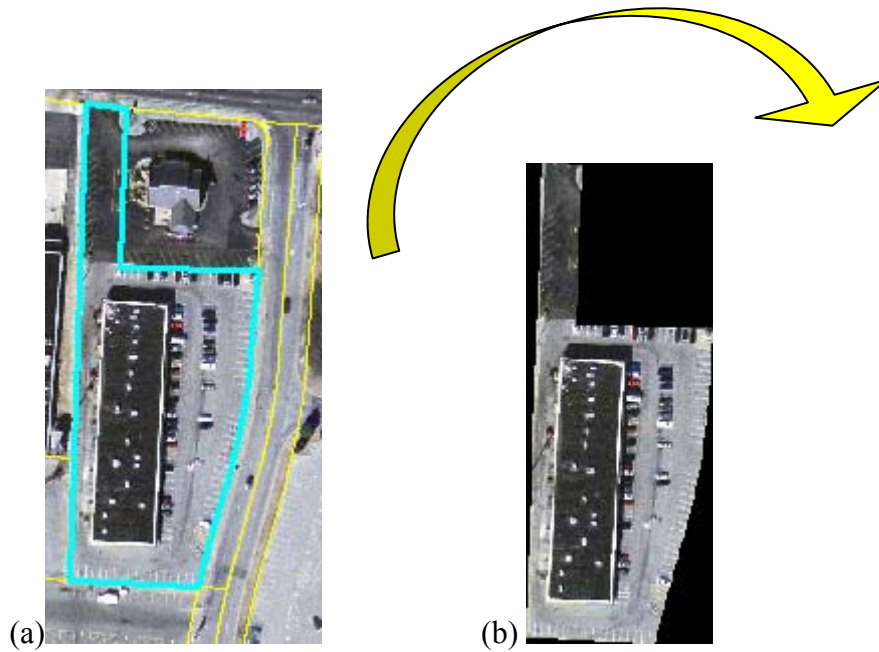
Figure 13 – Image Subsetting process.   (a) Original image overlayed with parcels layer (yellow line). Highlighted parcel is sub-setted. (b) Subset result image. Background pixels in black

Specific implementation details for the subset procedure are provided in Appendix A.

4.1.2   Histogram Analysis and Image Segmentation

The initial subsetting of the image into small, parcel-sized images is followed by a series of image processing techniques to extract the buildings. The subset procedure localizes the search area to a finite section of the image. Hence, it allows us to assert a basic assumption: the building area should be dominant enough and can be identified using simple image processing techniques.

4.1.2.1 Histogram Analysis

The histogram analysis is based on locating peaks within the histogram, which in the small area of the parcel, correspond to objects.   For a multi-spectal image that consists of

more then one band, the histogram analysis should incorporate several bands. In this project, the three visible bands (RGB - Red/Green/Blue) were analyzed. Mostly due to the small geographic area of the parcel, high correlation was identified between the histogram of the three bands (Figure 14).
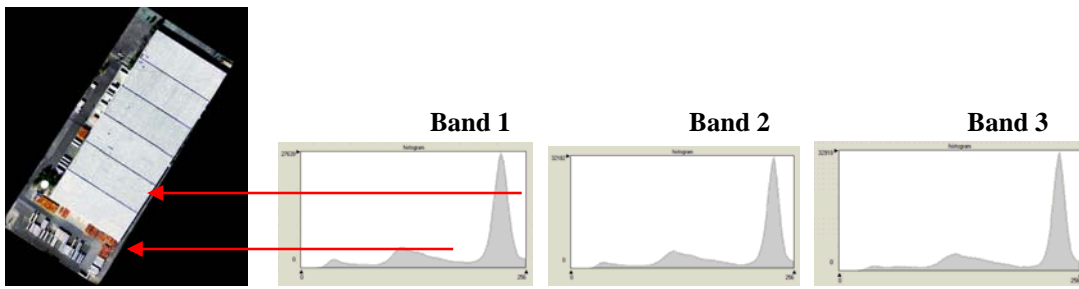
Figure 14 - Bands 1/2/3 for the image on the left. The high sine wave represents the building. Values span (dark) 0-255 (light).

In order to identify the threshold of each peak, the process locates the high point for the peak and then searches for the low "Saddle Point" (Figure15), where the slope becomes more moderate or completely shifts direction (for example in a V situation).
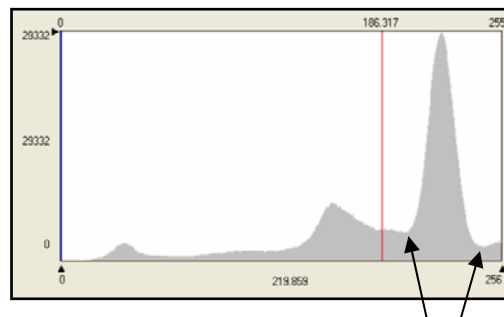
The "Saddle" Points of the peak

Figure 15 – Identifying the "Saddle" Points for each peak

Based on the "Saddle" points, a threshold is defined for each peak in each band. Then, all three bands are correlated to identify the object in all three bands.

51

Three different correlation scenarios were identified within the extraction process:

All three bands histograms are highly correlated (Figure 14)

Two bands are highly correlated and have a "Saddle" like geometry (indicating a bi-modal peak, two classes mixed together), while the third band shows a full sine wave. This scenario occurs when the roof is slightly tilted and there is a difference in the grey levels between the two sides of the roof due to the sun illumination direction (Figure15).
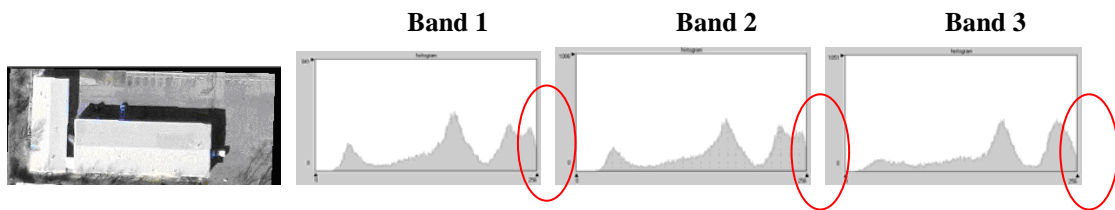
Figure 16 – Two sides roof on the left image. Bands 1 and 2 are highly correlated and have a "Saddle" geometry for the building. The third band shows a full sine wave for the building.

A "Saddle" shape is defined as two peaks that are close to one another (not more then 50 values away), their peak value is not significantly different (maximum of 1/3) and the saddle point is not less then 50 percent of the maximum number of pixels (the actual peak). In case this scenario is identified, the two peaks within the saddle area for bands 1 and 2 are combined into one peak and treated as the same object. Figure 17 illustrates the affect of the peak combination on the object segmentation. In section (b) of figure 17, each peak within the saddle was treated as a different peak, hence a different object (the right side of the roof of the building on the left was segmented but eliminated in the post processing). As a result, each side of the roof was segmented separately. In section (c) of figure17, the peaks were combined and treated as the same object. As a result, the entire roof was extracted as the same feature.

Figure 17- (a) Original image. (b) Two sides of the roof segmented separately. (c) Two sides of the roof segmented as one object in both buildings.

Two bands histograms are highly correlated and have peaks that merge several objects. The third band separates the peak, hence represent multiple objects. Figure 18 and Figure 19 show a building and areas around the building that share the same grey level values within band 1 and band 2. Those bands values lie within the same peak in the histogram. Band 3 better separates the peak in the histogram into 2 parts and allows to differentiate between the objects. Figure 20 illustrates the difference between the peaks in bands 1 and 2 versus the peak in band 3. In bands 1 and 2, the peak represents grey values of 2 objects, the building and the road around the building. In order to distinguish between the objects, we need to recognize the separation within band 3. Section (a) in figure 20 shows the segmentation result when only bands 1 and 2 are taken into consideration. Section (b) in figure 20 shows the segmentation result of bands 1, 2 and 3.
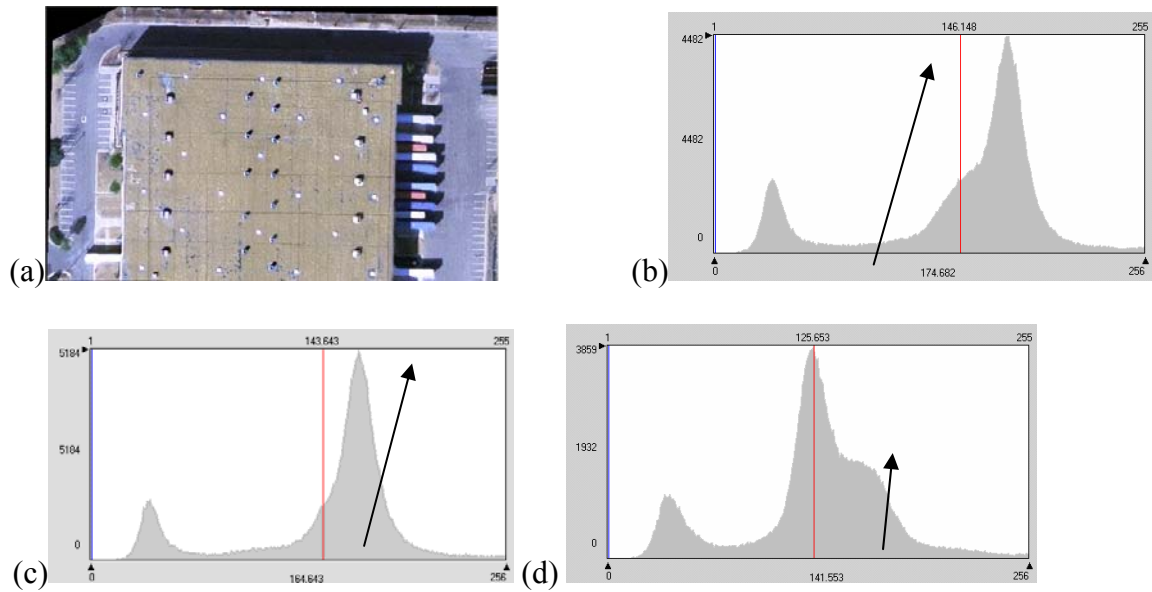
Figure 18 – (a) Original image (b)Band 1 histogram (c)Band 2 histogram (d)Band 3 histogram. Band 3 separates the peak into 2 parts.
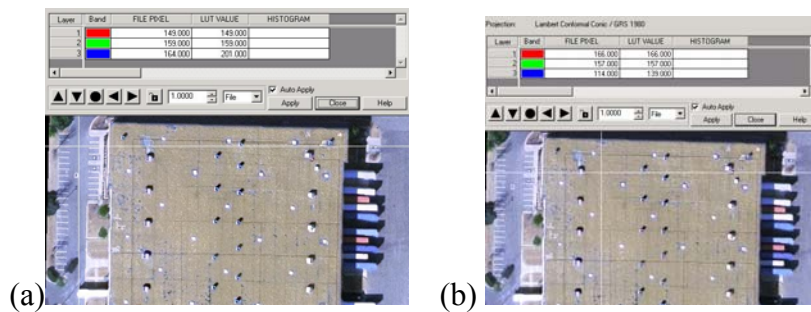
Figure 19 – (a) point around the building. (b)point on the building. Bands 1 and 2 have a value within the same peak and band 3 has a different value.

(a)   (b)
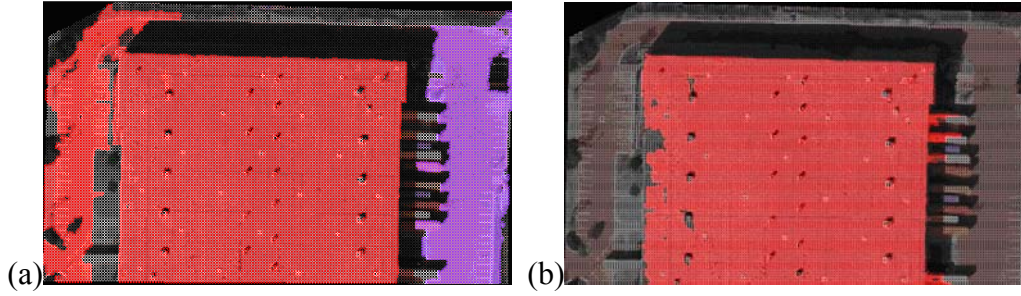
Figure 20 – (a) Red segment represents a peak within bands 1 and 2 (b) Red segment represents all 3 bands. The building is better separated from the surrounding objects in (b).

Due to the locality of the search for buildings within the parcel area, the histogram analysis allows an easy correlation between objects and peaks within the histogram. The highest peak in the histogram may not necessarily represent the built area within the parcel. An evaluation for the number of peaks recommended for building extraction analysis is provided in section 4.2.5. Based on manual observation and evaluation of the multiple peaks within the histogram, we allow for up to four different peaks. Each peak information includes the peak value, the left threshold value and the right threshold value. The peak information is referred to as a "peak interval".

4.1.2.2 Feature Segmentation

Based on the histogram analysis, peaks are segmented back into the image space. Each pixel is checked against the peaks information. If the pixel value is found to be located within an interval of a specific peak (in all 3 bands), the pixel new value is the peak value.   A peak may represent a whole feature (Figure 21-a), a section within a feature (Figure 21-b) or several objects with the similar spectral characteristics (Figure 21-c).
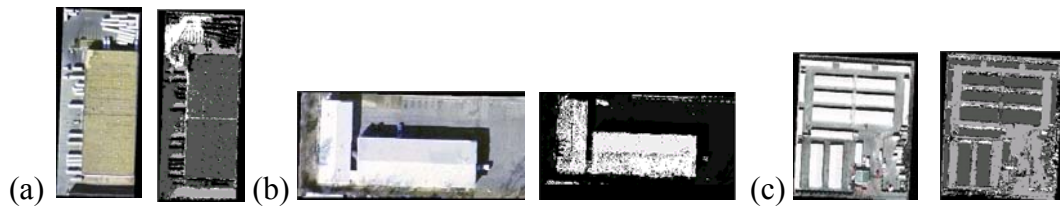
Figure 21- segmentation result. (a) One peak represents the entire building. (b) Different peaks represent the building as multiple sections. (c) Multiple features share the same spectral characteristics – same peak value.

The result of the segmentation is a new image. Each pixel value in the image represents a value of a peak or a background value (0). The pixels are later clustered into distinct objects.

4.1.2.3 Shadow Segmentation

Shadows are the darkest areas in the image. Initial observation of parcel-sized images revealed that shadows can be adequately identified as the lower 20% values within the grey level variance in the image. For example, if an image values span between 20 and 240, then all pixels with values between 20 and 64 are flagged as the shadows. The analysis is done in all 3 bands and the result is a "Shadow Image" (Figure 22).
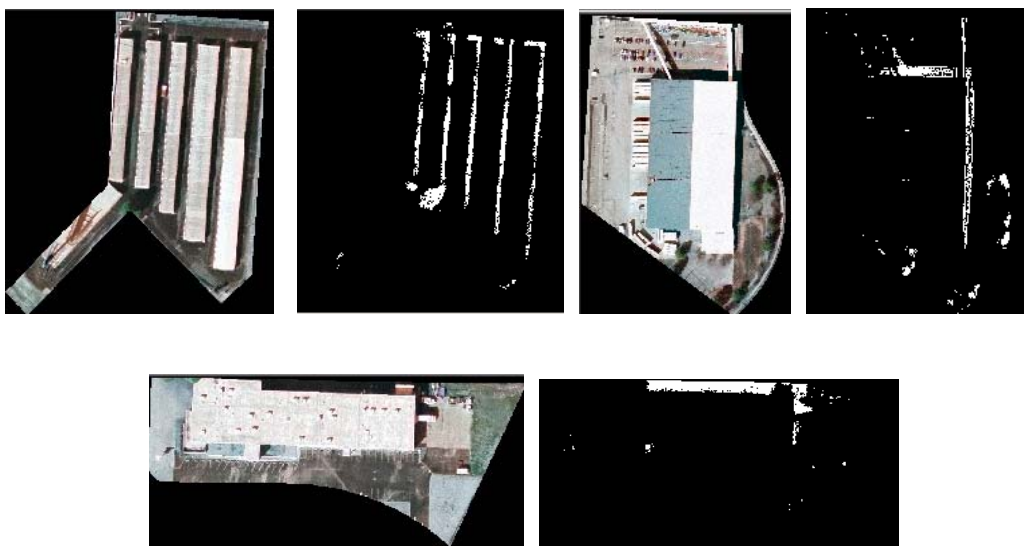
Figure 22 – Original image on the left and "Shadow Image" on the right.

Pixels that are flagged as shadows are later eliminated from possible feature analysis, therefore, in this step, it is necessary to realize that there may be an overlap between shadow pixels and feature pixels when the feature is dark. For buildings with relatively dark roofs (figure 23), the feature values may be represented within the histogram as adjacent to or interleaved with shadow values.

(a)   (b)
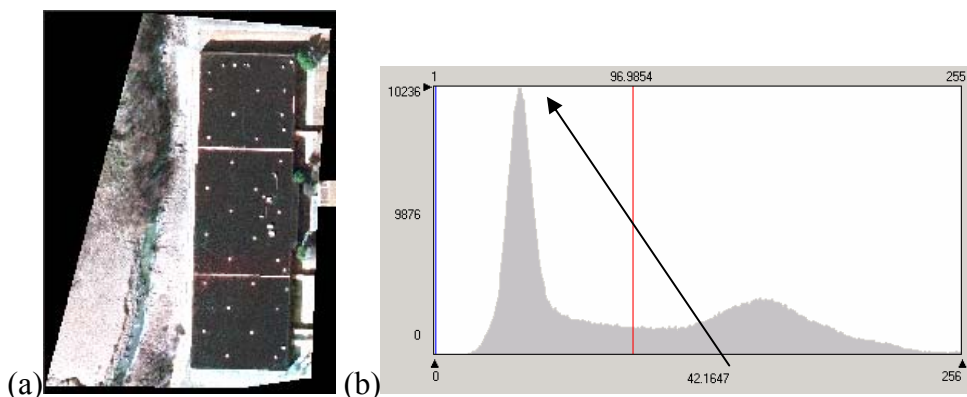
Figure 23 – (a)Dark building. (b) Band 1 histogram. The building roof and the shadow share similar spectral characteristics.

In Figure 23, the building is represented by values between 30 and 50, while the shadow pixels are represented by the values 20 to 43. The shadow area and the building roof share similar spectral characteristics and hence have overlap area within the first peak in the histogram.   Since grey level values in the image range from 0-255, the initial shadow variance is defined as 0-50 (20% of the variance of values in the image).   The initial result of the segmentation is shown in figure 24. Section (a) in figure 24 represents the feature segmentation result. Very similar result is depicted in figure 24 section (b) which represents the "Shadow Image". In order to restrict the area of the building and still allow adequate segmentation of the shadow area, the shadow area is restricted to 15 percent of the grey level values in the image, instead of 20 percent or to the lower limit of the feature range. As a result, in the example provided in figure 27, the shadow range changes from 0 to 50, into 0 to 37 and the feature range changes from 30-50, into 37-50. Narrowing the shadow range, hence changing the feature range, allows a better distinction between shadow and features. In the example, figure 24(c) represents the final object segmentation result. The dark shadow on the top of the building is flagged as a shadow and eliminated (see section a), while the side shadow area is attached to the building segment.   This is a better result then the initial segmentation result of the building and shadows as illustrated in figure 24 sections (a) and (b).

(a)   (b)   (c)
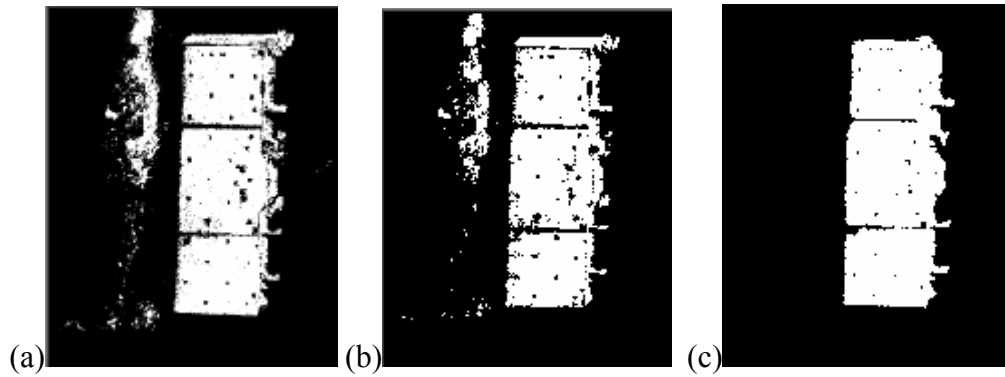
Figure 24 – Segmented result of a dark building. (a) Objects segments (b) "Shadow Image"

(c) Final object segment

4.1.2.4 Segments Post Processing

The result of the segmentation is two images. One raster image with values that represent the different peaks (thus potential buildings) and another that represents shadows ("Shadow Image"). In order to extract buildings in the image, the pixels have to be clumped into distinct objects. The post-processing of the segmentation is aimed at achieving segments that can be further analyzed to determine whether they represent buildings. To that goal, the post-processing consists of several steps including clumping connected pixels with same peak values into continuous segments, clumping connected shadow pixels into shadow segments to allow further shadow analysis and finally closing small gaps (islands) within the segments. Example for the post-processing result is illustrated in figure 25. The segmentation result, as seen in figure 25, section (b) is individual pixels, each with the value of a peak within the histogram. Section (c) shows the result of clumping the pixels into objects, where each color represents an object. Section (d) shows the result of the shadow segmentation for the image.

59
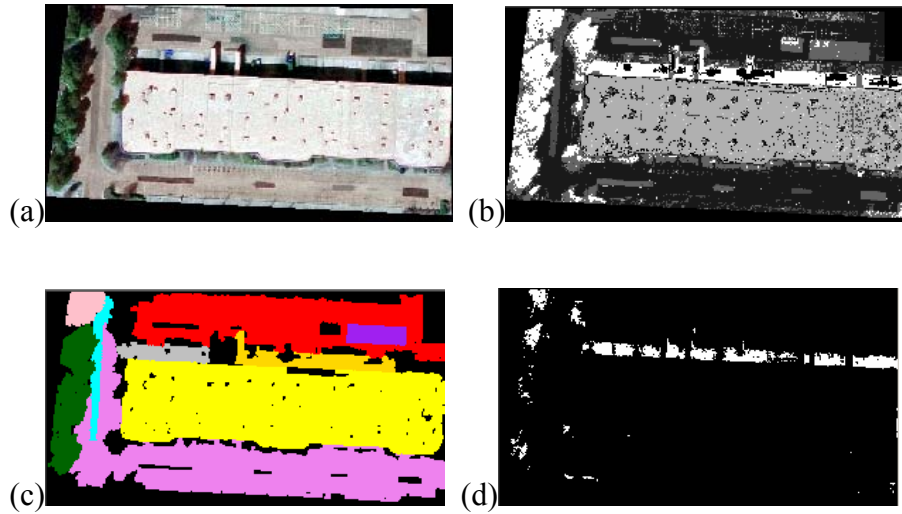
Figure 25 – Segment Post-Processing. (a) Original image (b) Result of feature segmentation (c) Result of segment post-processing. Each color represents a clumped segment (d) Result of shadow segmentation

Figure 26 illustrates the same post-processing result within a parcel that contains 4 different buildings. It is apparent from both examples (figures 25 and 26) that the transition from segmented pixels to objects is necessary and provides the analysis process with both valid buildings segments as well as miscellaneous objects around the buildings such as parking lots, trees and roads.
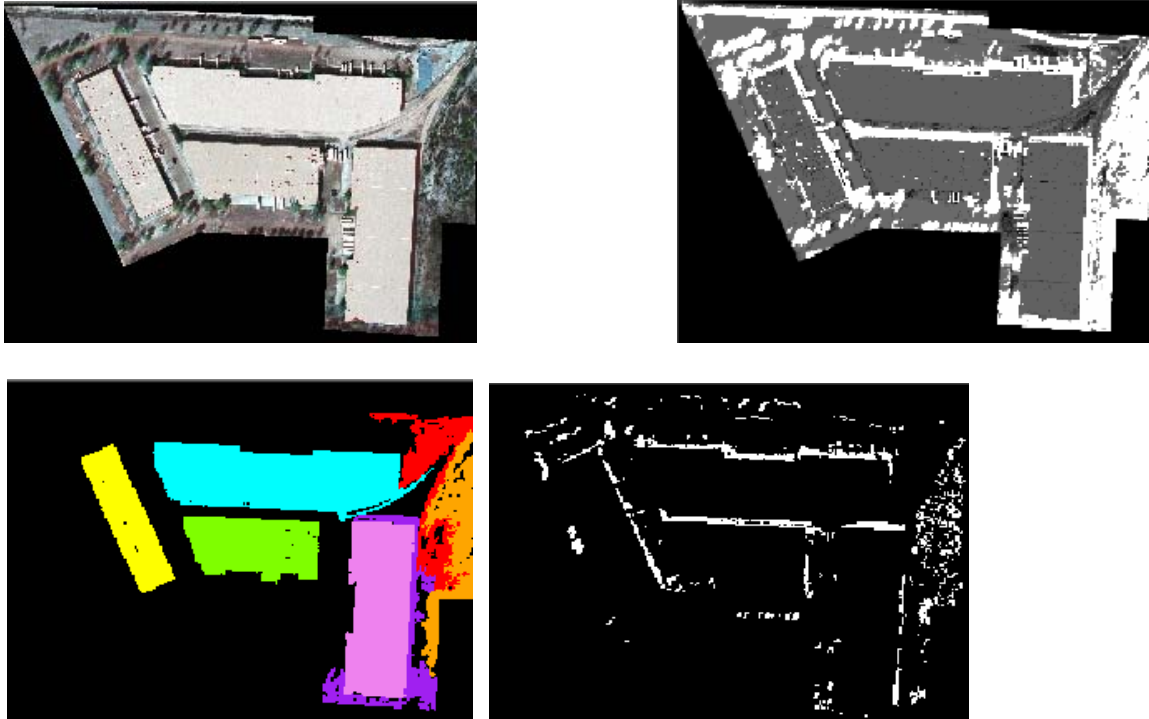
Figure 26 - Segment Post-Processing in a multi-building parcel. (a) Original image (b) Result of feature segmentation (c) Result of segment post-processing. Each color represents a clumped segment (d) Result of shadow segmentation

The clumping procedure is a simple method that is used to search for similar pixels and create a continuous "clump". For each pixel, the neighborhood of the pixel is searched for pixels with the same value. A neighborhood can be four or eight pixels (figure 27). If a neighbor has the same pixel value, it becomes part of the clump and the search continues to that pixel.
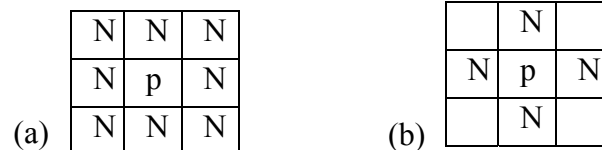
Figure 27 – Neighborhood of a pixel. (a) 8 pixel neighborhood (b) 4 pixel neighborhood

Any pixels that are clumped into very small segments (less then 2 percent of the parcel area) are eliminated. Many small segments are created at this stage and pose a burden on down the line analysis. Moreover, this ensures that relatively small, insignificant segments that are not dominant within the parcel, are eliminated at an early stage.

The following elimination steps in sections 4.1.3, 4.1.4 and 4.1.5 aim to eliminate the miscellaneous segment while keeping the building segments for the final result.

4.1.3   Eliminate By Parcel Attribute Analysis

Tax assessor data provides valuable information that can be used within the building extraction process. This section elaborates on the tax assessor information used for elimination of miscellaneous objects. As any other source of information, the tax assessor data has to be updated and as close as possible to the time the imagery was acquired. Examples of the elimination will be provided as well as observed limitations.

Figure 28 details selected attributes from the tax-assessor database. Each parcel has a unique ID (PAR_ID), number of stories (NO_STORIES), the area of the building (SQ_FEET) and other information such as the year the building was built and land use. Tax assessor information varies in details and format between jurisdictions. Initial inquiry regarding the nature and details of the information is required. In the information provided, as illustrated in figure 28, the actual expected area of the building outline is the total area of the building divided by the number of stories, hence equals to $\frac{SQ\_FEET}{NO\_STORIES}$. Each parcel is identified based on the PARCEL_ID.

SQ_FEET	NO_STORIES	OCC_TYPE	OCC_DETAIL	YEAR_BLT	PAR_ID
10744	1	RES3	- MULTI-FAMILY RESIDENTIAL	1972	050056 00040
10744	1	COM7	- OFFICE MEDICAL	1997	D0207 00301
10740	1	COM1	- SHP-CTR-STRP	1965	046008 00018
10738	1	RES3	- MULTI-FAMILY RESIDENTIAL	1963	025034 00019C
10735	2	RES1	- SINGLE-FAMILY RESIDENTIAL	1997	080008 00372
10734	1	COM7	- OFFICE MEDICAL	1984	056062 00010C
10728	1	RES3	- MULTI-FAMILY RESIDENTIAL	1973	079013 00066
10727	1	RES3	- MULTI-FAMILY RESIDENTIAL	1962	017048 00015
10721	2	COM4	- FUNERAL HOME	1964	037035 00001C
10720	1	COM1	- STORE-RETAIL	1967	073011 00008
10719	1	COM1	- STORE-RETAIL	1997	044011 00056
10716	1	RES3	- MULTI-FAMILY RESIDENTIAL	1990	080008 00419
10715	1	IND1	- MANUFCT MILL	1970	024001 00038
10712	1	RES3	- MULTI-FAMILY RESIDENTIAL	1959	022021 00024
10710	1	COM4	- OFFICE LOW	1973	060174 00113L
10704	1	RES3	- MULTI-FAMILY RESIDENTIAL	1920	020076 00034
10703	1	COM2	- AUTO DEALER	1979	074089 E00009
10701	1	COM2	- WHS STORAGE	1986	089028 B00073
10695	1	COM1	- SHP-CTR-STRP	1958	073014 00002C
10690	1	COM7	- OFFICE MEDICAL	1963	001035 00002
10690	1	COM2	- WHS STORAGE	1979	050109 00002
10680	1	RES3	- MULTI-FAMILY RESIDENTIAL	1970	078028 00297
10680	1	COM4	- OFFICE LOW	1984	093106 E00012
10680	2	COM2	- WHS STORAGE	1969	060176 00328
10680	1	COM2	- WHS STORAGE	1970	088063 00031
10675	3	COM8	- FAST FOOD	1900	002039 00011
10675	1	COM2	- WHS STORAGE	1976	073007 00036
10674	2	COM4	- OFFICE LOW	1986	065010 B00013

Figure 28 – Selected attributes from the tax-assessor database

Figure 29 illustrates the use of building area within the elimination process. Section (a) shows the original image and section (b) is the segmentation result. Each segment in section (b) has a color value and associated number of pixels which represent the size of the segment. The size or area of the segment is reflected as a number in the histogram column in figure 29, section (d). The histogram reflects the number of pixels in the segment, and since a pixel size is 1x1 ft, the number of pixels is in fact the area of the segment. The black color, value 0, represents the background and is completely ignored. Section (e) details the tax assessor information for the parcel. In this example, the commercial (OCC_TYPE is COM2) parcel contains a building that was built in 1990, has 38810 square feet and only 1 level (NO_STORIES equals 1). The two results that have values close to the area in the database are the light grey segment (44853 pixels) and the

63

white segment (28319 pixels). The dark grey segment that has 110487 pixels is eliminated based on the size, and the output of that step is illustrated in section (c) and consists of two segments.

(a)   (b)   (c)

(d)

Row	Histogram	Color
0	144985	
1	110487	
2	44853	
3	28319	

(e)

SQ_FEET	NO_STORIES	OCC_TYPE	OCC_DETAIL	YEAR_BLT	PAR_ID
38810	1	COM2	· WHS STORAGE	1990	093400 00501
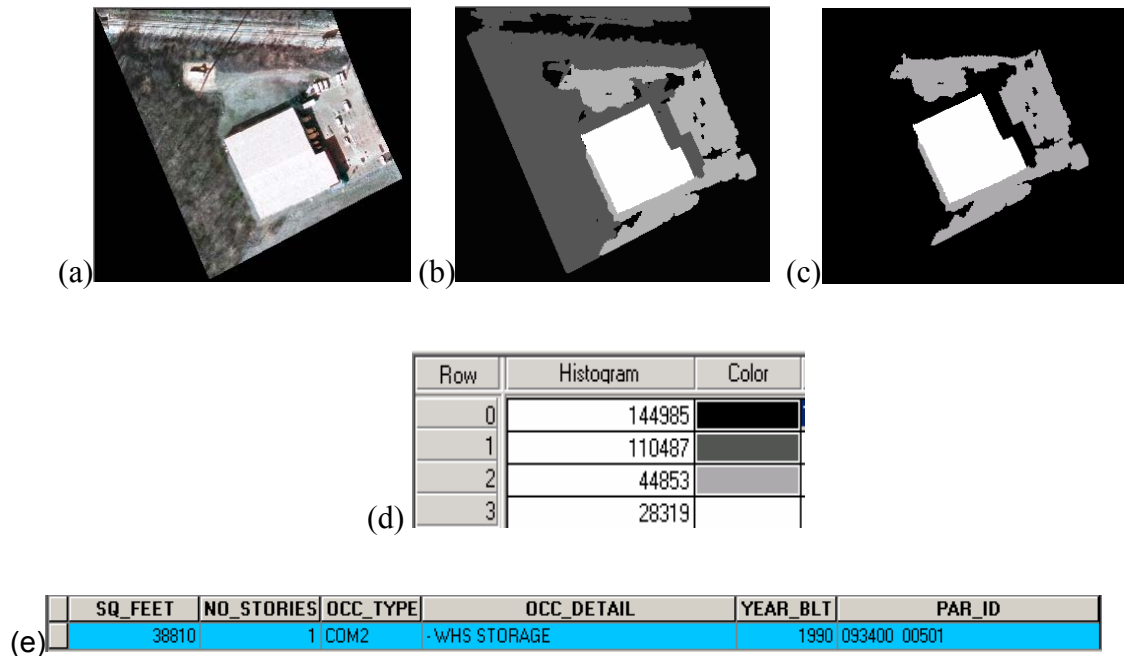
Figure 29- Eliminate by Parcel-Attribute result. (a) Original Image. (b) Objects segmented in the image (c) Objects that remain after the size elimination process (d)) Size (Number of pixels) of each segmented object (e) Parcel details in the tax-assessor db

As shown in figure 29, it is not uncommon to have a discrepancy between the value of the area in the database and the extracted size of the segment from the image. There are several reasons for that discrepancy:

Currency of the Database – when the database is not up-to-date or there is a time gap between the database generation and the time the image was acquired, discrepancies between the database and the image are expected.

Extraction artifacts - Due to the nature of the automatic process, the extracted segments might have errors of omission or commission.

Figure 30 shows several cases of omission errors. Section (a) shows two segments, the grey segment representing the building and the white segment representing the surroundings. Since the building has many AC units on the roof, the segment as extracted from the image has many islands. As a result, the size of the segment is smaller then the size in the database. The size in the tax-assessor database is 97317 sqft while the extracted segment size is 92175 sqft. Section (b) of figure 30 illustrates the omission error caused by shadows of nearby trees. The database area for the building is 5120 sqft while the extracted segment size is 3596 sqft. Sections (c) and (d) illustrate a very common artifact of the extraction process. Due to tilted roof, different grey levels on the roof or division of the roof into several sections, the actual roof area in extracted as different sections. In section (c), the roof is divided into two sections due to the different colors on each side. In section (d), due to the sun illumination direction, the different sides of the roof are extracted as two separate segments. The area in the database is 14000 sqft, while the white extracted segment is 7028 and the grey extracted segment is 7044 sqft.

(a)

(b) (c)

(d)
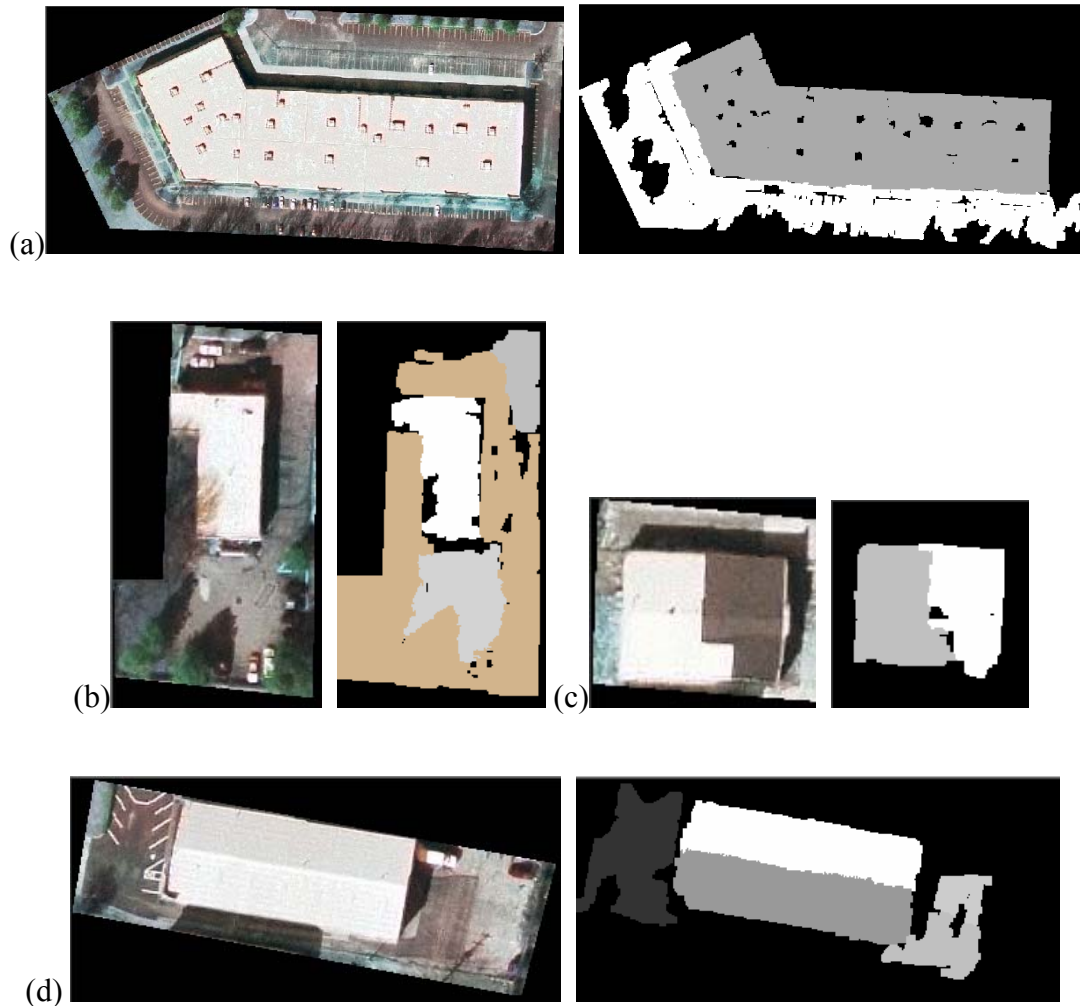
Figure 30 – Extraction artifact - errors of omission. Original image on the left and extracted segment on the right

Tilted roofs, as indicated in section 4.1.2.1, depending on the difference in color between the two slopes, can either be extracted as one roof segment or as two segments. The greater the difference in spectral values between the slope sides, the lower the possibility of extracting the roof as one segment.

Errors of commission usually occur when objects surrounding the building share similar spectral values with the building. In figure 31, there are two examples of additions to the extracted segments. The black building on the left has two extensions

due to similarity to the shadow area surrounding the building. The building area in the database is 16875 sqft, while the extracted segment is 18217. The grey building has a small addition on the lower right part resulting in difference between 15000 sqft in the database and 15729 sqft for the extracted segment. Due to the commonality of a 2 sides roof, in case two segments appear close in size (not more then 20% difference), touch each other and have the same orientation, the cumulative area will be evaluated against the tax assessor area. If the areas match, the segment will not be merged, but will remain for the next analysis.
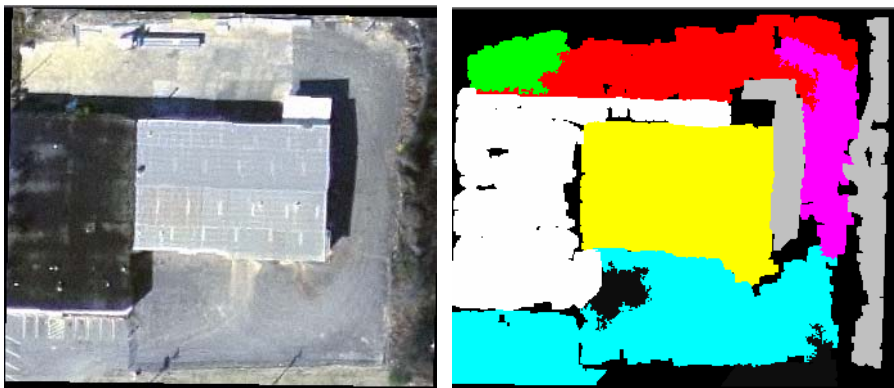
Figure 31 – Extraction artifact – errors of commission. Original image on the left and extracted segments on the right

As indicated previously, there are cases when the roof is divided into several sections. That usually occurs when there are division lines on the roofs (appear as lines) or when there are several levels on the roof that may seem divided due to shadow cast. Figure 32 illustrates such a case where the roof seems to be divided into two sections, a large section on the left and a narrow section on the right. The dark line between the sections result in two different segments, since the pixels, although segmented with the

same peak value, are not connected. The red area indicates the large segment that was not eliminated based on the area.
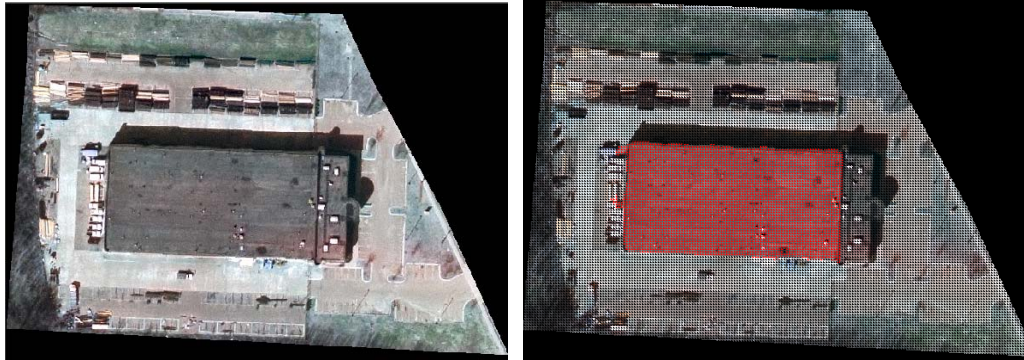
Figure 32 – Extraction artifact – only one building section extracted

Errors in the database – Erroneous information in databases is not uncommon and not limited to geographic datasets. Human errors, data entry mistakes and miscalculations are just few reasons for those errors.

During the project, it became evident that in some cases, the area of the building footprint was not the sqft area (in the database) divided by the number of stories, but rather the sqft area itself. In order to identify how common those cases are and decide on the best analysis approach, the discrepancy between the area of the digitized buildings and the area in the database was evaluated.

The building dataset consists of a point at the centroid of every parcel. The point may be linked one or several buildings information based on the number of buildings in the parcel. The digitized buildings dataset consists of polygon buildings outlines. In order to evaluate the discrepancy in the area, the ratio between the digitized building area and the area in the tax assessor database was calculated. Ratio close to one indicates small discrepancy between the datasets. For a large number of parcels, it is quite complex to

link multiple building entries that are physically represented by one point at the centroid of the parcel, to a specific polygon outline in the digitized building layer. As a result, the evaluation was limited to parcels that consist of a single building. Each point from the tax assessor information was linked to a digitized building outline.   That link was based on geographic join between the layers. Each point and polygon was first joined to the parcel it resides within. Based on the parcel id, each point was attached to a specific polygon. The initial join to the parcel eliminates the possibility of erroneously joining a point to a closest polygon that is outside of the parcel.

The evaluation was separated to buildings that have multiple stories and buildings with one level, in order to isolate the affect of the number of stories. For multiple stories buildings, the discrepancy was first calculated between the digitized area and the area in the database divided by the number of stories. Then, the difference was calculated between the digitized area and the area in the database. For the one level buildings, the discrepancy was calculated between the digitized area and the area in the tax assessor database.

In the single building parcels, 6367 buildings in the database were identified as multi-stories buildings. The ratio between the digitized area and the area in the tax assessor area is shown as a graph in figure 33.

It can be seen from the graphs in figure 33 and figure 34 that the majority of buildings have a ratio of 2 (3355 buildings) and 3 (2587 buildings). Only a total of 80 buildings can be considered close to a ratio of 1 (1 +/- 0.2)

Area Ratio between the digitized area and tax assessor database area for multi-stories buildings
(digitized area / (db_area/number of stories))
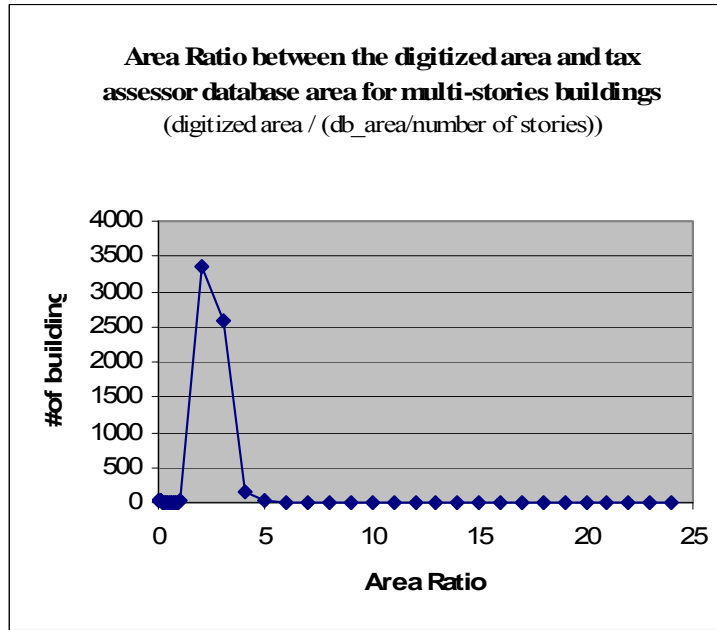
Figure 33 – Ratio between digitized area and tax assessor area for multi-stories buildings

Figure 34 shows in greater details the area ratio within the 0 to 5 range. For the ratio of 4, the number of buildings drops to 155.

Area Ratio between the digitized area and tax assessor database area for multi-storie buildings
(digitized area / (db_area/number of stories))
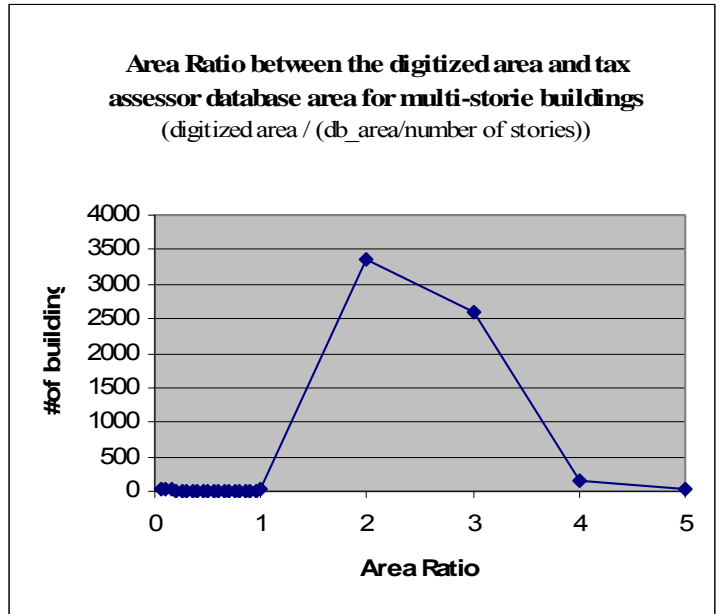
Figure 34 - Ratio between digitized area and tax assessor area for multi-stories buildings

An examination of the relationship between the area ratio and the number of stories in the building, as seen in figure 35, reveals no direct or special relationship between the number of stories and the area ratio.
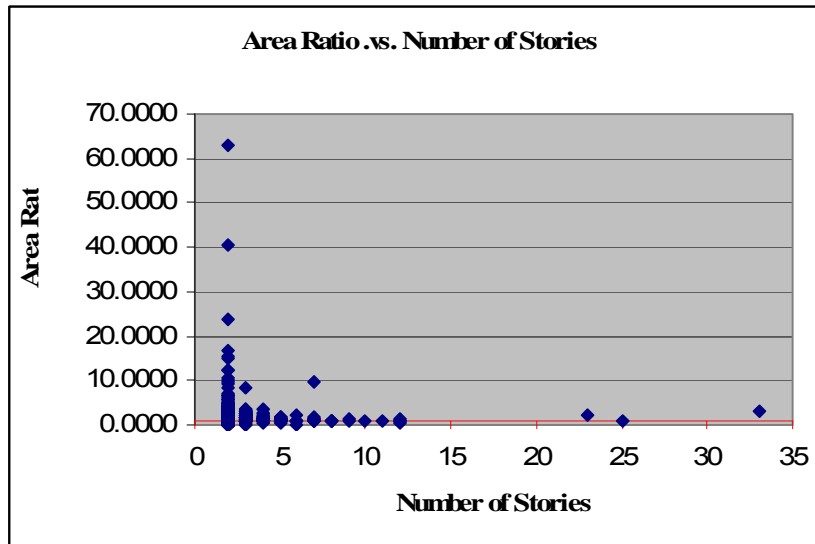
Figure 35 – Area Ratio .vs. number of stories for multiple stories buildings

The red horizontal line indicates a ratio of 1. It is obvious that no specific number of stories is linked to a specific range of area ratios. Since we have a large number of buildings with 2 stories, there is a greater range (0-63) of values for the area ratios of those buildings. For a 2 stories building, values greater then 4 or smaller then 0.25, might indicate a plausible error either in the database information or the digitized building dataset. A total number of 458 buildings, which represent 7.2% of all the buildings, have such a large discrepancy (60 buildings with values greater then 4, the rest lower then 0.25).

To better evaluate the difference between the area of the digitized buildings and the area in the tax assessor database, a calculation of this difference was performed. That

difference was then divided by the digitized area and the expected result should be 0 when the tax assessor area and the digitized area are exactly the same. The average difference between the areas was 2438 sqft, for an average building area of 4857 sqft (the difference accounts for about 50.2% of the average building area). A plot that shows the difference ratio was created for 6240 buildings (out of the 6367 buildings) that have a difference ratio that equals or is less then 1, and a second plot for the remaining 127 buildings with a difference greater then 1.

Area discrepancy between the database and digitized buildings for multi-stories buildings
(digitized area - db_area/num of stories)/(digitized area)
(6240 buildings with ratio between 0-1)
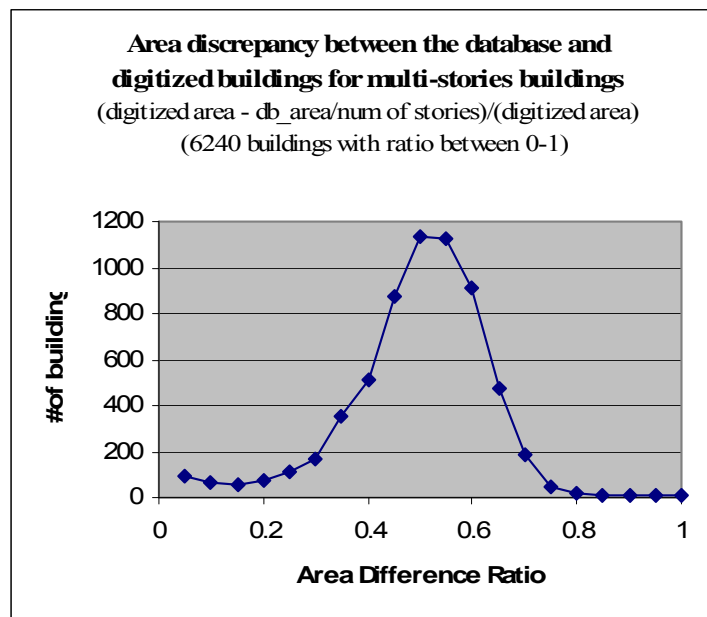
Figure 36 – Ratio between the difference in area and the digitized area for multi-stories buildings

The Gaussian distribution shown in figure 36 is centered at 0.5-0.6 area ratio and represents the majority of the sample data (98%). 5478 buildings (86% of the total number of building) have a ratio of 0.6 or less. Since the difference was calculated as an

absolute value, 0.5 ratio represents an area that is half or one and a half of the reference area. The other 127 buildings area ratio is shown in figure 37.
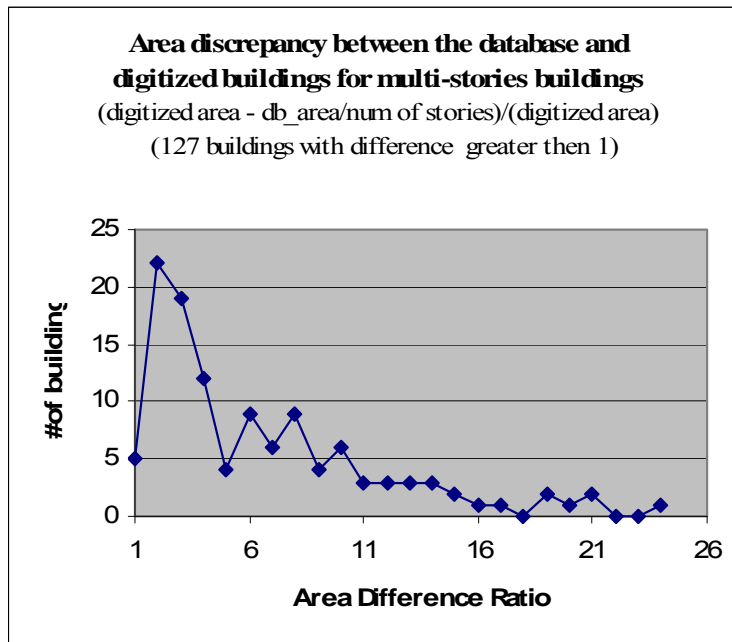
Figure 37 - Ratio between the difference in area and the digitized area for multi-stories

buildings. Only ratios greater then 1.

The great area ratios for the 127 buildings, as represented in figure 37, are the result of either errors in the database or currency problems of the tax assessor database.

Based on the area discrepancy analysis above, a threshold that allows a difference in areas in the range of 0.5-1.5 appears reasonable. This range means that the area of the segment needs to be within the range of 0.5-1.5 of the area as retrieved from the tax assessor database. Moreover, there are many cases where the area in the database is the area of the building footprint rather then the total built area of the building. Hence, the analysis based on the database area should account for those cases. As a result, the

analysis process checks the area of the extracted segment, first, against the area divided by the number of stories and if no segments are found compatible, against the total area.

The difference between the areas was also examined for one-story buildings. For the one-story buildings 12442 were examined. The evaluation revealed an average area difference (between the digitized buildings and the tax assessor database) of 720.54 sqft for an average building area of 9792.35 sqft. The difference accounts to about 7.6% of the average building area, which is much lower then the 50.2% found for the multi-stories buildings. The difference in percentage when calculated separately for each building was found to be 11.7% (difference divided by the digitized area).
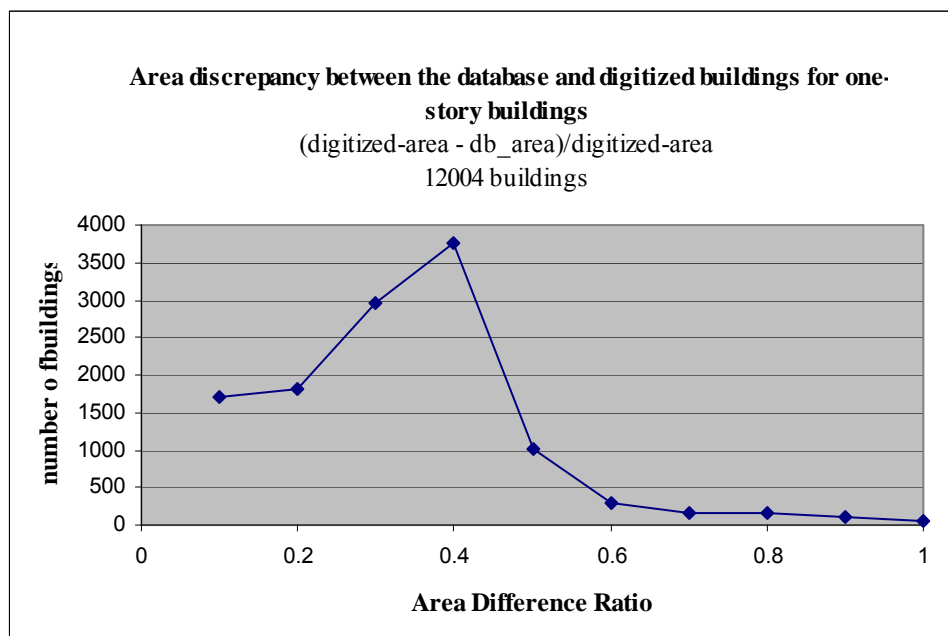
Figure 38 - Ratio between digitized area and tax assessor area for one-story buildings

(12004 buildings ratios between 0-1)

Figure 38 shows the ratio of the difference on a scale of 0 to 1. A scale of 1 means the digitized area is twice the size of the area in the database. Any ratio greater than 1 for one-story buildings indicates an error in the data. 438 buildings out of 12442 were identified with data errors and not shown in figure 38.

From both the multi-stories and the one story buildings it is evident that there is an inherent discrepancy between the digitized area and the area value in the database. Hence, the analysis of the segment extracted from the image has to account for more then random artifacts of the extraction process. In order to evaluate the use of tax assessor area within the extraction process, the threshold of the area can not be too large, as it will allow too many segments to remain as legitimate buildings. As a result, the process allows for an area discrepancy between 0.5-1.5 of the area in the tax assessor database.

4.1.4   Eliminate by shadow Analysis

The shadow analysis is based on the fact that due to the sun-illumination, three dimensional objects cast a shadow on the ground. The shadow should appear in a specific location relative to the building. As illustrated in section 4.1.4, the user provides the sun illumination direction. This is a relatively easy task that requires a scan of the image and determination of the location of the shadow. For example, in figure 37, the sun casts shadow on the northern and eastern sides of the buildings (north is up). The user should indicate a sun-illumination direction as south-to-north (S-N) and west-to-east (W-E) in the GUI (figure 39).

Figure 39 – Sun illumination direction. South to North and West to East

The shadow analysis is based on the initial "shadow image" as created as a result of the shadow segmentation (section 4.1.2.3). Since the shadow segment and the building segment are adjacent, a 10ft buffer is created around the segmented shadows. Any segment that overlaps with the buffer is then flagged as a possible building.

Next, the relative location of the shadow and the segment is examined. First, the location and orientation of each feature segment and shadow segment is determined. The orientation and the location characteristics of a segment are based on image moment analysis. Using image moments the process calculates central location, image ellipse and orientation of the segment (section 4.1.7). Figure 40 illustrates the analysis of the shadow and adjacent feature segment in order to determine whether the feature segment should remain as a possible building segment for following analysis. Section (a) in figure 40 is the original parcel-sized image. Section (b) shows the result of the shadow segmentation and section (c) shows the feature segments that remain after the size elimination analysis. There are 3 different colored segments in section (e). The red segment illustrates the overlap area between the buffered shadow and the red feature segment. The green and cyan represents the overlap area respectively. Section (d) illustrates an overlay of the shadows and the feature segments as an input to the shadow

analysis procedure. Very small shadow segments are eliminated. A buffer is created around each shadow segment to determine adjacency to a feature segment. Section (e) illustrates the result of the buffer overlap with the features. Each shadow overlap area is shown in the color of the feature it overlaps. Section (f) illustrates the result of the shadow analysis. As can be determined by the image in section (a), the shadow location is expected on the northern and eastern sides of the building.

Since the red shadow is located on the western side relative to the red feature segments, the red segment is eliminated. The green shadow is located on the northern and the eastern sides of the building hence remain for the following analysis. The cyan segment remains due to confusion regarding the relative location. As illustrated in section (e), the shadow is located on the eastern, north-east and western sides of the building. Since no geometrical constraints are used to determine the shadow geometry, all cyan colored shadow areas are analyzed. As a result, the building seems to have shadows on different sides, including the expected locations. To avoid elimination of legitimate building segments, the procedure keeps segments with some uncertainty.
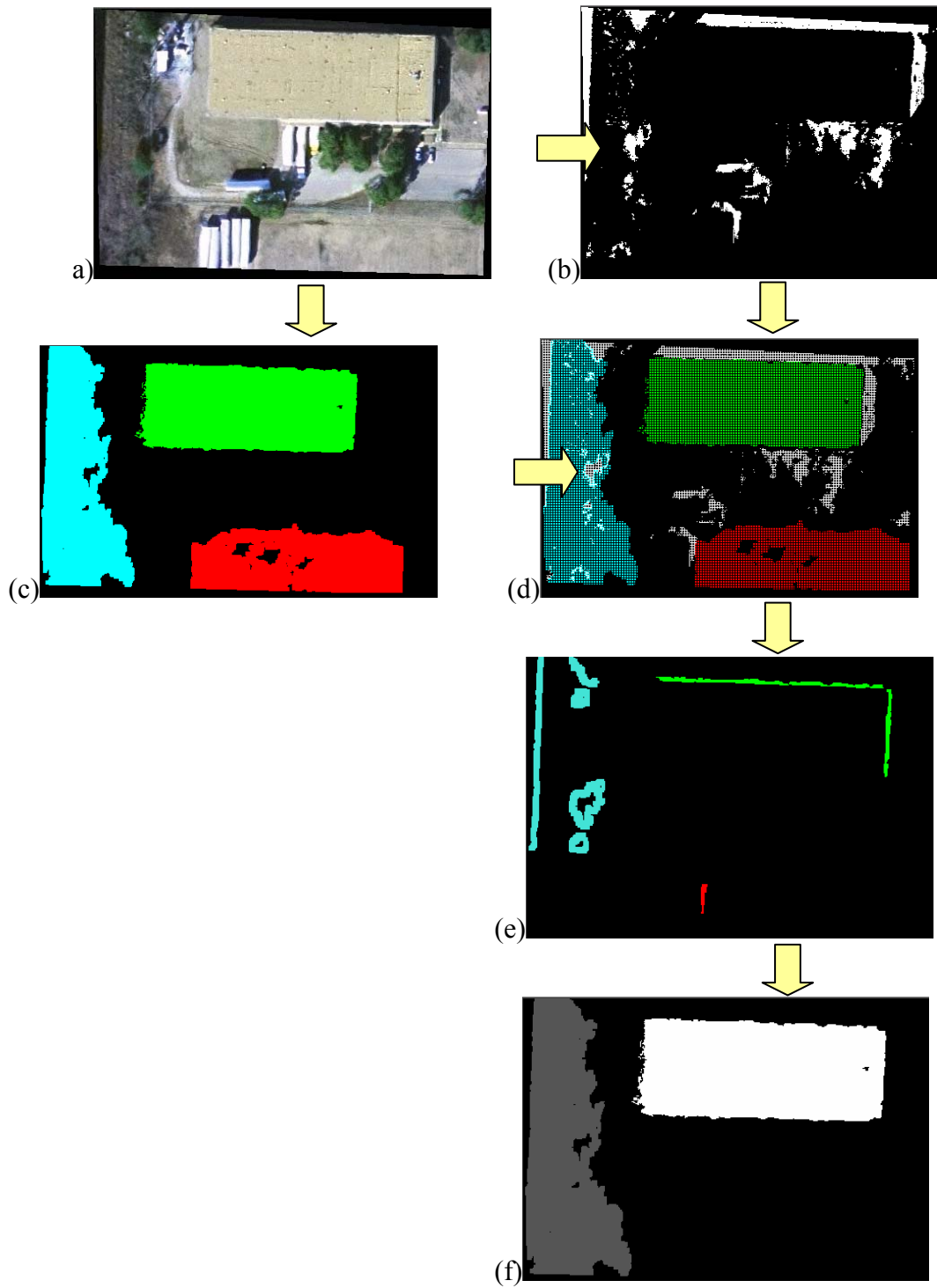
Figure 40 – Shadow analysis. (a)original image (b)Shadow segmentation result (c)Feature segments (d)Feature segments and Shadow overlap (e) shadow adjacent to the features (f)Shadow analysis result

Shadow analysis is a complicated task. Shadows from adjacent features may overlap, including trees. Shadows for buildings may not appear where expected due to occlusion of a near by feature. Moreover, sections of a building may not be detected due to shadow cast. All those issues are represented in 41. Due to the complexity of shadow analysis in an urban setting, the analysis, as detailed in this section, does not eliminate segments when there is a level of uncertainty regarding the authenticity of the segment as a building.

Figure 41 – Shadows around buildings. From left to right – commercial buildings, residential buildings, high-rise buildings.

Figure 42 illustrates the flow of the shadow analysis.
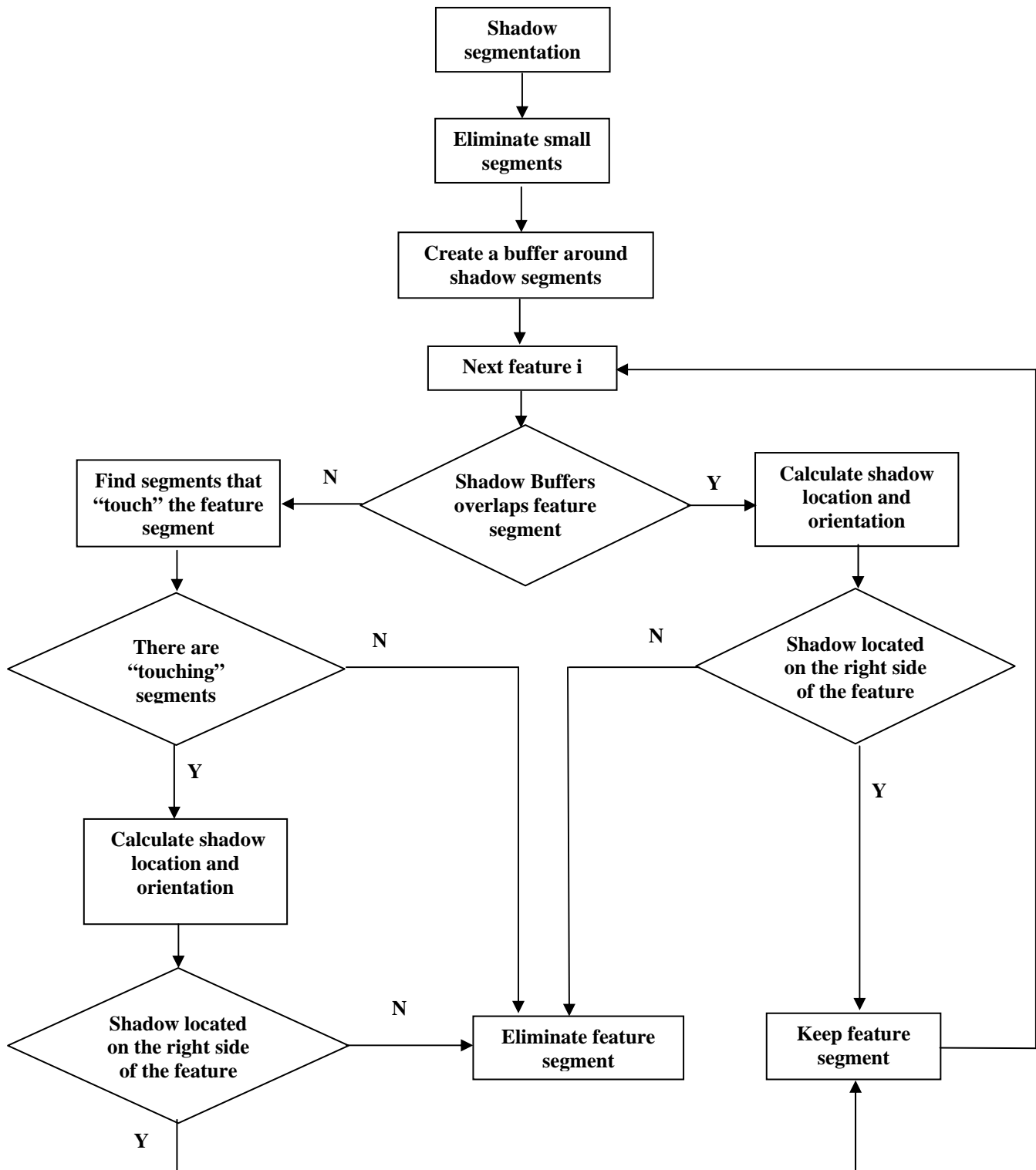
Figure 42 – Shadow elimination process flow

Figure 42 illustrates the flow of the shadow elimination process. Following the segmentation of shadows, pixels are clumped together into shadow segments as explained in section 4.1.2.4. A 10 ft buffer is created around each shadow segment. Each feature segment is then examined for possible overlap with shadow buffers (since there may be more then one shadow area around the building). In case a shadow buffer overlaps the segments, the feature is further examined for the relative location of the shadow and the feature. If the shadow is located on the correct side of the building (based on the sun-illumination input from the user), then the feature remains. If no shadow is located on the correct side of the building – the segment is eliminated. If no shadow buffer overlaps the building, topological relationship between the features is examined. If a feature segment that has no attached shadow but it touches another a feature segment that has a shadow attached on the correct side, then the feature segment remains. The touching segments have to be similar in size and orientation to qualify for this analysis. That constraint reduces possible confusion between different type of features, for example, building feature analysis and vegetation.

The result of the shadow elimination is an image with one or multiple segments that remain as possible building candidates.

The geometric analysis comprises of common measures to represent the geometric characteristics of a segment. The values of the measures as pertain to building and non-building segments are used to eliminate segments based on a probability index. In order to calculate probability index for each segment, values for each measure were calculated for manually selected building segments and non-building segments. All the selected segments are the result of an automatic segmentation, not digitized buildings. Hence, the selected segments represent the geometry of an actual result of an automatic process, rather then clean, manually processed segments.

As mentioned in section 3.5, the measures used for the geometric analysis are:

- Rectangularity - Rectangularity of shapes can be evaluated in different ways (Rosin, 1999; Rosin, 2003). The standard method for calculating rectangularity is using the MBR (Minimum Bounding Rectangle) of the segment. The rectangularity measure reflects the ratio between the area of the segment and the area of the MBR. This method is obviously sensitive to spikes in the shape and may produce a low rectangularity value for rectangular shapes with a sharp localized spike. The MBR method was concluded as one of the rectangularity calculation choices by Rosin (1999). Rectangularity can also be estimated using moment invariants (See Appendix D). The orientation calculations for this method were found to be sensitive to noise, and high rectangularity values can possibly be attributed to non-rectangular shapes if

they poses similar ratio of moments (Rosin, 1999). The moment method, during progressive noise-adding testing, was found to be not significantly sensitive to noise and kept high rectangularity values for rectangles. As a result, both the MBR and moment methods were calculated, and the highest rectangularity value for the segment was reported. More evaluation for geometric analysis using moment invariant is elaborated in section 4.1.7.

- Ellipticity - The ellipticity of a segment is evaluated using moment invariants (Rosin, 2003), which was shown to provide good classification results. Based on Rosin (2003),

$$E = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4} \qquad Ellipticity = \begin{cases} 16\Pi^2 E & \text{If } E \leq \dfrac{1}{16\Pi^2} \\ \dfrac{1}{16\Pi^2 E} \end{cases}$$

(Elaborated moment definition is provided in appendix D)

- Solidity – Solidity is defined as the Bounding Polygon Area divided by the actual segment area. That index measures the amount of holes within the segment:

$$Solidity = \frac{BoundingPolygonArea}{Area}$$

- Convexity – Convexity reflects the ratio between the perimeter of the convex hull polygon of the shape divided and the perimeter of the original shape. That index should give 1 for a complete convex shape such as a rectangle or a square. Using the convex hull of the shape, three different measures were

83

evaluated. The first convexity measure (convexity_g) was calculated based on the perimeter of the generalized polygon. As elaborated in section 4.1.5.2, that measure was not found sufficient for discerning between the building and non-building segments. As a result, other measures of convexity were examined: convexity_o, based on the perimeter of the original extracted segment, and convexity_area, based on the ratio of the convex hull area and the area of the generalized polygon.

Convexity_g was calculated as the ratio between the convex hull and the perimeter of the generalized polygon (See section 4.1.6):

$$Convexity\_g = \frac{ConvexHullPerimeter}{PerimeterOfTheGeneralizedPolygon}$$

Convexity_o was calculated as the ratio between the convex hull and the perimeter of the original segment polygon:

$$Convexity\_o = \frac{ConvexHullPerimeter}{PerimeterOfTheOriginalPolygon}$$

Similar to calculating convexity based on the perimeter, convexity can be calculated based on the ratio between the area of the convex hull and the area of the generalized polygon. The area ratio was calculated as follows:

$$Convexity\_area = \frac{ConvexHullArea}{AreaOfTheGeneralizedPolygon}$$

Area and perimeter are commonly used to represent and analyze shapes, internal representation as the area and external representation as the perimeter. Since one can be constructed from the other (Rosin, 2000) there is usually no advantage of one method over the other. Sometimes, however, one

method may make it easier to draw out certain conclusions or characteristics of the shape (Rosin, 2000). The convexity measure based on area and the measure based on the perimeter will be examined and used accordingly to determine the probability of the shape to represent a building.

- Compactness – Compactness reflects the ratio between the area of a shape and the area of a compact shape with the same perimeter. We use a square as the most compact shape rather then a circle. As a result, we need to define the compactness index accordingly:

$$Compactness = \frac{Area}{AreaOfShapeWithSamePerimeter} = \frac{16*Area}{Perimeter^2}$$

In order to calculate the geometric measures (rectangularity, convexity…), initial geometric computations are required. First, the convex-hull of the segment is calculated using the Graham algorithm (Graham, 1972). The Graham algorithm is explained in details in Appendix B. One of the convex Hull points is used as an anchor for determining the exterior ring of the segment. The exterior ring is the result of edge detection (section 3.1) followed by connecting points starting from the anchor (convex hull) point. The exterior ring is generalized (section 4.1.6). The generalized exterior ring is then used to evaluate the different measures.

Extraction artifacts and geometry analysis constraints impose evaluation of the analysis parameters based on both known, accepted values, as well as observation of extracted segments. In section 14-4-84 of the "Code of Ordinance and Character of

Memphis, TN" (available in http://municipalcodes.lexisnexis.com/codes/memphis/),
there is no restriction on the minimum width of a dwelling, but a restriction on the
minimum floor area – not less then 150 sqft. A manual scan of residential, offices and
commercial dwellings in the image as well as the automatic segmentation result
concluded a minimum width parameter for a segment as 20ft. The width of a shape is
determined based on image ellipse parameters (see appendix D) of each segment.

The size of islands within the segments was determined as not an indicator for a non-
building segment. Necessary post processing of segments, as elaborated in section
4.1.2.4, closes small gaps and reduces the size of larger gaps; building segment, due
either to features located on the roof or extraction artifacts, have large islands on the roof
(figure 43). As a result, both non-building and building roofs may have large islands and
can not be used as an effective geometric measure to discern between building and non-
building segments.
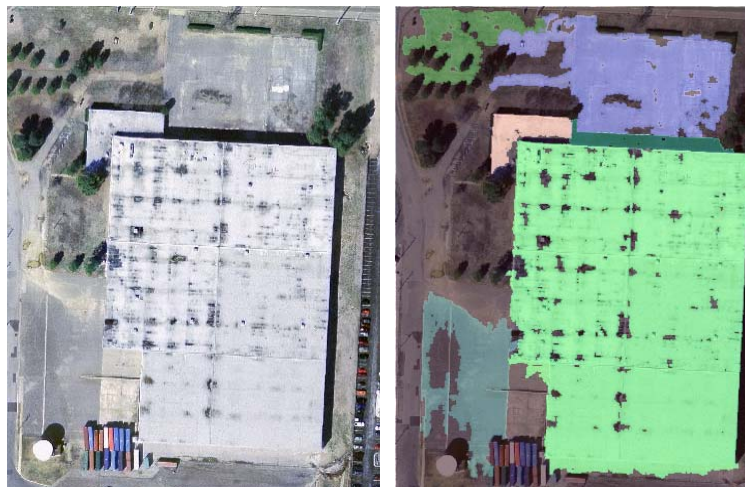
Figure 43 – Islands on a roof. Left – Original image; Right – extracted segments overlaid on the image.

Figure 44 provides few examples for the geometric measures as calculated for several automatically extracted segments. Sections (a), (b) and (c) represent each a building segment, highlighted in green. Sections (d), (e) and (f) represent each a non-building segment, highlighted in red. The database entry for each segment is attached to the right of the segment. Each entry provides the calculated values for the rectangularity based on MBR; rectangularity based on moments; convexity_g; convexity_o; convexity_area; solidity; compactness (see equations at the top of section 4.1.5). Figure 44 provides just few example of the wide variety of building and non-building segments as extracted from the imagery. Some non-building segments can be easily detected both visually and based on geometric measures, as illustrated in sections (d) and (e). Sections (c) and (f) represent a building and non-building segments respectively and illustrate the problem of discerning between the two classes due to similarity in geometric measures.

(a)

rect_dis	rect_momen	elps_index	convexity	convexity2	convex_are	solidity	compactnes
0.74942	0.9636	0.91049	0.99742	0.3717	1.00005	0.99	0.74275

(b)

rect_d	rect_mo	elps_index	convexity	convexity2	convex_are	solidity	compactne
0.7322	0.80542	0.86932	0.95047	0.35792	1.0536	0.96659	0.75832

(c)

rect_d	rect_mo	elps_index	convexity	convexity2	convex_are	solidity	compactne
0.7255	0.65602	0.86829	0.95776	0.40281	1.05617	0.97605	0.61306

(d)

rect_d	rect_mo	elps_index	convexity	convexity2	convex_a	solidity	compactne
0.5771	0.48844	0.56289	0.81771	0.26619	1.12676	0.88669	0.58155

(e)

rect_dis	rect_mom	elps_inde	convexity	convexity2	convex_ar	solidity	compactnes
0.34197	0.23764	0.31842	0.84239	0.25081	1.50677	0.982	0.4607

(f)

rect_dis	rect_mom	elps_inde	convexity	convexity2	convex_ar	solidity	compactnes
0.73459	0.80816	0.87511	0.99765	0.36593	1.00027	0.99	0.22801

Figure 44 – Calculated geometric measures for building (a,b,c; green) and non-building

(d,e,f; red) segments. The segments image is overlaid on the original image.

4.1.5.1 Evaluating geometric measures for buildings and non-building segments

In order to calculate a probability index for each segment, a range of valid values for each measure has to be evaluated separately for building segments and non-building segments. To that cause, 240 parcels were randomly selected. The parcels consist of 250 automatically extracted building segments and 334 non-building segments. Initially, the average and standard deviation (table 2) were examined for each geometric measure. In order to use those measures to calculate the probability of a segment to be a building segment, those measures should be distinct for the building and non-building classes respectively.

For all 250 building segments and 334 non-building segments, an average and standard deviation values for each measure (rectangularity, ellipticity, convexity, solidity, and compactness) were calculated. Table 2 shows the average values and the standard deviation values as calculated for all the measures. Building segments highlighted in green and non-building segmented highlighted in red.

Table 2 - Geometric measures for building (green) and non-building (red) segments.

	Building Segments Average	Building Segments Standard Deviation	Non-Building Segments Average	Non-Building Segments standard Deviation
Rectangularity-MBR	0.63852	0.20222	0.42466	0.22757
Rectangularity-Moments	0.78271	0.13690	0.44655	0.19537
Ellipticity	0.86896	0.08620	0.53547	0.27279
Convexity_g	0.95853	0.07402	0.86401	0.12962
Convexity_o	0.41457	0.06113	0.35140	0.07346
Convexity Ratio	1.07313	0.12621	1.51263	0.63973
Solidity	0.99000	0.00492	0.95708	0.09240
Compactness	0.73974	0.20044	0.48281	0.25242

The average and standard deviation values, as depicted in figure 2, reveal that those measures can not distinctly classify a segment as a building or non-building with a reliable probability. The rectangularity based on MBR measure provides an average of 0.64 for building segments and 0.42 for non-building segments. The average values may appear distinct, but with a standard deviation of about 0.2, the range of the class within just one standard deviation creates an overlap between the classes, hence classification confusion. The two convexity measures (convexity_g and convexity_o) reflect relatively small standard deviation values (0.06-0.13), but also fairly close average values (0.96/0.86, 0.41/0.35). The solidity measure averages of (0.99/0.95) reflect an obvious overlap in values between the building and non-building segments. Any class with a large standard deviation value can not be defined in a distinct enough manner to allow adequate classification and probability calculations.

4.1.5.2   Geometry parameters definition

As illustrated in section 4.1.5.2, the average and standard deviation are not sufficient to determine whether a segment represents a building with certain probability. In order to calculate the probability of a segment as a building or non-building shape, each segment (within the 240 selected parcels – see section 4.1.5.1) was visually classified. Based on the calculated values for each segment, a graph for building and non-building segment was compiled. The graphs, as illustrated in the following figures (figure 45 to figure 51) illustrate the distribution of the values for each class (building/non-building). The graphs visually provide more information about similar or different trends between buildings and non-building segment. The distribution of values and the extracted trends define certain

91

ranges of values. Each range provides a probability for a segment to be classified as a building or non-building.

Figure 45 illustrates the distribution of rectangularity (based on MBR) values between building and non-building segments. The green graph depicts the values for building segments and the red for the non-building segments. The obvious trends of both graphs reveal that rectangularity scores between 0.7-0.85 show a greater probability of representing a building. Values lower then 0.6 show greater probability of being a non-building feature.
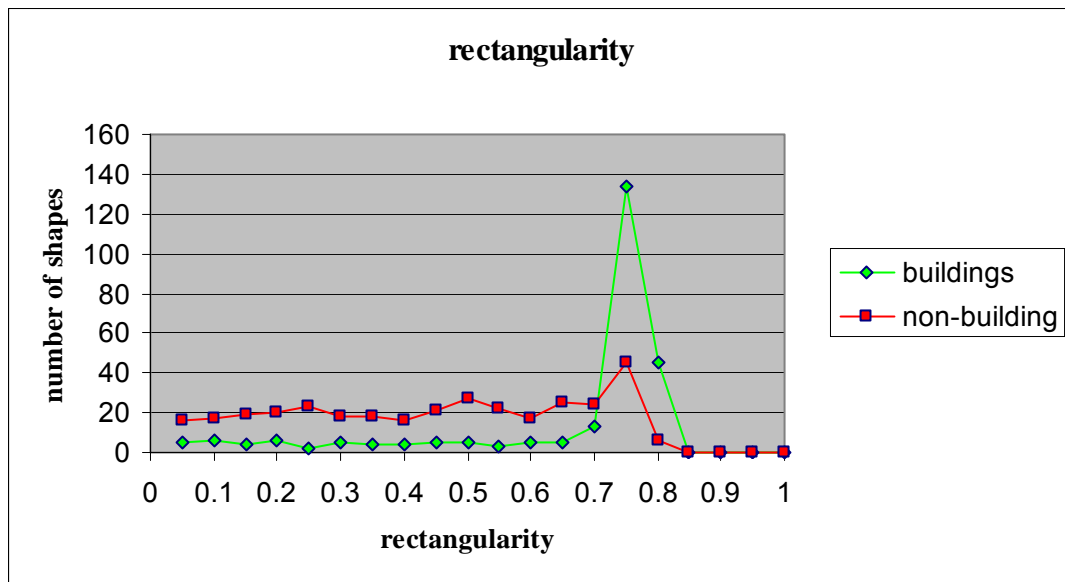
Figure 45 – Rectangularity (MBR) values distribution for building and non-building segments

Figure 46 illustrate the distribution of values for the rectangularity measure based on moment calculations. The graph depicts negative correlation, opposite trends between the building and non-building segments. Rectangularity values greater then 0.8, have a high probability of being buildings. Rectangularity values between 0.1—0.6, have a higher probability of being non-buildings then buildings. Rectangularity values between 0.6-0.8 can not indicate whether the segment is a building or not – a segment with a value in that range has the same probability of being a building or a non-building.
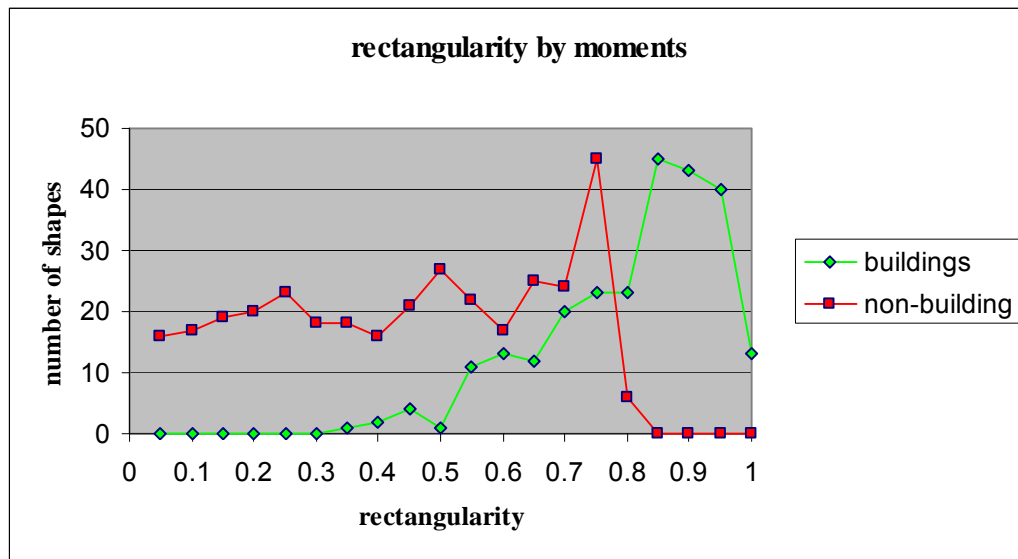
Figure 46 - Rectangularity (moments) values distribution for building and non-building

segments

.
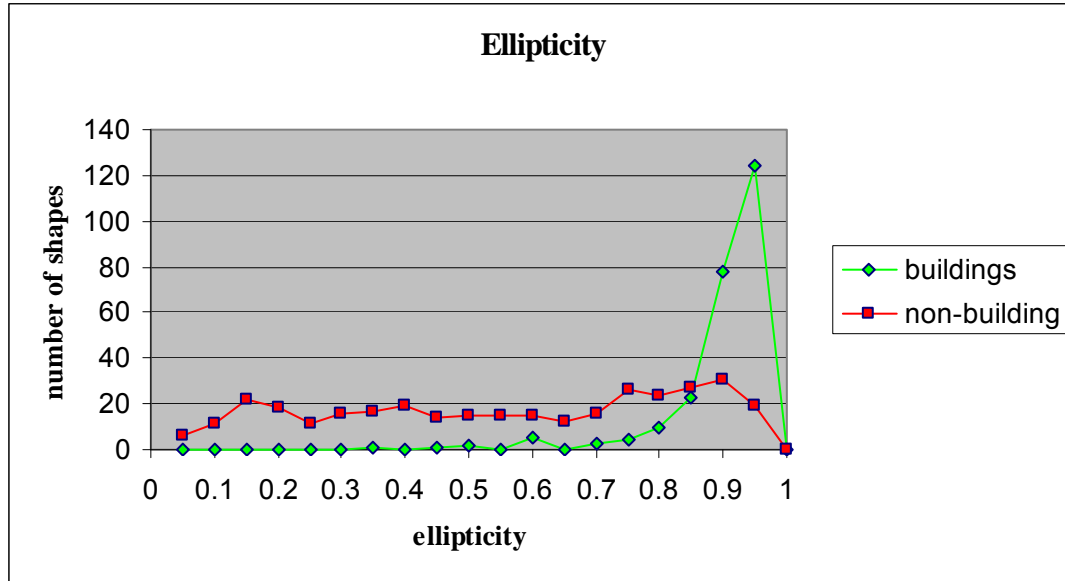
Figure 47 – Ellipticity values distribution for building and non-building segments

Figure 47 illustrates the distribution of ellipticity values for building and non-building segments. The graph shows that ellipticity values greater then 0.85 indicate a building feature with a higher probability. Ellipticity values lower then 0.6 indicate a non-building with higher probability

Figure 48 illustrates the distribution of convexity values (convexity_g – convex hull perimeter divided by the generalized polygon perimeter). Similar trend for building and non-building values is depicted in the graph. It can be inferred that convexity values lower then 0.8 might indicate a non-building feature with greater probability. Convexity values greater then 0.95 might indicate building features. Convexity values range between 0.8-0.95 shows slightly higher probability for non-building segments, but a growing trend for building segments. Due to the similar trend, a second convexity index was calculated based on the original bounding polygon rather then the generalized

94

polygon. The distribution of values for the second convexity measure (convexity_o) is illustrated in figure 49.

Convexity

Based on generalized polygon
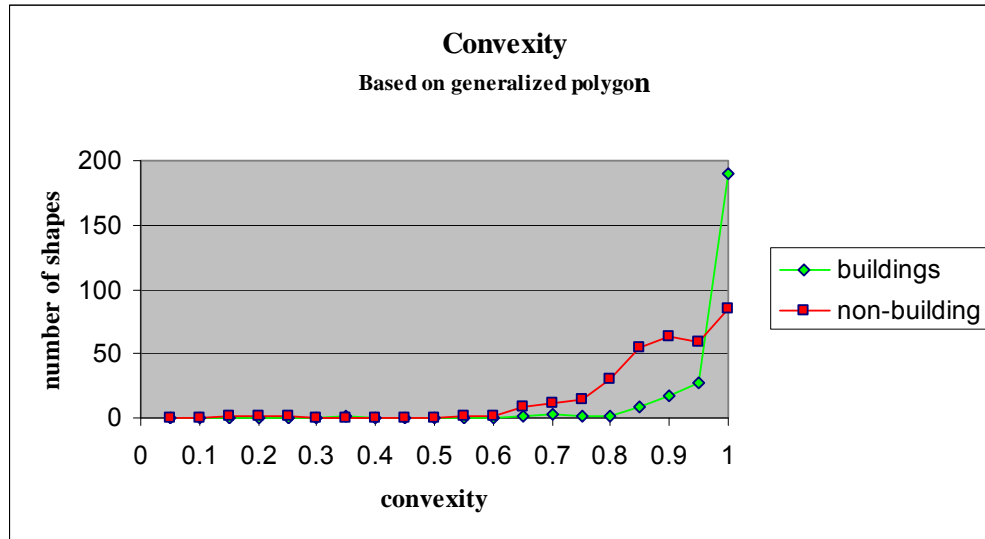
Figure 48 – convexity (generalized polygon) values distribution for building and non-building segments

The convexity index, as depicted in figure 49, indicates a similar trend between building and non-building classes. It may be inferred that very low values, lower then 0.3, might indicate a non-building feature. Due to the obvious overlap of values between the graphs, this index was not used for the confidence calculations.
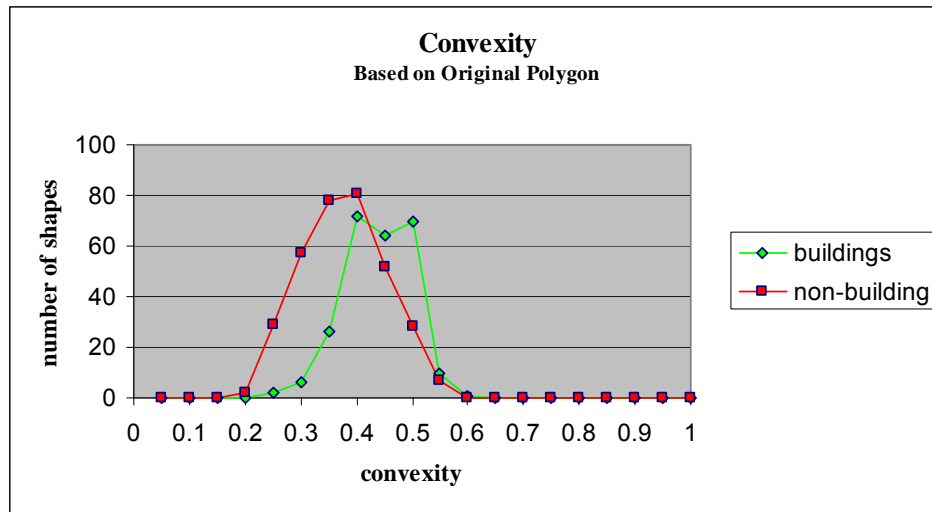
Figure 49 - convexity (original polygon) values distribution for building and non-building segments

Since the similar trend makes it hard to use this measure to distinguish between the classes of segments, a third ratio, based on the area of the convex hull and the area of the generalized polygon was calculated. The distribution for the area based convexity measure (convexity_area) is shown in figure 50.
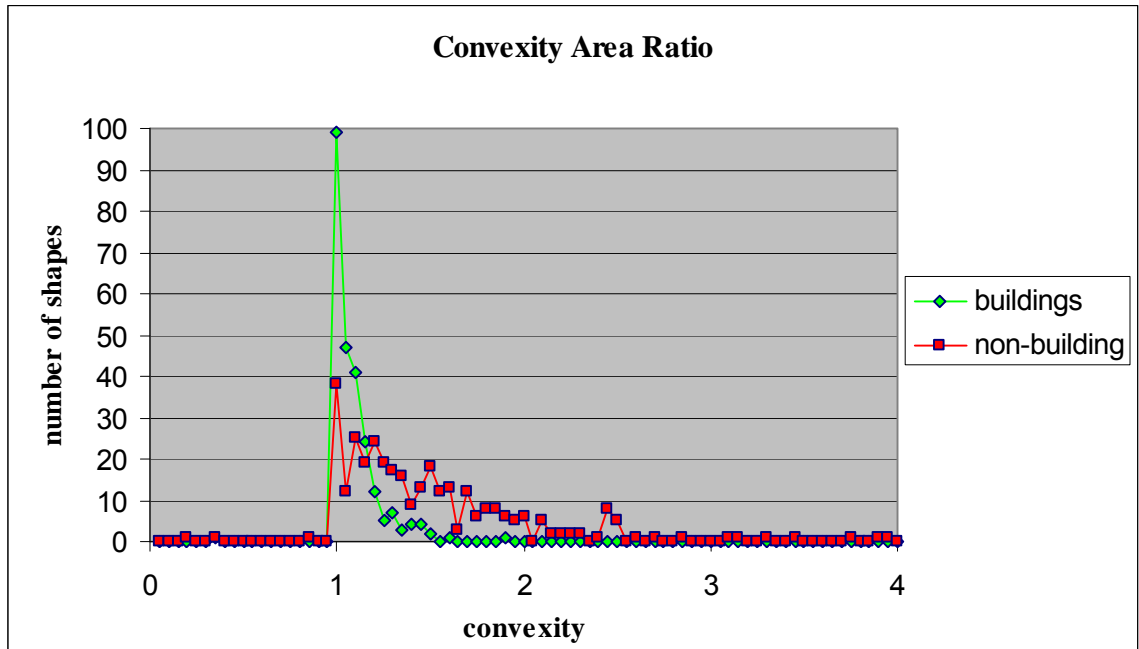
Figure 50 - convexity (area ratio) values distribution for building and non-building

segments

The convexity based on area, as depicted in figure 50, provides more information about the building and non-building segments. Values greater then 1.5 indicate a high probability of being non-building segments. Values between, about 1.35, to 1.5 indicate higher probability of being non-building features and values lower then 1.35 may indicate a building segment with higher probability.
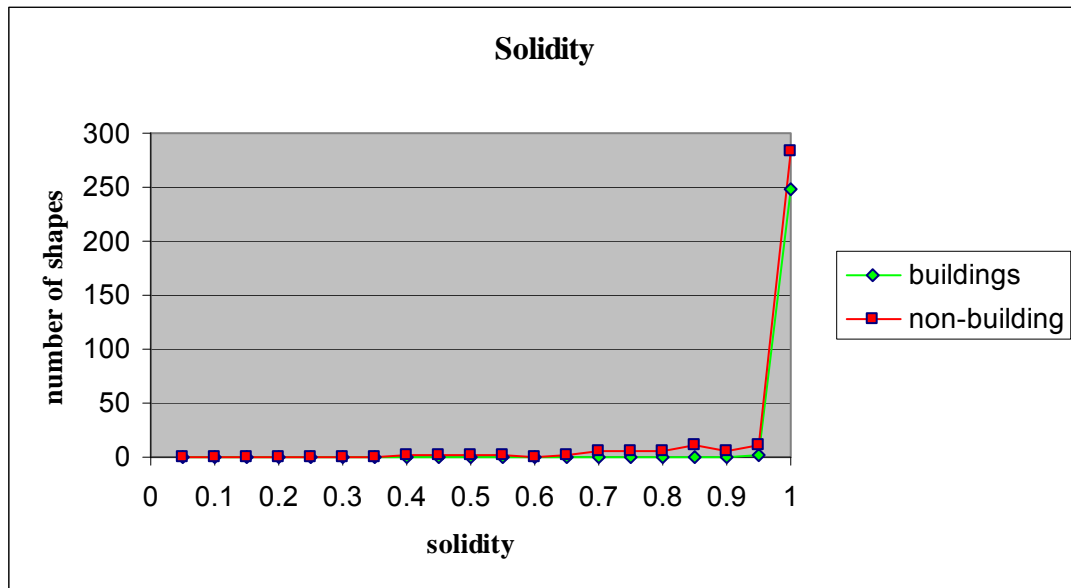
Figure 51 – Solidity values distribution for building and non-building segments

The graph in figure 51, illustrates the distribution of the solidity measure. The graph indicates very similar trends (with minor differences) between the building and non-building features. Hence, the solidity measure was not used in the geometry analysis for the segments.

The compactness measure, as illustrated in figure 52, indicates that very high values (0.85-1) represent building features with greater probability. Compactness values lower then 0.75 indicate a non-building feature and the
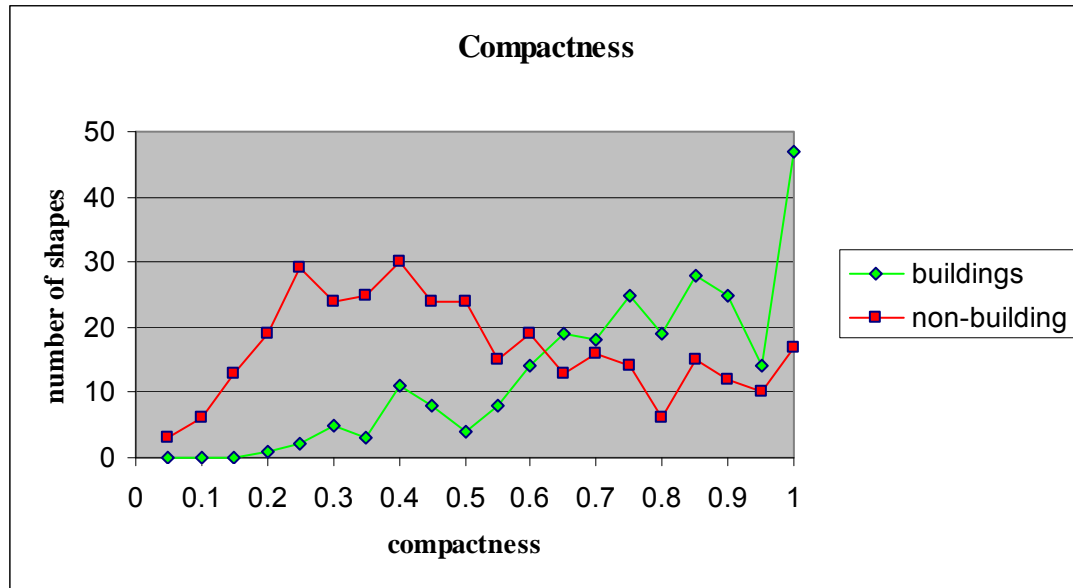
Figure 52 – Compactness values distribution for building and non-building segments

The compactness measure indicates that very high values (0.85-1) represent building features with greater probability. Compactness values between 0.6 and 0.9 have a higher probability of representing a building. The graph shows a growing trend for values between 0.5-0.6 for building features even though it is below the intersection point between the two graphs. Compactness values lower then 50 show a higher probability of representing a non-building feature.

Based on the values of the different indices and the trends of the graphs (figure 45 to figure 52), each range of values received a probability value. The probability values for all indices were combined to one confidence value that reflects whether the segment is a building or not. A script used to calculate the confidence is provided in appendix C.

4.1.6   Raster to Vector and Generalization

The result of the image segmentation is a raster image containing different values, each represents a segment. The segments have to be geometrically analyzed and converted to vectors as a final result. The vector polygons can be then analyzed by creating the convex hull, calculating the perimeter, area etc.

The raster to vector conversion is achieved by the following steps:

Edge detection – locating the edge of each segment, including the islands. Since the segment has only one value, the edge detection is a fairly straight forward process.

Calculating the convex hull for each segment (see appendix B)

Generating the exterior ring of the segment. Starting with a convex hull point and adding connected points ensures that islands are not part of the exterior ring.

Generalization – The goal of the generalization is to eliminate as many points as possible from the exterior ring and still capture the essence of the feature. Essence of the feature means that for example, small intrusions and protrusions may be eliminated, but the critical defining points of the shape should be maintained. Figure 53 illustrates a result of the generalization process. The extracted segment depicted in white, the convex hull of the segment is painted in red and the generalized polygon in yellow. Obviously, the defining corner of the building that was not part of the convex hull, but was included in the generalized polygon.
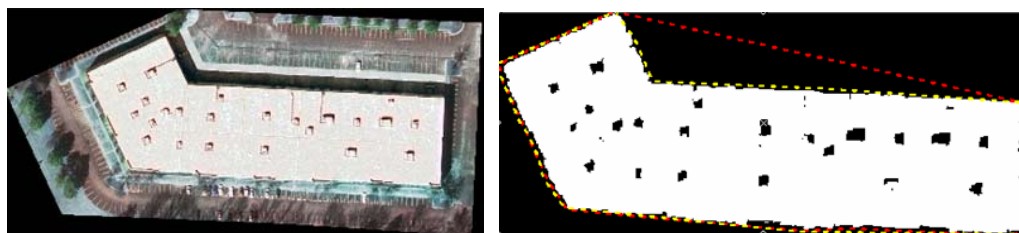
Figure 53 – Generalization example. Left – Original image. Right – extracted segment (white) convex hull (red) generalized polygon (yellow)

First, the convex hull and the corresponding exterior ring are located. Then, based on angular thresholds (between 60-120) followed by distance criteria (greater then 10 percent of the line length), the process iteratively adds points from the exterior ring between every two nodes of the convex hull. The integration of both angular and distance criteria for generalization of automatically extracted polygon has been examined in prior research projects. Lee and Shan (2002) illustrate the advantage of using both criteria for IKONOS imagery (see Table 1) extracted buildings. The authors use the Douglas-Peucker (1973), distance based methodology to eliminate redundant and extra points that cause a "zig-zag" affect of the line. An angular criterion follows the distance criterion to remove "unrealistic sharp changes" of the polygon. The authors conclude that the combination of both methodologies provides an effective simplification tool.

In this project, instead of eliminating points from the exterior ring, points are added from the exterior ring to the convex hull. To determine the drawbacks and advantages of each criterion (angular and distance), the angular threshold and the distance threshold were tested individually, including a fixed threshold for the distance (10ft). A fixed threshold has inherent drawbacks due to the wide variety of lengths and sizes of buildings. A 5 ft intrusion may be an important characteristic of a 20 ft building side, but

not as important on a 100 ft side. Between every two nodes of the convex hull, points are examined for the angles. Based on initial visual evaluation of buildings in the images and testing automatically extracted segments, an angle threshold of 90 degrees +/- 30 degrees (60-120) was determined. The threshold is a result of both extraction artifacts and legitimate building shapes (figure 53). The distance criterion is secondary to the angular criteria. A point within a distance greater then 10 percent of the total distance between the two nodes may be considered for the distance criteria. For buildings outline generalization, methodology that only uses distance criterion is inferior to the angle criteria, since the angle criteria manages to eliminate more points of the exterior ring, while keeping the critical turning points and corners on the ring. The entire process is iterative in a sense that it begins with two nodes of the convex hull and examines exterior ring points located between them. Once a point was selected to be part of the generalized polygon, it replaces the previous point (a convex hull node or a previously selected point). The process continues until it reaches the second convex hull point.
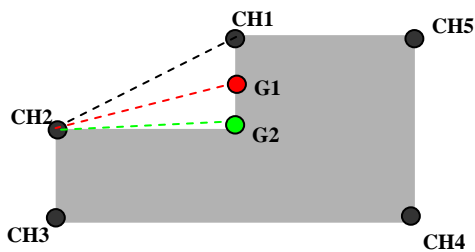
Figure 54 – Generalization process. Convex hull points in black; Intermediate result in red; final result in green.

Figure 54 illustrates the generalization process. The initial convex hull result is depicted by points CH1-CH5. A direction is determined for the convex hull ring and the

102

exterior ring (all the points on the edge of the shape) are sorted within the same direction. As a result, every two nodes of the convex hull have an attached, known list of points between them on the exterior ring. In this example, the direction is CH1-CH2-CH3-CH4-CH5. Based on the angle created by CH1-G1-CH2 (about 110 degrees) and the distance between G1 and the CH1-CH2 line, point G1 is added between the nodes CH1 and CH2. Point G1 replaces point CH1, and the search for the next point continues until we reach point CH2. One drawback of adding the distance criterion is that point G2 may not necessarily be placed exactly on the corner. Derived from the distance criterion, the point may be placed close to the corner within the defined distance. It should be noted that a slightly different result may be received based on the direction of the points (clockwise/counter clock wise) due to the nature of point-adding process.
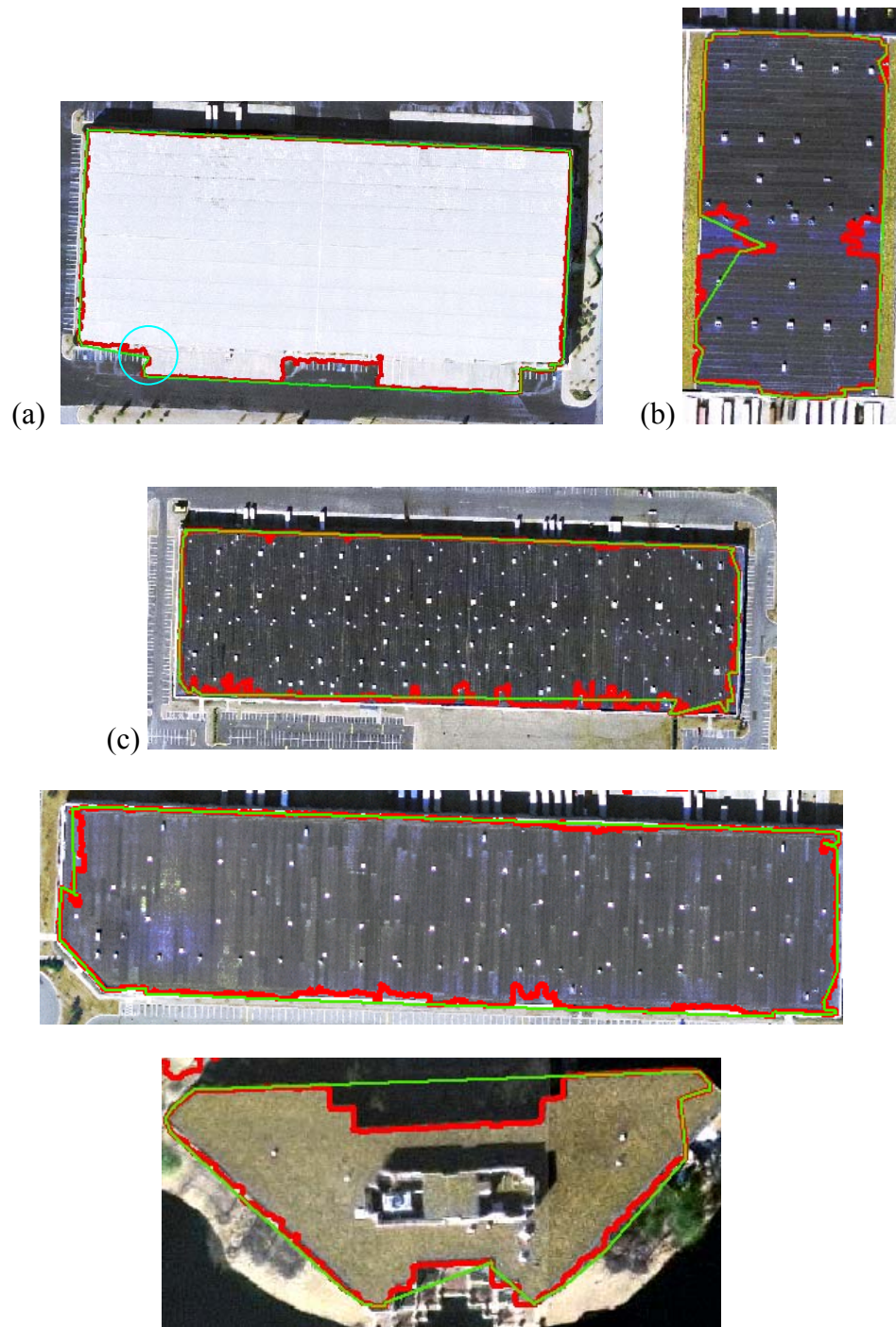
Figure 55 – Generalization results overlaid on the image. Exterior ring highlighted in Red; Generalized polygon – highlighted in green.

Figure 55 illustrates several examples of the generalization process, successful results as well as artifacts of the process. Section (a) depicts a relatively large building with a

small intrusion. The small intrusion is eliminated due to the distance criterion. The corner highlighted shows a corner that is not located exactly on the corner point, but very close to it, maintaining the essence of the shape. Section (b) illustrates a typical problem of an automatic extraction process. The generalization result, due to the intrusion on the left side, does not include a section of the building on the left side. Sections (c) and (d) illustrate common rectangular and near-rectangular building examples. The generalization result is a successful result that manages to dramatically reduce the number of points on the outline. Section (e) shows a less common building shape. The exterior ring defines the outline in great details and is fairly accurate. The generalization result eliminates the upper intrusion and does not exactly follow the outline at the bottom of the shape. As depicted in all examples in figure 55, the generalization result heavily relies on the convex hull result. Artifacts of the automatic process may add or remove points to the convex hull location, hence, may have a direct affect on the final result. The angular and distance criteria may eliminate or keep parts of the buildings and should be re-considered on a project-purpose basis (need the essence of the building; exactly the outline; eliminate small intrusion etc).

4.2   Result Evaluation

The methodology was tested on aerial imagery acquired over Memphis, TN.   Three different types of buildings (commercial, residential, high-rise) were tested, each with intrinsic, distinct characteristics. Due to the inherent nature of each type of buildings, a separate section is dedicated to commercial (4.2.2), residential (4.2.3) and high-rise (4.2.4) buildings. More general topics that are common to all types of buildings are elaborated in sections 4.2.5 to 4.2.8. Section 4.2.9 provides a comparison between an automatic process result and manual digitization of buildings.   Two different methodologies for manual digitization are provided and quantitatively evaluated.

4.2.1   Memphis Test-Bed

The Memphis test-bed area includes two data sets: aerial imagery and tax assessor parcel data set.

The aerial imagery data sets include ortho-rectified image mosaics. Table 3 provides specific details about the imagery as retrieved from the metadata files.

Table 3 - Aerial imagery over Memphis - metadata

Image source (Origin Agency)	National Geospatial-Intelligence Agency (NGA), U.S. Geological Survey (USGS)
Projected coordinate System	UTM, NAD83 Datum
Camera	Leica ADS40 Digital Camera System ISTAR Digital Image Processing Software Z/I Imaging Orthopro 4.0 Windows NT/2000 Systems
Spatial Resolution	0.3m (approximately 1ft)
Spectral Resolution	3 bands, natural color image (RGB)
Single image area	1500 meter
Acquisition date	Feb 2004
Image Original Format	GeoTiff
Mosaic ground Area	9 x 4.5 km
Design accuracy	2.12m for X and Y

The Shelby county tax assessor database, compiled in 2004, contains multiple data sets, including parcels and a building inventory. The database contains 346,393 land parcels. There are a total of 306,003 dwellings within 291,552 parcels, each parcel containing one or more built structures. An extract of parcels and buildings points from the Shelby county tax assessor database is illustrated in figure 56. The building inventory includes a point at the centroid of each parcel polygon. The point layer was compiled

from the tax assessor database and is linked to one or multiple buildings, based on the parcel ID. Each parcel has a unique parcel identifier, designated as the primary key in the parcel database. The attribute schemes for the parcel and building inventory are provided in appendix J.

Figure 56 – Extract of the parcel and building datasets in Memphis, TN

The attribute data includes zoning information that was used to select subsets of parcels for testing. The following three sections provide testing result for each type of building. Initial subset of parcels is selected, followed by building extraction results and discussion.

In order to quantitatively evaluate the result of the extraction process, a digitized building polygon layer was used as ground truth. The building layer was provided by the Shelby County, TN. The digitized-ground truth data set was used as the reference layer

compared with the extraction result building layer. A quantitative evaluation calculates the discrepancy between the layers and allows drawing conclusions regarding the accuracy and efficiency of the extraction process. As part of the evaluation, two types of errors were defined, Type 1 and Type 2. Type 1 error represents the number of segments that were extracted but are not buildings. Type 2 error represents the number of buildings that were not identified. Those buildings were either not extracted or extracted and erroneously eliminated. The two types of errors define very distinct results of the extraction process and may entail different manual post-processing. The time, effort and overall cost of the post processing can be evaluated differently for the two errors. Moreover, the importance of Type 1 and Type 2 errors may vary by application. The nature of the application may derive specific quality measurements for the extraction process. For example, an application may state that it is crucial not to identify features that are not buildings as legitimate structures, therefore allow a larger number of Type 2 errors while minimizing Type 1 errors. Other application may be more concerned with actual extraction of as many buildings from the imagery and with the accuracy of the extracted features. In this case, a larger number of Type 1 errors may be acceptable. The importance of the errors also depends on the amount of effort and cost as defined for the post processing. Limited or no post processing may dictate a distinct extraction approach that differs from other, less restrictive applications. Each type of error requires a distinct set of manual operations. Type 1 errors require simple one delete operation while Type 2 errors entail digitizing the entire building. As a result, each error may have a different weight attached to reflect the importance of the error during the post processing. Since

this project is not aimed at a specific application, both errors were evaluated with the same weight.

Besides the errors, a partial extraction result describes segments that represent buildings, but were extracted with an error of more then 50 percent of the building (for example, 45 percent of the building was extracted). Successful extraction represents the remainder of the segments, which cover 50 or more percent of the buildings. The average of the discrepancy for the buildings is also calculated. The average value is the calculated mean of undershoot and overshoot for all the buildings. The evaluation process was initially aimed for all building types, but adjustments have to be made as the results vary greatly between the distinct building types as shown in sections 4.2.2 (commercial), 4.2.3 (residential) and 4.2.4 (high-rise).

4.2.2 Commercial Parcels Testing

Most commercial and industrial buildings are located in the Memphis down town area. Three image mosaics that span over the downtown area were identified to include substantial number of commercial and industrial buildings for the testing. Initial selection of a subset of parcels was based on occupancy type. Parcels that have a built structure (not vacant) and have occupancy type of either commercial or office were selected. Within the three mosaic images, 1079 parcels were selected. Figure 57 provides an overlay of the selected parcels over the three image mosaics.
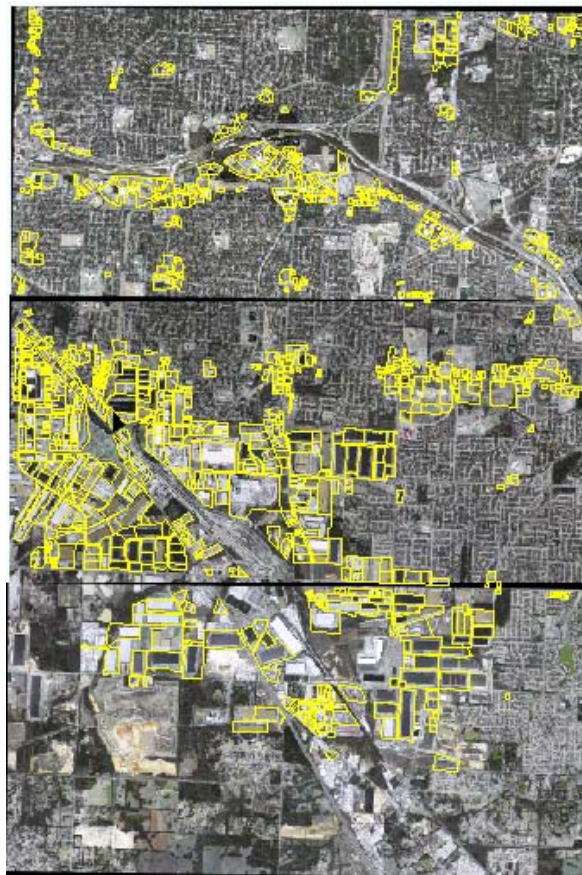
Figure 57 – Commercial parcels in downtown Memphis,TN overlaid on orthophoto images.

4.2.2.1   Testing Results

The result of the extraction process is a vector, polygon layer. The polygon layer was compared to the ground truth, building dataset. A quantitative evaluation of the result compared the vector layer with the building dataset. The discrepancy between the layers was calculated as follows:

Calculating the symmetric difference between the layers – First, union operation was performed on the polygon result and the building dataset. The result of the union operation is segments inside and outside the intersection of the two layers. The result segments that are outside the intersection of the two layers represent the overshot/undershoot areas. Each of those segments is linked (joined) to a building.

Calculate the overshoot/undershoot area for each building.

Summarize the undershoot/overshoot for every building.

Calculate the ratio between the undershoot/overshoot and the building

Table 4 presents the results for the commercial buildings testing. The testing parcel subset includes 1079 parcels that contain 1128 buildings. The result vector layer was compared with a building layer and quantitatively evaluated as illustrated in table 4. The "Method" column represents the sequence of processes as used for the testing. "All eliminate" means a full run that includes all three elimination steps: elimination by size, shadow and geometry characteristics. "No size" indicates that the elimination by size was not part of the process. Similarly, the methods "No Shadow" and "No Geometry"

were tested. Type 1 error represents the number of segments that were extracted but are not buildings. Type 2 error represents the number of buildings that were not identified. Those buildings were either not extracted or extracted and erroneously eliminated. Partial extraction represents segments that represent buildings, but were extracted with an error of more then 50 percent of the building (for example, 45 percent of the building was extracted). Success represents the remainder of the segments, which cover 50 or more percent of the buildings. The following three values were calculated in order to better evaluate the degree of error and discrepancy between the result and the ground truth building layer. The average of the over/undershoot for the buildings is also calculated. The average value is the calculated mean of undershoot and overshoot for all the buildings. Last, the number of buildings with an area error less the 10 and 15 percent is provided.
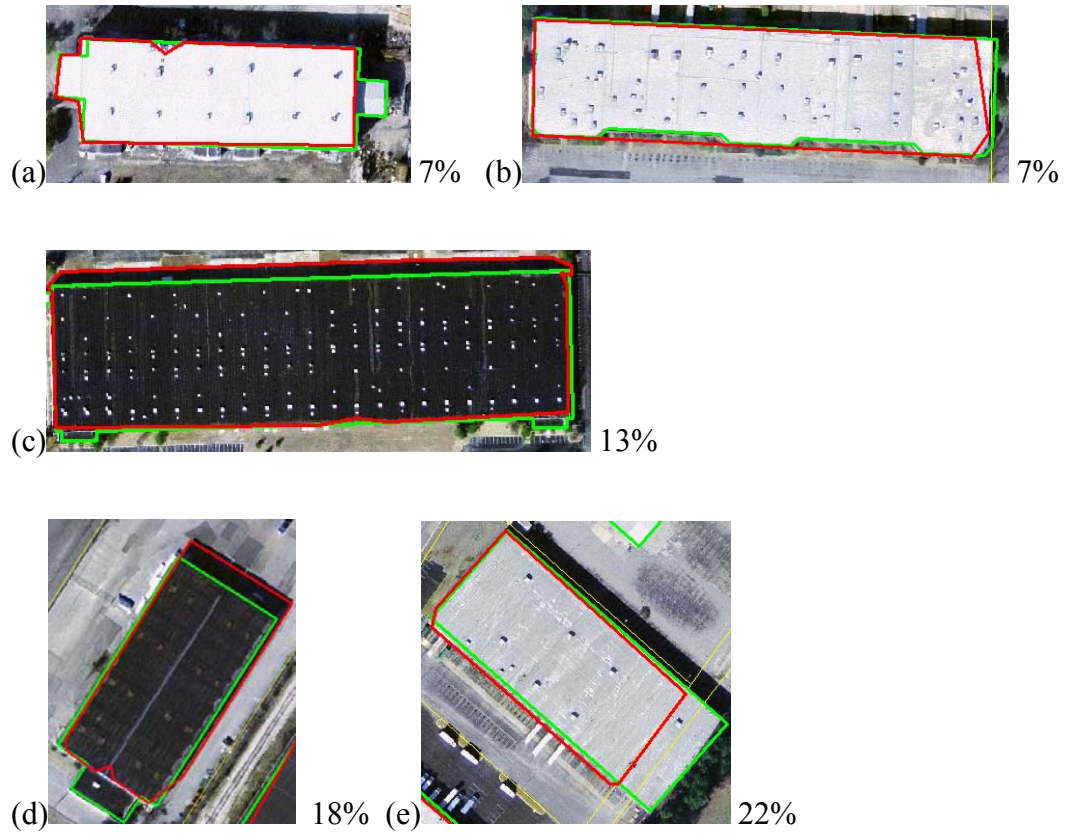
Figure 58 – Area discrepancy between automatic extraction result (red) and digitized building dataset (green).

Figure 58 provides several examples of extracted buildings and the calculated area discrepancy. Sections (a), (b) and (c) in figure 58 illustrate an area discrepancy within the range of 10 and 15 percent. Those examples demonstrate that buildings with area discrepancies within 10, 15 percent (and even close to 20 percent in section d) capture the building area in a manner that is acceptable for a variety of applications. Hence, an area error of 10 and 15 percent was selected and presented as an acceptable result for extracted buildings in table 4 below.

114

Table 4 - Commercial buildings testing result

Method	Type 1	Type 2	Partial	Success	Average over/under	<10%	<15%
All eliminate	96	169	23	936	9.7	732	887
No size elimination	213	110	47	975	10.01	748	908
No Shadow elimination	272	100	35	997	10.2	759	927
No Geometry elimination	1556	153	40	939	9.72	735	890

For the full run, 96 (about 10%) segments were flagged as Type 1 error, "false positive" identification, where a building was extracted but does not exist. 169 (about 15%) segments were flagged as Type 2 error, "false negative" identification, where the building exists but was not identified. Out of the 1128 buildings, 732 (about 64%) were extracted with under/overshoot percentage of not more then 10% and 887 (about 79%) with under/overshoot percentage of not more then 15%. In the second scenario ("No Size"), the process added 59 legitimate parts of buildings, but Type 1 errors increased to 213. In the third scenario ("No Shadow"), the process identified 69 additional legitimate parts of buildings, but Type 1 errors increased to 272 segments (for instance, rectangular parking lots tend to be identified as buildings). The final scenario ("No Geometry") identified 16 additional legitimate parts but received a significant increase in Type 1 errors. There are several interesting conclusions that may be drawn from the results, as

well as multiple questions to be raised. Clearly, the full run, that includes all elimination steps, increases the overall accuracy of extraction and reduces the amount of manual post-processing. The full run may have the lowest number of identified buildings, due to erroneous elimination of legitimate buildings by the elimination steps, but it has the lowest number of type 1 errors. That is significant since type 1 errors indicate the amount of post processing effort.

The increasing number of type 1 errors in the scenarios following the full run is reasonable. Since the segmentation step includes four different ranges (peaks) within the histogram, a parcel that contains 2 buildings with different colors, at least two non-building segments are expected. Those segments would be eliminated or appear as type 1 errors. The considerable difference between the number of type 1 errors between the "no size" (213), "no shadow" (272) and the "no geometry" (1556) can be attributed to the fact that the same segment may be eliminated in different steps. When the "no size" elimination step is not used, a segment may be eliminated in any of the other elimination procedures. The no geometry step is the final procedure in the analysis and indicates the number of segments that were not eliminated within the size and shadow elimination steps. The number of type 1 errors for the "no geometry" step, indicates that for the 1079 parcels, one or two segments in each parcel remained for the geometry analysis. Figure 59 provides examples of common segmentation results. The left images are the original parcel-sized images and the right is the segmentation result. Clearly, more then five segments are possible candidates that do not represent buildings. Hence, it is not unreasonable to have one or two remaining for the geometry elimination analysis.

116
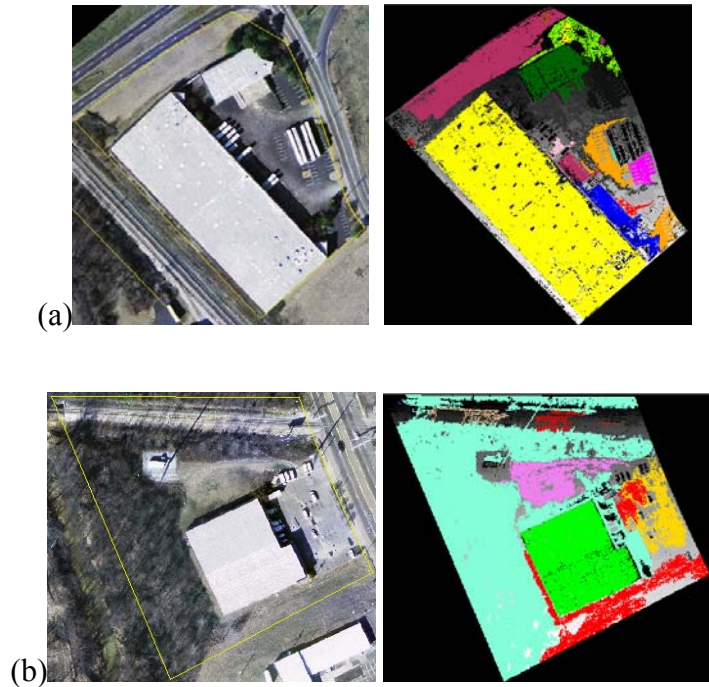
Figure 59 – (Left) Original parcel-sized image (Right) segmentation result

The examples in figure 58 show that segments that are not eliminated by the size or shadow eliminations (due to removing the elimination step or due to analysis confusion) may be eliminated by the geometry analysis. For example, the orange segment in the middle of section (a) or the pink segment in section (b). By eliminating segments that otherwise would have been eliminated in previous elimination steps, the geometry elimination step reduces the number of type 1 error segments for the "no size" and "no shadow" elimination. In addition, since the geometry elimination is the final step, it remains with the largest number of type 1 errors. The geometry analysis appears to be a significant step in the elimination analysis with considerable effect on the final result. The "no size" and "no shadow" analysis, maintain an important role in the overall analysis and should be integrated in the process to reduce the post processing effort.

4.2.2.2   Extraction Failure Factors

In table 4 (Section 4.2.2.1), 169 buildings are classified as type 2 errors, buildings that were not extracted in the full run process. 23 buildings were partially extracted using the same scenario. In order to evaluate the reasons for these extraction failures, each building was manually inspected and classified as one or more of the following categories: slope (sloped roof), complex (complex roof signature), size (small size relative to the parcel size), shadow (complicated shadow cast), compound (compound buildings).

Buildings that were not extracted were classified as follows (the same building may belong to more then one class):

- "slope" class –   26 buildings.

- "complex" class –   98 buildings

- "size" class –   26 buildings

- "shadow" class –   51 buildings

- "compound" class – 17 buildings

The 23 buildings that were partially extracted were classified as follows:

- "slope" class –   5 buildings.

- "complex" class –   18 buildings

- "size" class –   9 buildings

- "shadow" class –   2 buildings

"compound" class – 9 buildings

The results of the classification indicate that a significant number of buildings were classified as "complex" (having complex roof signature).   A set of buildings with complex roof signatures is provided in figure 60.
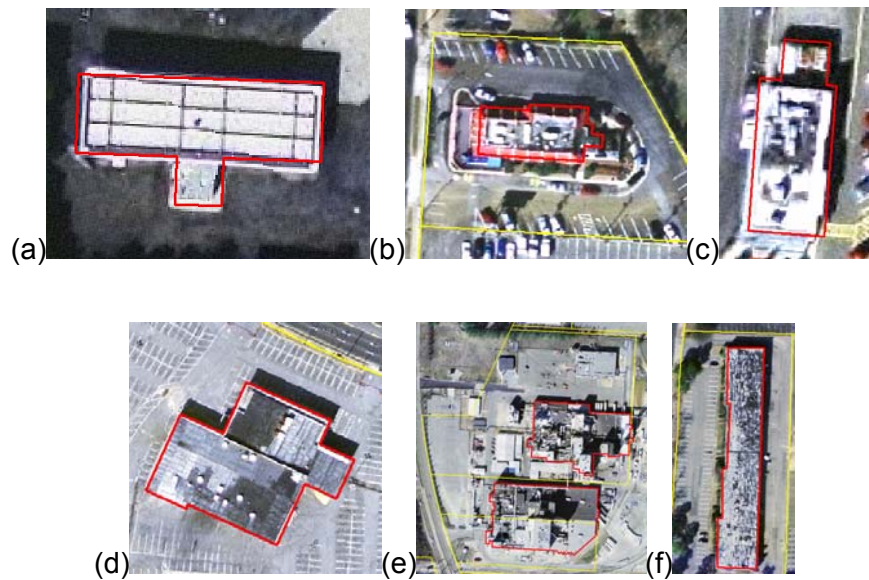
Figure 60 – Buildings with complex roof signature

The examples in figure 60 include multiple buildings with complex roof signatures. The complexity may be attributed to objects located on the roof (sections b and e), color patterns (section a), multiple level roof (section e) or simply variety of grey levels due to shadow, material discoloration etc.   The variety of grey levels on the roof prevents clustering of pixels into continuous segments that can be further analyzed as possible buildings.

A considerable number of buildings were classified as "shadow". Those building features have a shadow cast that obstructs the shadow analysis (4.1.4). Figure 61 provides three examples of extraction artifacts that interfere with the shadow analysis. The examples show a gap around the building that separates the extracted segment and the shadow. As a result, the shadow is not correlated with the roof segment, hence, eliminated during the shadow elimination step (the process analysis concludes that the segment has no shadow attached in the correct location).
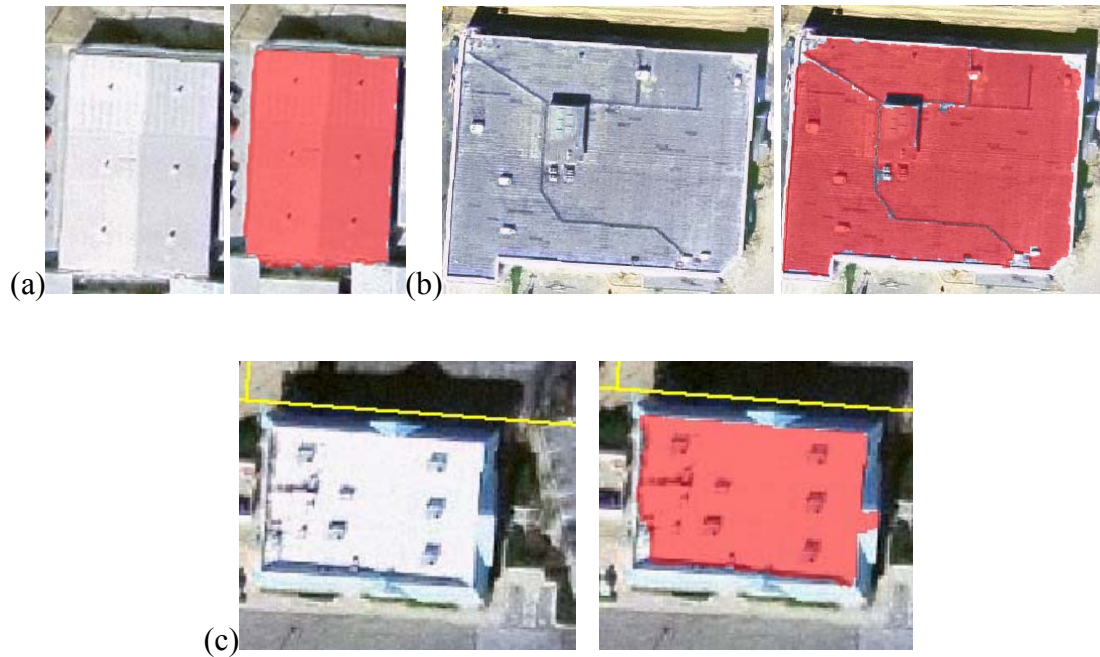
Figure 61 – Gap between the extracted building segment (red) and the shadow

Example (a) shows a gap between the segment and the shadow, created by a lower level roof on the building entrance. Example (b) shows a white belt around the roof of the building. In example (c) the roof shadow creates a lighter color shadow gap between the building and the shadow. All the above examples tremendously complicate the shadow elimination analysis.

The "size" class represents several scenarios. One scenario includes buildings that were eliminated since their area was small relative to the parcel area. Buildings extracted as multiple sections were eliminated when a section area was significantly different then the database area (the entire building area). There may be a discrepancy between the database record and the area extracted from the image. Extraction artifacts and analysis problems due to the size of the building and the database attribute are discussed in section 4.1.3.

There are various factors that might affect the result of the extraction process. Clearly, the nature of the roof signature has a significant and direct affect on the segmentation, and as a result, on the entire extraction process. Other contributing factors may be the size of the building relative to the parcel, shadow analysis, the slope of the roof and building compound architecture. Some factors can be further investigated and possibly resolved (size, shadow and compound) while some (complex signature) will remain an obstacle to any automatic extraction process.

4.2.2.3 Multi-building parcels

97 out of the 169 buildings (57%) that were flagged as "type 2" errors (full run scenario), are located in multi-buildings parcels. Multi-building parcels are defined as parcels that have 2 or more buildings. 69 out of the 97 buildings are located in parcels that have 3 or more buildings, and 42 are in parcels that have 4 or more buildings. Out of the 936 successfully extracted buildings, 163 (17.4%) are located in multi-building parcels. 42 are located in parcels that have 3 or more buildings and 17 in parcels with more then 4 buildings. Since multi-building commercial parcels are common, this

section presents statistics that examines whether multiple buildings within a parcel lowers the probability for successful extraction.

Total of 260 buildings are located in multi-building parcels. 37% of those buildings were not extracted. Those buildings comprise of 57.4% of the buildings that were not extracted. 17% of the extracted buildings are located in multi-buildings parcels. Clearly, the failed extraction rate of buildings within multi-building parcels (37%) is much higher then the overall extraction rate (15% type 2 error). Buildings that reside in multi-buildings parcels cover a smaller area of the parcel and may become as a result less dominant features in the parcel.

The buildings that were not extracted within the multi-building parcels maintain a mean area ratio between the building and parcel of 4.7%. The actual average values range between 0.1-29.5%, with 76 out of 97 have a ratio of less then 5%. 62 out of the 97 buildings are at least half the size of the average building area within the parcels.

163 buildings were identified and successfully extracted within the multi-building parcels. 15 buildings have a ratio (building area divided by the parcel area) of less then 5% with a mean ratio value of 35.1%. 155 buildings are at least half the size of the average buildings area within the parcel.

One of the fundamental assumptions made in the project is that the building is a dominant feature within the localized parcel area. A small, insignificant feature would not have a considerable presence in the image histogram and would not be easily identified. When only one building resides in a parcel, it usually covers a large area within the parcel. That is obviously not the case when there are multiple buildings within

the parcel. Multi-building parcels may reduce the overall extraction rate of the buildings. As a result, it is suggested to evaluate the mean ratio between the buildings and the parcels area and adjust parameters for the size elimination accordingly.

4.2.2.4 Parcel-sized images

The extraction process limits the search for the building area by initial image partitioning. Cutting the image using the parcel layer assumes that the building is located within the parcel boundaries. Clearly, inconsistencies might occur when there is a significant time gap between the compilation of the parcel layer and acquiring the images. The images and the parcels layer used in this project were both captured and compiled in 2004. Some inconsistencies and discrepancies are expected between two independently created datasets. For a parcel layer and an image, the discrepancy might be represented as a parcel boundary that intersects a building outline. An evaluation of the consistency of such artifacts is presented below.

In order to evaluate the phenomenon, a spatial selection is performed between the parcel layer and the digitized building outlines. The building layer contains the 1128 buildings used for the testing (see 4.2.2.1) within the 1079 parcels. The result of the spatial selection is a subset of buildings that intersect with the parcel boundary.

Out of the 1128 buildings, 170 buildings (about 15%) outlines intersect with the parcels boundaries. 45 out of the 170 are compounds (figure 62) that comprise of several sections and reside within multiple parcels. Compound structures can be separately extracted within multiple parcels.

Figure 62 – compound buildings residing in multiple parcels (yellow lines represent parcels)

The 45 compound buildings include 11 "OFFICE CONDO" structures. Those office buildings reside in multiple parcels and cover almost the entire parcel (Figure 63). Office structures can be described as the union area of the office parcels.

Figure 63 – office condo structure divided between multiple parcels

An automatic process may not easily extract the multiple sections of the same structure since each section individually does not necessarily resemble a building segment. As a result, it is recommended to initially merge office parcels that are touching the boundaries.

Besides the compound structures, 125 buildings have sections outside the parcels boundaries. For those structures, the percentage of the area that is located outside the parcel was calculated. Since the process confines the extraction to the area inside the parcel, significant sections outside the parcel may clearly compromise the extraction

process. In order to calculate the percentage, the total area that resides outside the parcel was divided by the total area of the building. The following results were calculated:

- 62 buildings have 2 % or less outside the parcel boundary.

- 84 buildings have 5 % or less outside the parcel boundary.

- 100 buildings have 10 % or less outside the parcel boundary.

- 111 buildings have 25 % or less outside the parcel boundary.

From the percentage calculation, about 50 percent of the buildings have minor or insignificant portions of the structure ($<=2\%$) outside the parcel. The vast majority of those buildings (80-90%) can be extracted successfully (10-15% error) or at least partially extracted (50% or more of the building is extracted). Inspection of the extraction result shows that out of the 170 buildings that lie partially outside the parcel boundaries, 165 were identified and extracted. 154 structures were extracted successfully (area error $<=15\%$) and 11 structures extracted partially (area error $<= 50\%$). The remaining 5 structures that were not extracted, failed due to complex signature of the roof top (See Figure 64)
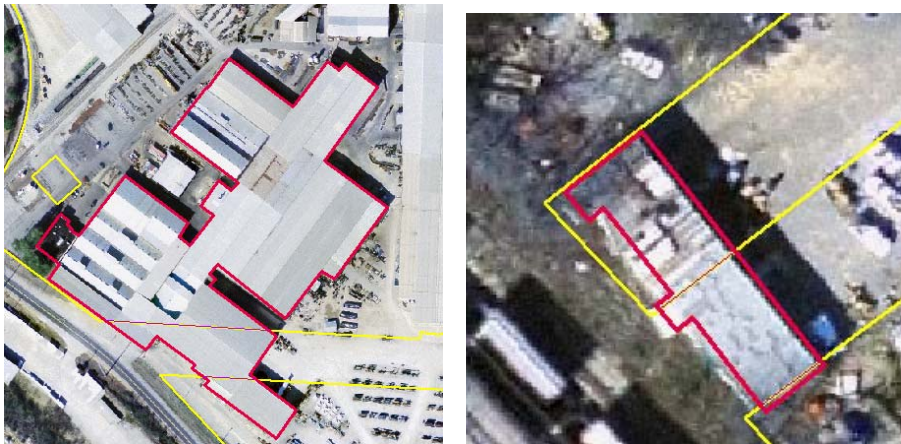
Figure 64 – Complex roof signature. Buildings reside in multiple parcels (yellow)

Parcels, as already indicated, can dramatically simplify the extraction process. From the evaluation presented in this section it is clear that the majority of buildings (85%) reside within the parcels boundaries. Compound structures can be extracted in parts within multiple parcels. Other buildings that cross between boundary lines, maintain, for most cases, a significant portion of the building within the parcel area. As a final note it can be concluded that the parcels layer can be used to localize the search for the buildings by dividing the image. The loss of information is secondary to the obvious benefit to the extraction process and can be further reduced by applying a buffer around the parcels. The actual buffer size should be determined for each parcel individually based on the parcel size and the possible expected structure.

<u>4.2.3</u>   <u>Residential Parcels Testing</u>

As mentioned before, the testing was performed on three types of structures: commercial, residential and high-rises. Each type of structure has specific functionalities and exists in a different environment. The structure characteristics and the environment have a direct affect on the extraction process. For example, trees around the building are a prominent problem for residential buildings extraction versus high-rise buildings. This section provides a discussion regarding the challenges of extracting residential buildings footprints. Figure 65 provides two examples of residential single family houses (a) and apartment complex (b). In both examples, trees and shadow casts obstruct the view of the buildings from above.

(a)   (b)

Figure 65 – Residential parcels (yellow) overlaid on 1ft image

Testing, similar to the commercial testing (section 4.2.2), was performed for residential parcels. The testing includes 2580 residential parcels in the same image mosaics. The outcome of the testing can be regarded as poor results. Only 829 buildings out of the 2580 were detected and partially extracted. The mean overshoot and

127

undershoot for the residential buildings was calculated as 55% with 92% of the houses having an overshoot/ undershoot greater then 50%. Tree occlusion, sloped rooftops and townhouses that span multiple parcels were all contributing factors to the results.

4.2.3.1 Residential Parcels Analysis

Clearly, an analysis of the core problem is required to understand and possibly suggest better extraction methodology. Since the entire methodology is based on the result of the initial segmentation, a further investigation of segmentation methodologies was performed. In order to assess the result of the initial segmentation, two areas were selected. One area (Figure 66 section a) is characterized by considerable number of trees, while the other (Figure 66 section b) does not seem to have as much vegetation around the buildings. Within each area, 100 parcels were randomly selected as depicted in figure 66.

(a)　　　　　　　　　　　(b)
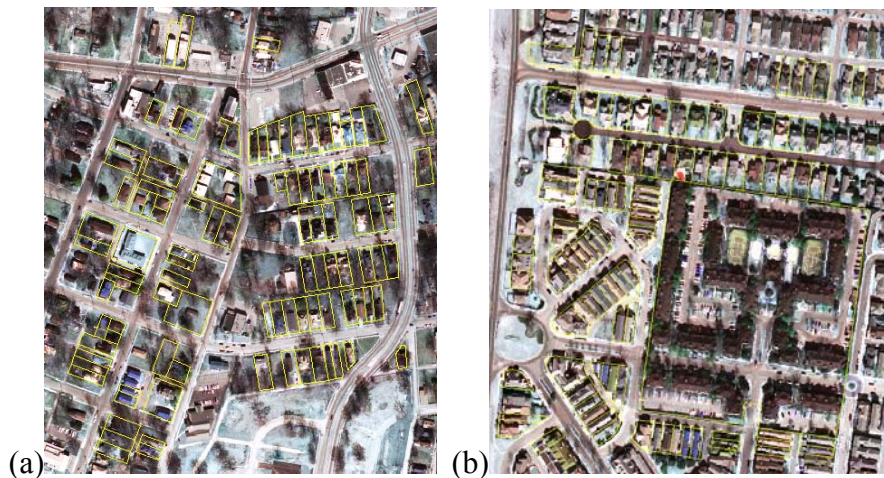
Figure 66 – 100 parcels overlaid on residential area (a) Area characterized by many trees

(b) Area with little or no vegetation

The segmentation assessment includes a comparison of two methods: the segmentation as implemented in the project using the histogram peak analysis and the Isodata classification method (section 2.3.1). The Isodata, unsupervised classification algorithm, requires several input parameters, including the number of classes. A careful examination of multiple parcels revealed that several classes are in particular dominant within a residential parcel. Those classes include the built structure, driveway, shadow and vegetation.

The result of the testing includes two segmentation images for each parcel. One image is the result of the histogram analysis and the other is the result of the Isodata algorithm. All images were manually examined to evaluate which segmentation method provides a better result. Table 5 presents the result of the segmentation methods. Test 1 column provides the number of buildings within the more vegetated area. Test2 was performed in a less vegetated area (significantly less trees around the buildings).

Table 5 - Segmentation Result of Residential Parcels

	Test1	Test2
Vacant	4	0
Histogram analysis – better	6	5
Isodata – better result	30	50
Isodata – better result but	26	30
Only Isodata	5	0
No method	29	15

The last row in table 5 indicates how many buildings were not segmented properly by any method. 29 buildings were not extracted within the more vegetated area and 15 buildings within the less vegetated area. In both scenes, the main reason for the failure was the trees. Deciduous trees obscure parts of the house or complicate the roof signature with branches or with shadow casts. Since test2 area does not include as many trees as test1 area, more buildings were adequately segmented (50 versus 30). The Isodata algorithm seems to perform slightly better then the histogram analysis on the residential parcels. It may indicate a relatively Gaussian distribution within the classes, since the algorithm assumes normal class distribution. The better segmentation result for 50 percent of the houses is not sufficient for an automatic extraction. The process will not yield a satisfactory result with 35% (test1) or 50% (test2) segmentation rate. The consistent poor segmentation result for both methods emphasizes the complexity of the residential parcel scene.

Following the automatic segmentation results, that were not found to be adequate for building extraction, a semi-automatic approach was investigated. Section 4.2.3.2 below elaborates on the algorithm and presents testing results.

4.2.3.2 Region Growing Algorithm

Region growing is a semi-automatic image segmentation procedure (Tremeau, and Borel, 1997; Haralick and Linda G. Shapiro, 1992). Region growing techniques have been used for feature extraction, including buildings (Li, Wang and Ding, 2006; Müller and Zaum, 2005; Ohlhof, et al, 2004). The algorithm is based on initial selection of a seed point. Every pixel connected to the seed is examined and added to the region to create a continuous, larger region. A pixel that is added becomes a seed, thus expands

the growing to its neighbors. The decision whether to add a pixel to the region is based on spectral distance between the seed and the pixel, as defined by the user. In many applications, such as ERDAS-IMAGINE, the user can define the extent of the neighborhood (4 or 8 connected pixels), the Spectral Euclidean Distance (the distance between the value of the pixel and the mean value of the seed) and the maximum number of pixels allowed in the region (Figure 67).
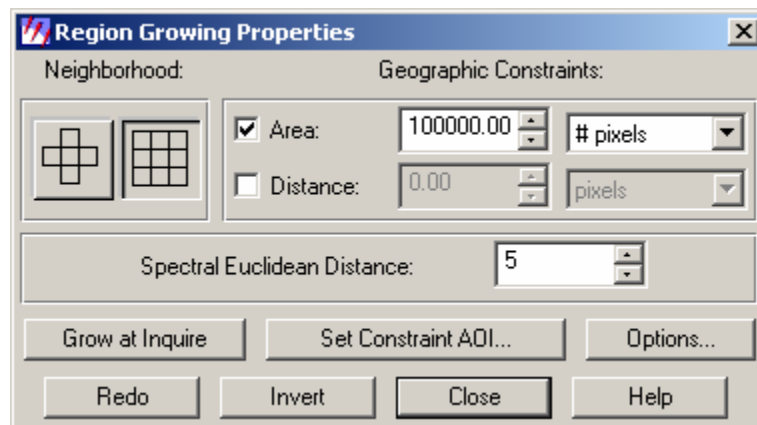
Figure 67 – Region Growing GUI in ERDAS-IMAGINE

The region growing algorithm result is highly depended on the parameters settings as shown in examples below (Figure 68).
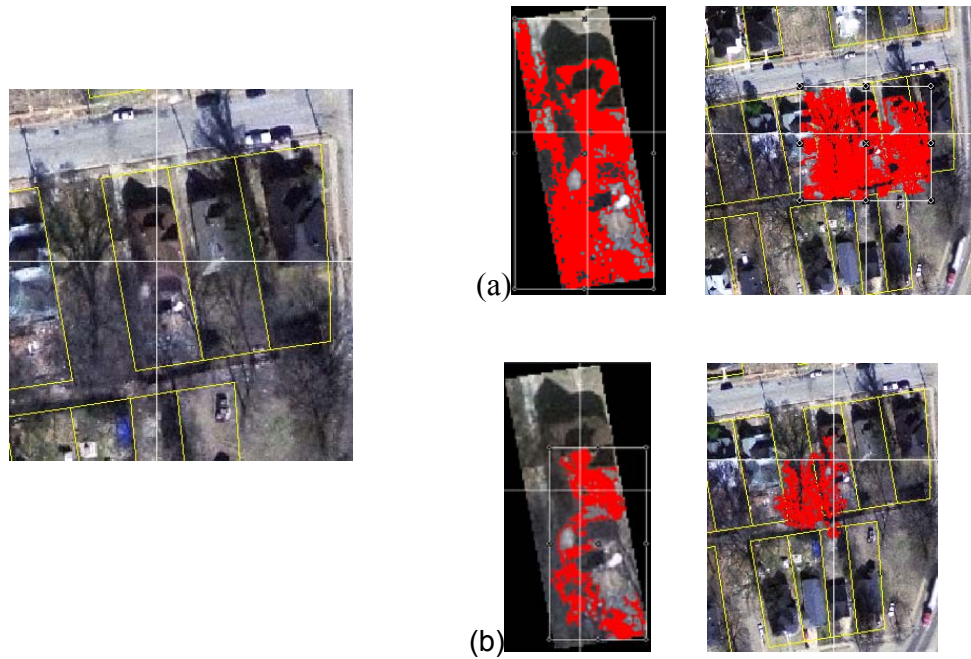
Figure 68 – Region Growing Result. (left) Original Image (Right) (a) Spectral Distance = 50

(b) Spectral distance=20

Figure 68 illustrates the different outcome of the region growing process for different input spectral distances. The left image shows the original image of the building and the overlaid parcels layer (yellow). On the right there are two sections of images. Section (a) illustrates the result for a spectral distance of 50. The growing algorithm was examined on both the parcel-sized image (on the left) and the entire image (on the right). Clearly, the pixels extend the parcel boundary lines when tested on the entire image. Section (b) was similarly tested for a spectral distance of 20.

Region growing performs adequately in relatively simple scenarios as indicated in figure 69.
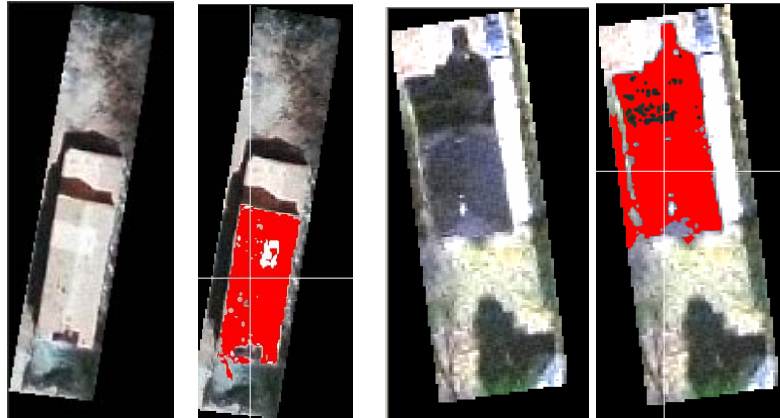
Figure 69 – Region growing segmentation result (Spectral distance =50)

The region growing procedure was tested on 100 residential parcels (test2 area with spectral distance of 50). In 60 out of the 100 parcels, the region growing result expands beyond the actual building. If the building is confined within the parcel area, the region can only grow inside the parcel. Clearly, when the region grows outside the parcel, many pixels that append the building region should be removed. The region growing results emphasize the advantage of using the parcel region rather then the entire image for the extraction.

The three segmentation algorithms (Histogram analysis; Isodata; Region growing) were used to compare segmentation for several residential parcels scenarios. The comparison and evaluation are provided in the following section.

4.2.3.3   *Three algorithms segmentation testing*

In order to extract buildings or any other feature from images, an initial segmentation of the image is required. As indicated by the results in section 4.2.3.1, residential parcels present great challenges for the initial segmentation. There is clearly high level of complexity and segmentation confusion that requires human intervention. The two

automatic segmentation procedures (Histogram Analysis; Isodata) as well as the semi-automatic procedure (Region growing) were further examined. The performance of the three procedures was evaluated for several residential parcels scenarios. The specific scenarios were identified to provide representation of the major obstacles during the segmentation process. The comparison may provide a better understanding and selection of the preferred procedure for a specific scenario.

The various scenarios identify multiple sections of the structure, shadows, trees and the sun illumination (1 or more illumination sides). The scenarios were tested and divided into "successful" scenarios and "failed" scenarios. The successful scenarios are:

- Two building sections; insignificant shadows on the roof; no trees occlusion; two sides of illumination.

- Two building sections; shadows; no trees occlusion, one side illumination

- Two building sections; shadows; no trees; one side illumination.

- Several building; insignificant shadows on the roof (slope roof); trees occlusions on one building; 1 side illumination.

- Mainly single section building; insignificant shadows on the roof; no trees occlusion; two sides illumination

The failed scenarios are:

- Two building sections; significant shadows on the roof; trees occlusion, one side illumination

- Single section building; shadows; slight trees branches occlusion; one side illumination

- Two building sections; shadows; trees occlusion; two side illumination.

- Several slope building; shadows; no trees; two side illumination

Each scenario is illustrated below and examined using the three segmentation algorithms.

Scenario 1: Two building sections; insignificant shadows on the roof; no trees occlusion; two sides of illumination
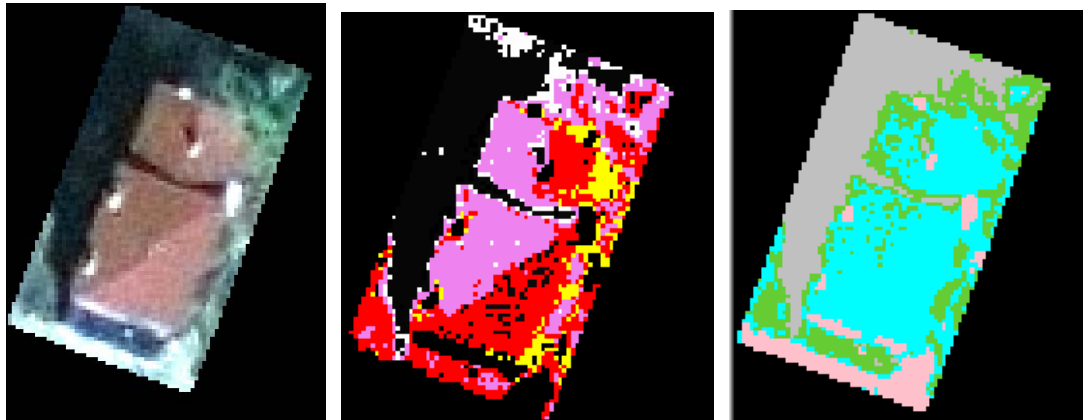
Figure 70 – scenario 1 [original image] [Histogram analysis Result] [Isodata Result].
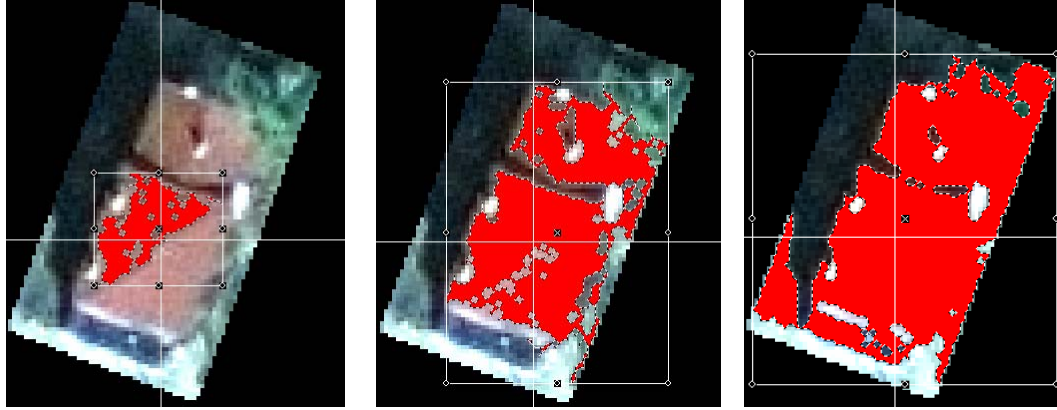
Figure 71 - scenario 1: region growing result (spectral distance = 20, 50, 100).

Figure 69 illustrates the original parcel image and the result of the automatic segmentation procedures. The image shows a fairly simple scene with limited shadow casts on the roof, no trees and a tilted roof. The histogram analysis result provides an adequate segmentation of only one side of the roof. The Isodata algorithm provides a better segmentation that appends parts of the surroundings to the building segment. Figure 71 provides results of the region growing algorithm for three spectral distances (the seed point is the intersection point of the vertical and horizontal lines). It is clear that spectral distance of 50 provides the best representation of the building including small portions of the surroundings and gaps within the segment. The building region may be partially extracted using the Isodata or the region growing procedures, but requires user post-processing corrections.

The histogram plots for the image are shown in figure 72. Compared to histogram plots of commercial parcels (section 4.1.2.1), the residential parcels histograms are not as "smooth" and there is less evident correlation between the bands (mostly between the red and the other bands).
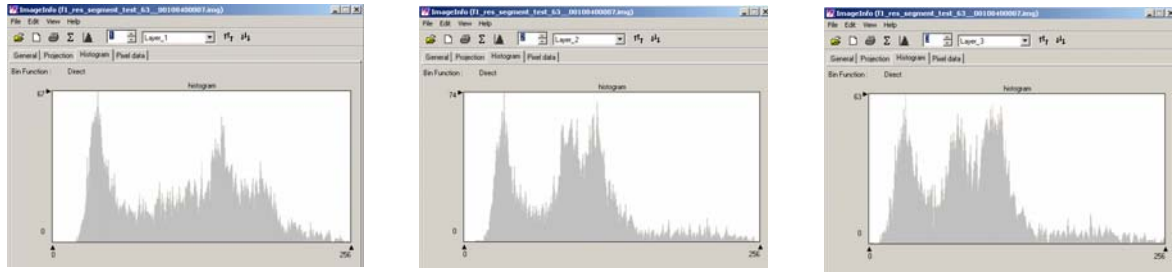
Figure 72 – scenario1. Histogram plots of the (left to right) Red, Green, Blue bands

Since the histogram plot is less "smooth", it is more complicated to automatically analyze the histogram and accurately allocate the different peaks and ranges.   That directly affects the segmentation of the corresponding objects in the image. The reduced correlation between the bands can be attributed to the dominant presence of vegetation and its representation within the bands histogram.

The segmentation result of the Isodata algorithm can be examined using feature space plots (figure 72). Feature space plots show the spectral location of object pixels on a two bands plot. Each axis represents a specific band. The plot illustrates the correlation between the bands and an overlay of the segmented classes. A higher correlation between the bands is represented by a near-line plot. Each ellipse represents a class as segmented in the image. The overlap between the ellipses may illustrate the level of confusion between classes.

137
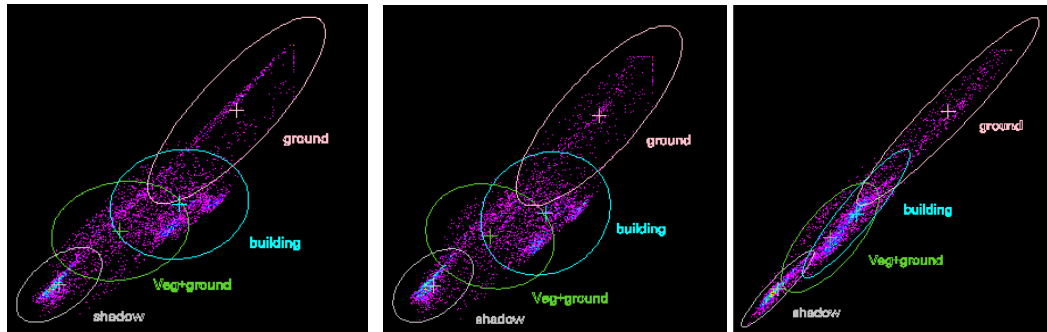
Figure 73 – scenario 1. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The colors of the classes in the feature space plot correspond to the colors of the Isodata result in figure 70. As indicated by the histogram plots in figure 72, bands 2 and 3 are highly correlated while the other combinations are more scattered. The classes representing the shadow and the ground are on the two far ends of the plots, have relatively small overlap with other classes and hence are segmented fairly accurately. Moreover, there is an evident concentration of pixels within the building ellipse area (cyan colored pixels) that indicate a significant number of pixels that represent that class (The same can be seen for the shadow class that is fairly dominant in the image). The building (cyan) and the vegetation (green) have greater overlap, are closer on the spectral plot and as a result produce segmentation confusion. The result is a building segment that includes vegetation and ground pixels, and vegetation segment that includes building and ground pixels.

The histogram and the feature space plot indicate that it is possible to define fairly distinct classes within the parcel for scenario 1. As with any segmentation procedure, some confusion and classification errors occur. The fairly simple scenario provides the opportunity to segment the building area. The Isodata and the region growing perform better but may require manual post-processing.

138

Scenario 2: Two building sections; shadows; no trees occlusion, one side illumination
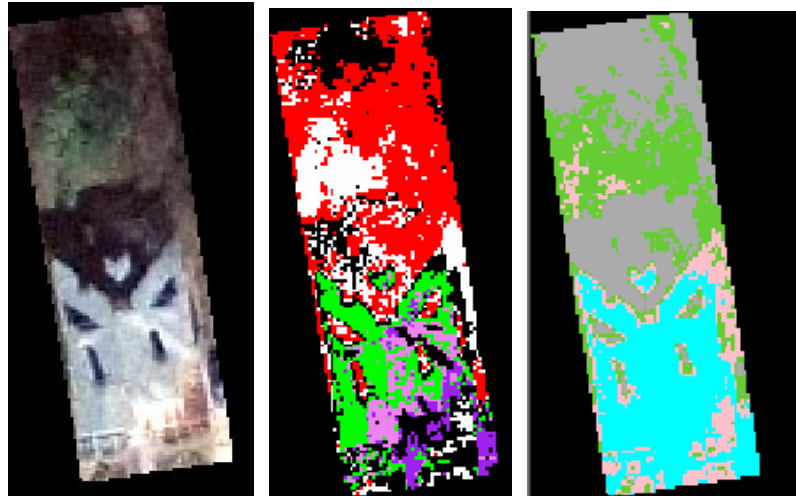
Figure 74– scenario 2 [original image] [Histogram analysis Result] [Isodata Result].
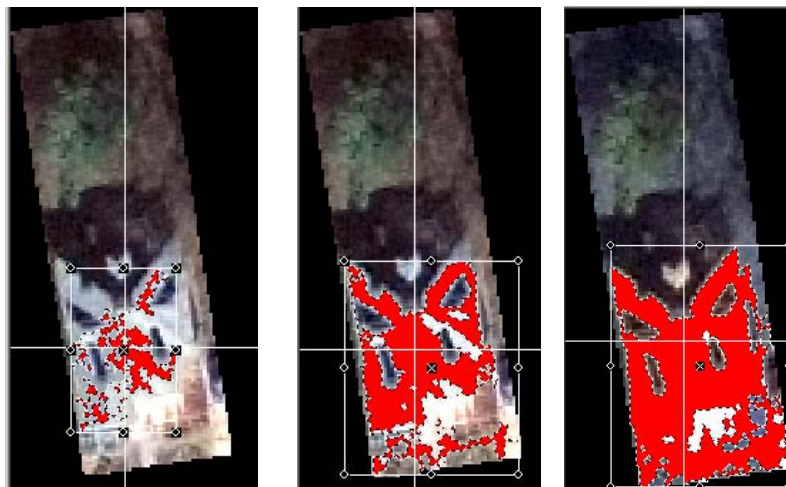
Figure 75 - scenario 2: region growing result (spectral distance = 20, 50, 100).

Figure 73 illustrates the original parcel image and the result of the automatic segmentation procedures. The image shows a scene with multiple shadow casts on the roof and no tree occlusion. The shadows are a major obstacle for the segmentation as they create a completely different grey level representation for several sections of the

house. The histogram analysis result provides a poor segmentation while the Isodata algorithm provides a better segmentation of the "no-shadow" section of the house. The Isodata result appends parts of the surroundings of the building to the building segment. Figure 75 provides results of the region growing algorithm for three spectral distances (the seed point is the intersection point of the vertical and horizontal lines). The region growing result for the 50 and 100 spectral distances may be considered for the building segmentation. No result provides a good segmentation as the 50 distance misses parts on the lower right side of the house and the 100 distance adds ground pixels to the building. The 100 distance result would be considered better and is the closest result to the Isodata procedure in figure 74. The building region may be partially extracted using the Isodata or the region growing procedures, but requires user post-processing corrections to add the "shadowed" sections of the house.

The histogram plots for the image are shown in figure 76. Compared to the histogram plots of scenario 1, there seem to be a greater correlation between the bands. The residential parcels histograms are again, not "smooth" and as a result produce a poor segmentation result for the histogram analysis segmentation.
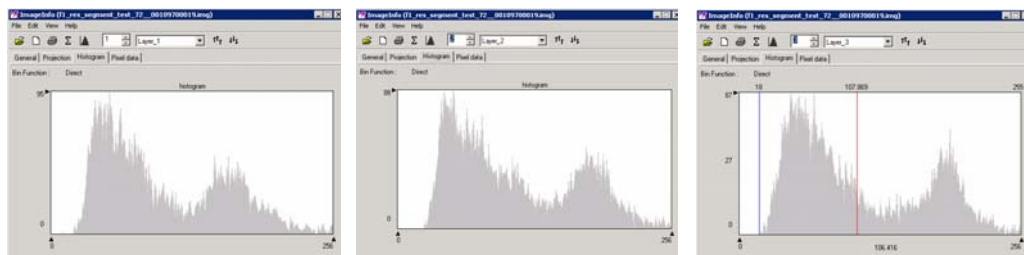
Figure 76 – scenario2. Histogram plots of the (left to right) Red, Green, Blue bands

Examination of the feature space plots (figure 77) emphasized the correlation between the bands. There is a better correlation between bands 1-3 and 2-3, consistent with the histogram plots of the individual bands (closer to a line, less scatter).
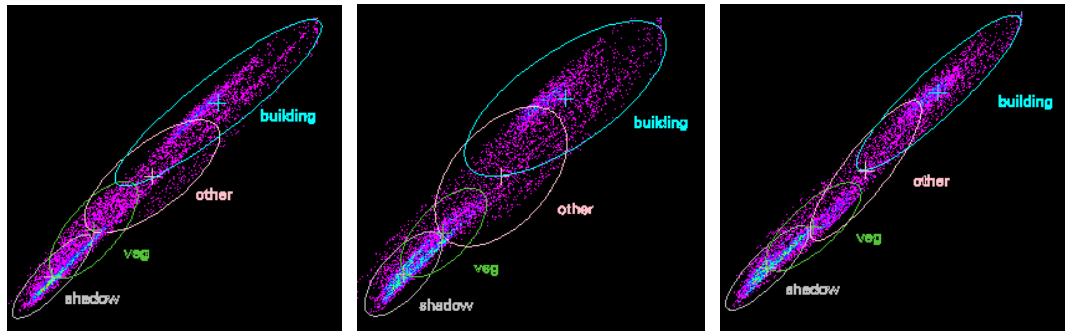
Figure 77 – scenario 2. Feature space plots for bands combinations: 1-2, 1-3, 2-3

In this scenario, the building ellipse (cyan) resides on the top end of the spectral plot and overlaps with the class "other" that represents miscellaneous objects around the building (ground, asphalt…). The overlap explains the addition of objects on the lower right part of the image to the building segments. The cyan colored pixels within the building ellipse indicate that there is a large number of pixels in that class. This is consistent with the classification of building pixels in the image (figure 74).

The histogram and the feature space plot indicate that it is possible to define fairly distinct classes within the parcel for scenario 2, with the exception of shadows on the roof and relatively minor additions. The Isodata and the region growing perform better and may require manual post-processing.

Scenario 3: Two building sections; shadows; no trees; one side illumination
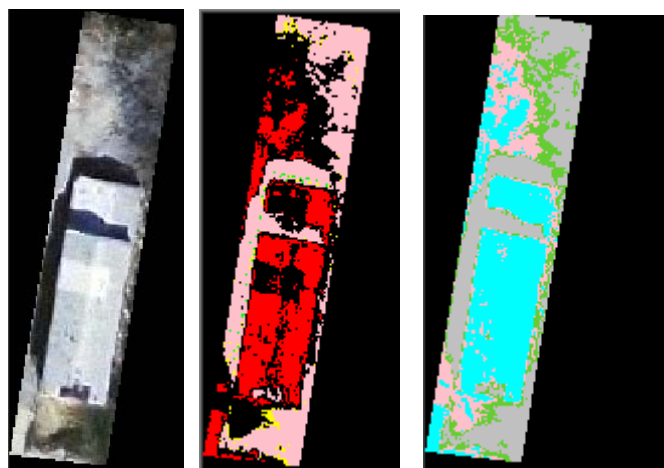
Figure 78 – scenario 3 [original image] [Histogram analysis Result] [Isodata Result].
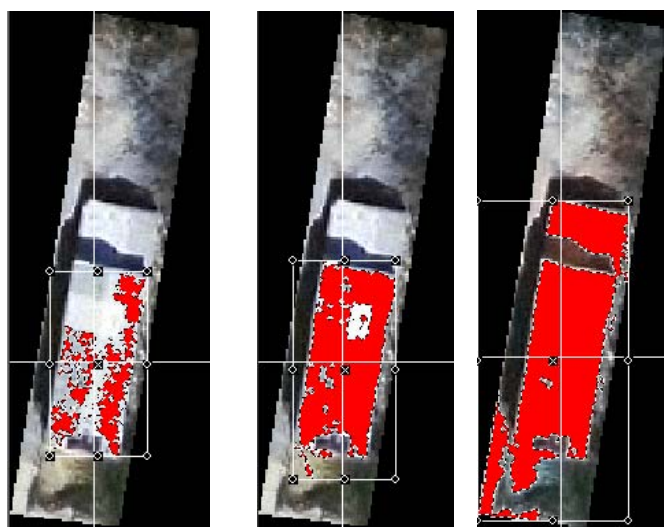
Figure 79 - scenario 3: region growing result (spectral distance = 20, 50, 100).

Figure 78 illustrates the original parcel image and the result of the automatic segmentation procedures. The image shows a fairly simple scene with shadow casts between the two sections of the building, no trees and a tilted roof. The histogram analysis result provides an adequate segmentation for the two sections of the roof. The Isodata algorithm provides a better segmentation that includes parts of the surroundings

within the building class. The buildings class in this scenario includes separates segments. The building roof is comprised of two distinct segments due to the shadow gap. Figure 79 provides the results for the region growing algorithm for three spectral distances (the seed point is the intersection point of the vertical and horizontal lines). It is clear that spectral distance of 100 provides the best representation of the building as it connects the two sections of the roof. The segment is solid with some additions and no major gaps. The building region may be partially extracted using any of the three algorithms, but requires user post-processing corrections.

The histogram plots for the image are shown in figure 80. The histogram plots show greater correlation between bands 1 and 2, consistent with the feature space plots in figure 81.
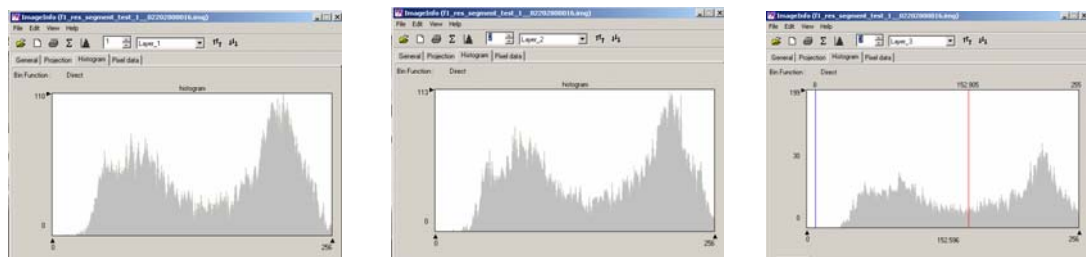
Figure 80 – scenario3. Histogram plots of the (left to right) Red, Green, Blue bands

As indicated in the feature space plots, the building and the ground share the same class (cyan). The ground pixels share spectral characteristics with the building and the dry vegetation and branches around the buildings. As a result, besides the ground distinct class, ground pixels appear in the building and vegetation classes. The cyan pixels (in the segmented image) that are not connected to the building segment illustrate that confusion.

143

As illustrated in the above examples, the building ellipse comprises of cyan colored pixels that indicate a concentration of the cyan class pixels (building and ground) in the image. That is consistent with the significant number of building pixels in the Isodata classification result in figure 78.
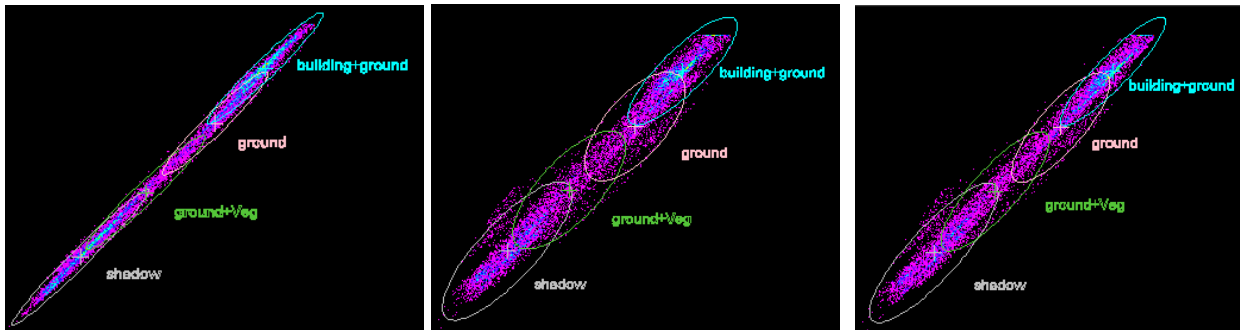
Figure 81 – scenario 3. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The histogram and the feature space plot indicate that it is possible to define the building with minor ground additions for scenario 3. Both the histogram analysis and the Isodata algorithms do not consider the actual location of a pixel within the image prior to the segmentation. As a result, both methods segmented erroneously several regions within the image as part of the building class. Due to shadow casts, the region growing procedure may require several seed points, one per section, to provide segmentation of the entire roof.

Any of the three algorithms may be used to segment the buildings for scenario 3, with some misclassifications.

Scenario 4: Several buildings; insignificant shadows on the roof (slope roof); trees occlusions on one building; 1 side illumination
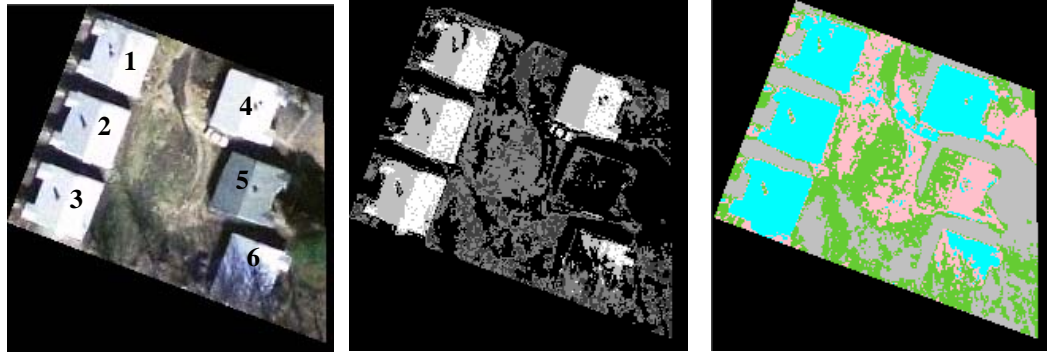
Figure 82 – scenario 4 [original image] [Histogram analysis Result] [Isodata Result].
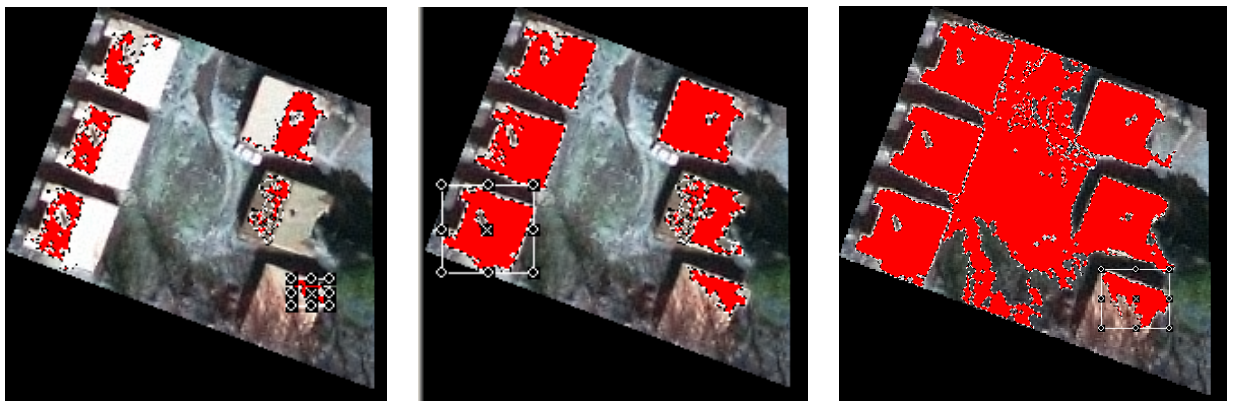
Figure 83 - scenario 4: region growing result (spectral distance = 20, 50, 100).

Figure 82 illustrates the original parcel image and the result of the automatic segmentation procedures. This image differs from the previous examples since it contains multiple houses with different circumstances. Houses 1-4 are very similar in color and roof tilt. House number 5 has a distinct color and relatively simple scene with no trees and shadow. House 6 illustrates a more complex scene with significant trees and shadow occlusion. The histogram analysis technique performs well for the four simple buildings with the exception of segmenting separately the two sides of the roof. House number 5 is not dominant enough within the image to be individually segmented. Only a small part of house number 6 can be clearly seen from above and segmented. The

145

Isodata provides a slightly better result since it combines the two parts of the roof to create one continuous roof segment for the simple four houses. The region growing results requires a seed point per house. A spectral distance of 50 provides the best segmentation result and extracts the entire roof of the four houses. For a spectral distance of 100, many additional pixels were erroneously added to the region. Clearly, the four houses, which represent fairly simple extraction scenarios without trees or shadow, can be extracted with any of the methods. House number 6 can not be extracted without a human intervention due to significant occlusion. House number 5 appears to have more variety of grey levels on the roof, hence, no segmentation method managed to successfully segment the house. The region growing result for a spectral distance of 50 illustrates the signature complexity. The region growing methods extracts houses 1-4 are but fails for house number 5.

The histogram plots for the image are shown in figure 84. The histogram plots in this scenario appear more flat, since the value 255 (white) has the largest number of pixels in the image compared to the other grey levels. The peak on the right side of the histogram represent the light section of the roofs, while the peak on the left side of the histogram represents the shadows and darker regions in the image. Clearly, bands 1 and 2 are more correlated then any other band combination. That correlation is consistent with the feature space plots in figure 85.
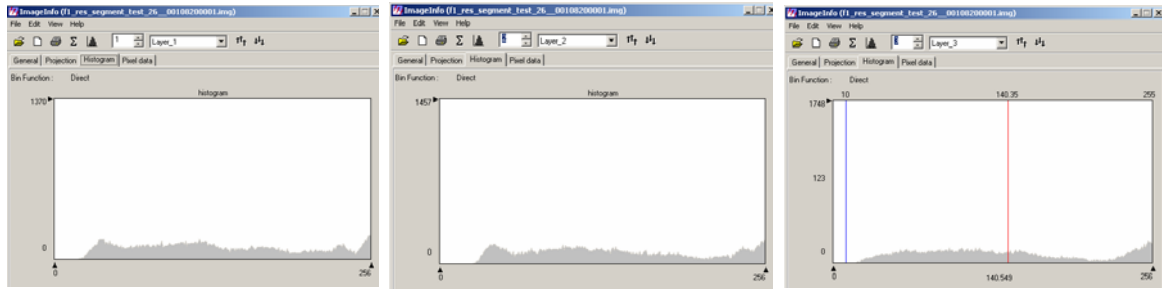
Figure 84 – scenario 4. Histogram plots of the (left to right) Red, Green, Blue bands

The segmentation result of the Isodata algorithm can be examined using feature space plots (figure 85).
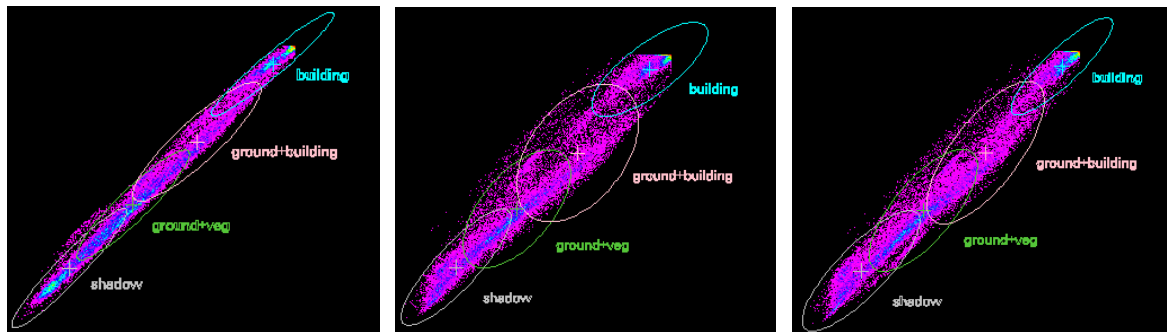
Figure 85 – scenario 4. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The building class (cyan) represents the light color houses and has a small overlap with the ground, as depicted in the Isodata result in figure 82. The ground and building class (pink) represents house number 5 that shares spectral characteristics with the ground object. There is a clear concentration of pixels in the cyan class (cyan and blue pixels) that represent the four houses. Opposite to those houses, the pink class corresponds to house number 5 and represents smaller number of pixels in the image.

147

The histogram and the feature space plot indicate that it is possible to distinctly define a class for the light color, simple houses (1-4). The trees and more complex signature for house number 5 prevent any of the methods from adequately segmenting the roof.

Scenario 5: Mainly single section building; insignificant shadows on the roof; no trees occlusion; two sides illumination
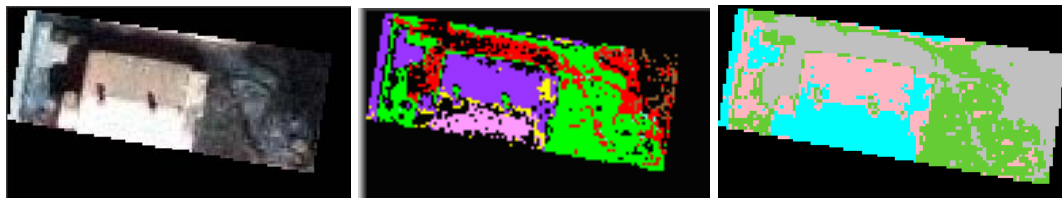
**Figure 86 – scenario 5 [original image] [Histogram analysis Result] [Isodata Result].**
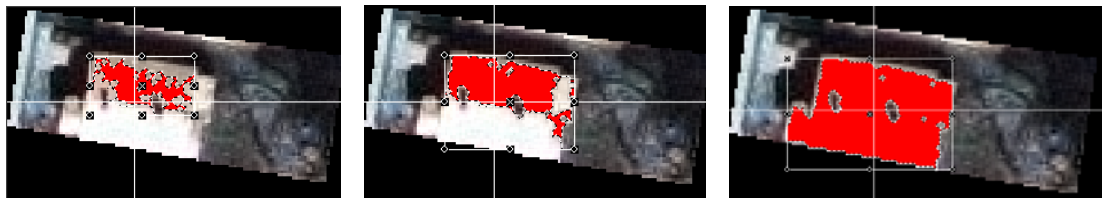
Figure 87 - scenario 5: region growing result (spectral distance = 20, 50, 100).

Figure 86 illustrates the original parcel image and the result of the automatic segmentation procedures. The image shows a fairly simple scene with limited shadow casts on the roof, no trees occlusion. The histogram analysis and the Isodata procedures segment the roof in two parts. Those algorithms segment other regions of the image within the same building class. The Isodata algorithm seems to perform slightly better and extracts the roof area as two classes. Figure 87 provides the results for the region growing algorithm. It is clear that spectral distance of 100 provides the best

148

representation of the entire roof. Due to the significant difference in grey levels between the two sides, a spectral distance of a 100 is required. The region does not append pixels beyond the house area due to the vast difference between the house and the surroundings. That is a distinct and not very common case.

The histogram plots for the image are shown in figure 88. The plots show higher correlation between bands 1 and 2. The correlation is consistent with the feature space plots in figure 89.
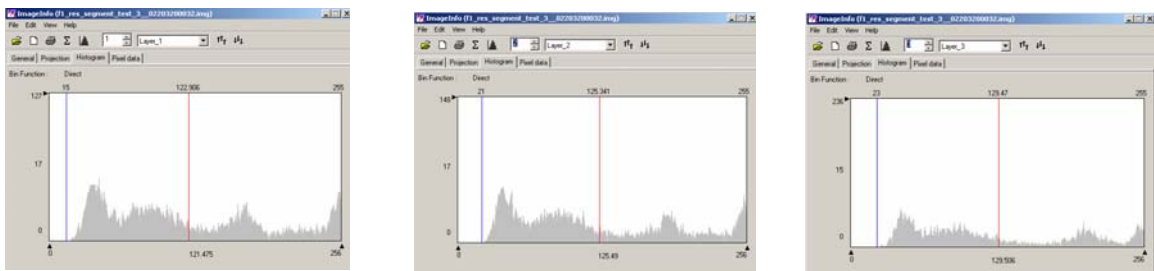
Figure 88 – scenario 5. Histogram plots of the (left to right) Red, Green, Blue bands

The confusion between the building and the ground classes is clearly depicted in the segmentation result (figure 86). The pink class and the cyan class that correspond to the two sides of the roof, have an area of overlap (figure 89). The overlap between the classes provides a spectral explanation for the misclassification of pixels within the building classes.
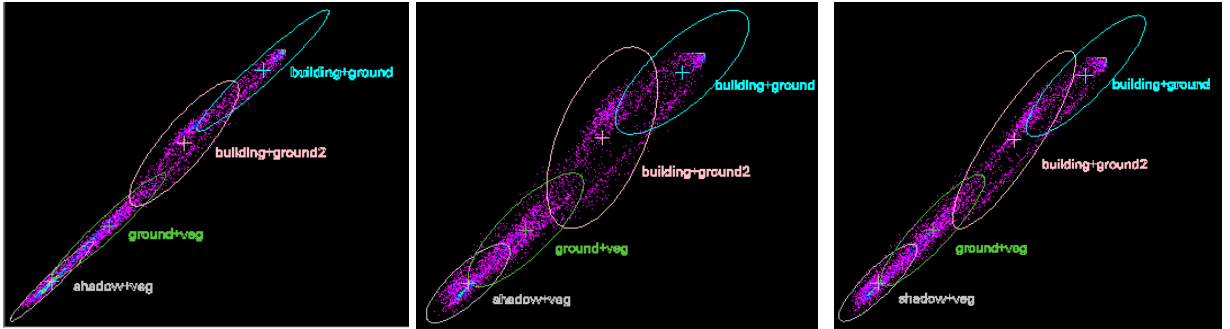
Figure 89 – scenario 5. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The histogram and the feature space plot indicate that it is possible to define the classes of the roof. Each class represents one side of the roof and includes pixels that do not belong to the roof. The region growing with a spectral distance of 100 provided the best segmentation. The large spectral distance, as shown in previous examples, is usually too broad and significantly extends the roof area.

The following testing scenarios provide examples of very poor segmentation results. The complex scenes are evaluated to determine which factors prevent successful segmentation and further extraction of the structure.

Scenario 6: Two building sections; significant shadows on the roof; trees occlusion, one side illumination
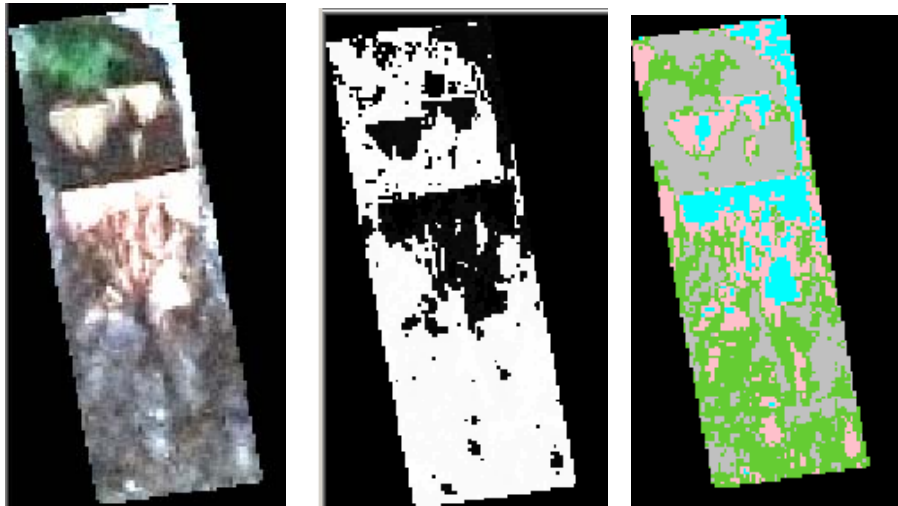
**Figure 90 – scenario 6 [original image] [Histogram analysis Result] [Isodata Result].**
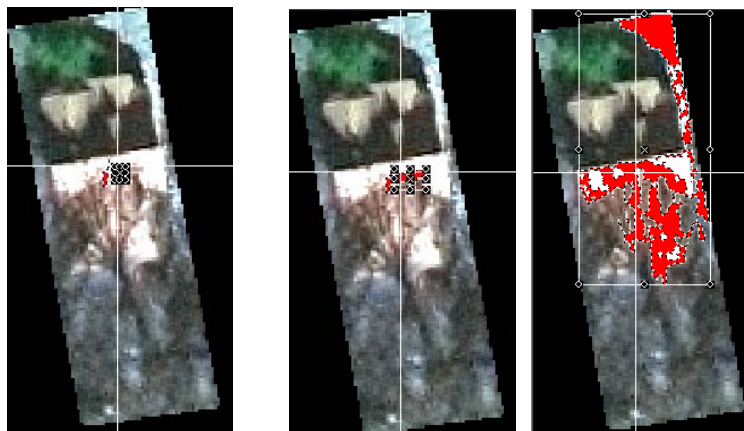
Figure 91- scenario 6: region growing result (spectral distance = 20, 50, 100).

Figure 90 illustrates the original parcel image and the result of the automatic segmentation procedures. The image illustrates a complex scene with considerable shadow casts on the roof and trees occlusion. Obviously, when the roof can not be

clearly seen from above, it can not be extracted. In the case of the house in figure 90, as well as in many other cases, a human may be able to determine the corners of the footprint. The same can not be easily asserted for an automatic process. All the tested procedures, including the semi automatic region growing algorithm, failed to adequately segment the roof.

The histogram plots of the image are shown in figure 92. The large peak on the left (dark) part of the histogram explains the large white segment in the histogram analysis segmentation result (figure 91). The pixels in the image are basically divided between the large peak and the rest of the histogram. As a result, all the darker pixels are segmented as one class.
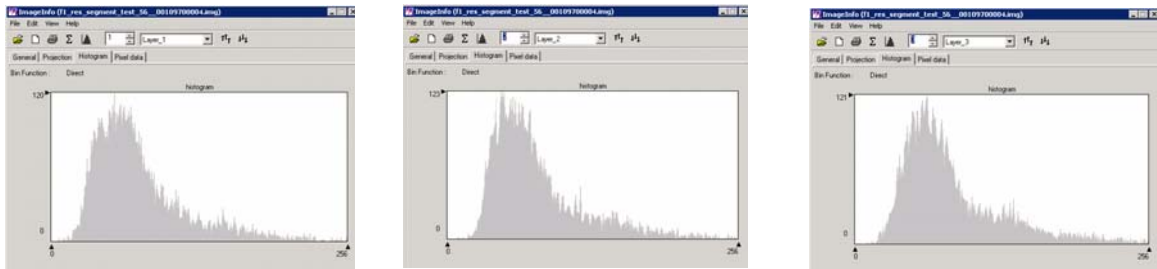
Figure 92 – scenario 6. Histogram plots of the (left to right) Red, Green, Blue bands

The feature space plots in figure 93 are consistent with the evident high correlation between the bands, especially bands 1 and 2.
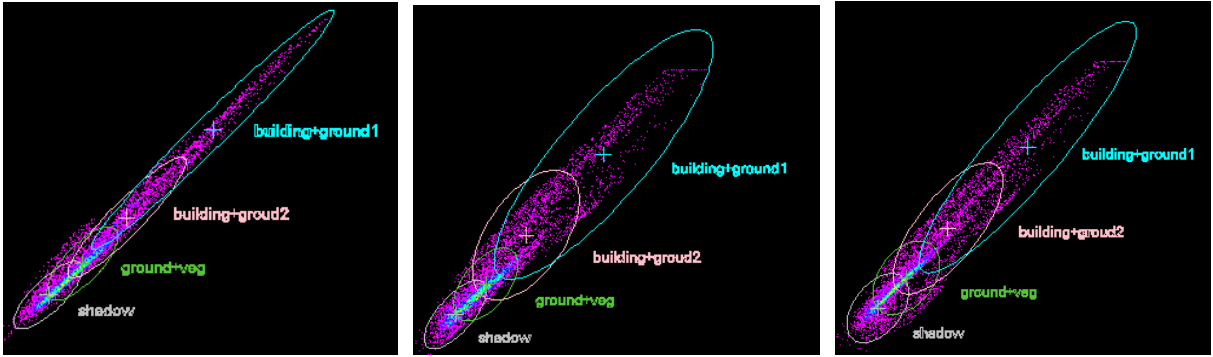
Figure 93 – scenario 6. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The ellipses in figure 93 show very large building and ground (cyan and pink) classes. There is significant overlap between the building and ground classes which explains the poor segmentation result. The ground and vegetation (green) class is completely overlapped by neighboring classes. The overlap explains the misclassification between the vegetation (trees and branches) and the shadows. The majority of the pixels in the image concentrate within the lower left part of the plot (cyan) which represents the shadow, ground and vegetation classes. The buildings classes represent only a small number of pixels in the image due to lack of visibility of the building from above.

The histogram and the feature space plot indicate that the majority of pixels in the image represent the shadows and vegetation (branches). As only small section of the house can be clearly seen from above, it is very difficult to automatically extract the building. Moreover, the shadows on the roof prevent the use of a semiautomatic process such as region growing due to the extreme variability of pixel colors on the roof. Scenario 6 illustrates an example of a parcel scene that requires a manual extraction of the building footprint.

Scenario 7: Single section building; shadows; slight trees branches occlusion; one side illumination
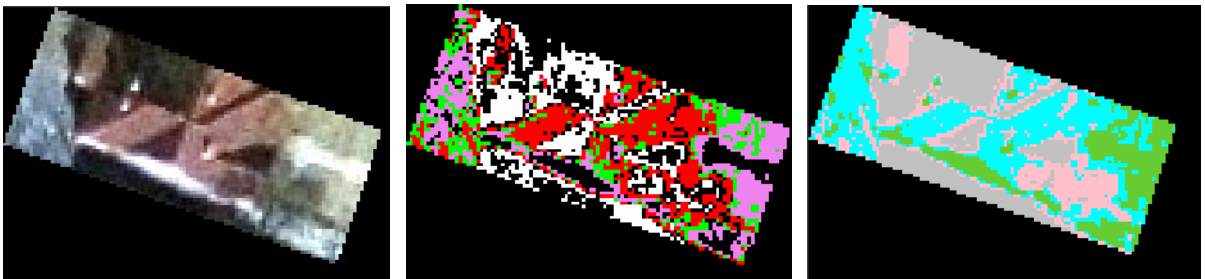
Figure 94– scenario 7 [original image] [Histogram analysis Result] [Isodata Result].
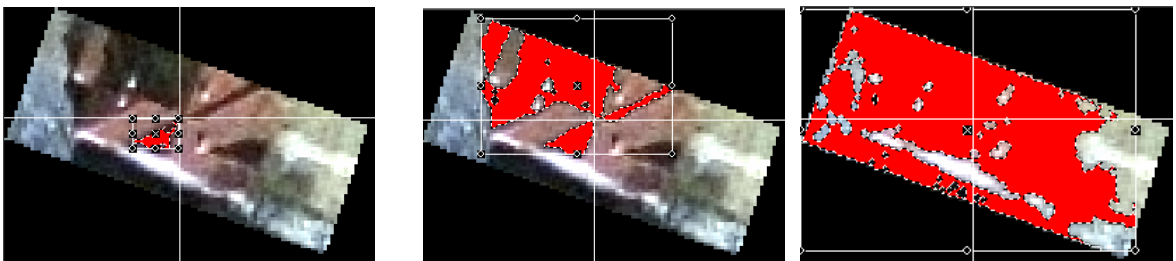
Figure 95- scenario 7: region growing result (spectral distance = 20, 50, 100).

Figure 94 provides the original parcel image and the result of the automatic segmentation procedures. The image illustrates a fairly simple scene at first sight. The entire house is clearly seen from above and there is no trees occlusion. The factor that turns this image into a complex scene is the shadows. The roof comprises of multiple tilted sections at different heights. The sections cast shadows due to the sun illumination, and as a result, divide the roof into lighted and shadowed regions. The result is multi-

class representation of the roof. The cyan, pink and grey classes interleave on the roof.

Evidently, the region growing algorithm fails to correctly segment the entire roof. The

result either covers a small section of the roof (spectral distance 20, 50) or adds regions

around the house to the roof segment (spectral distance 100). The histogram analysis

result shows a similar result, but with less continuous segments. The histogram plots in

figure 96 provide some explanation to the complex histogram analysis. Bands 1 and 2

are fairly correlated, but are represented by abrupt changes in the histogram which makes

the peak analysis much more challenging. The third band is not as correlated to bands 1

and 2 which is consistent with the feature class plot in figure 97.
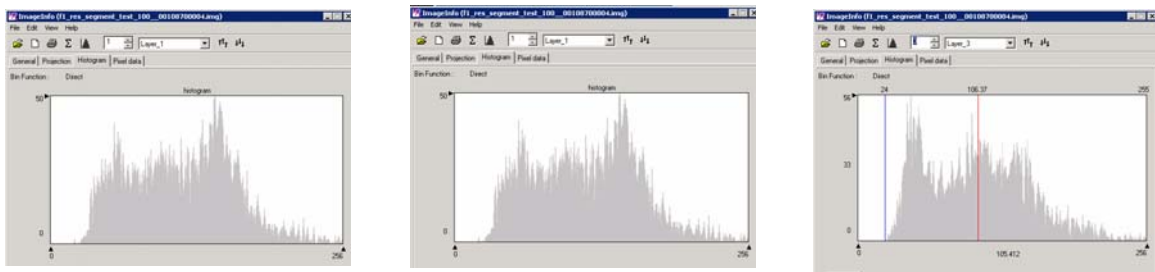
Figure 96 – scenario 7. Histogram plots of the (left to right) Red, Green, Blue bands

The feature space plots in figure 97 show higher degree of correlation between bands

1 and 2. There a significant overlap between classes and significant confusion between

the ground and building classes. The shadow class includes all shadow areas, and for this

example comprised of large sections on the roof.
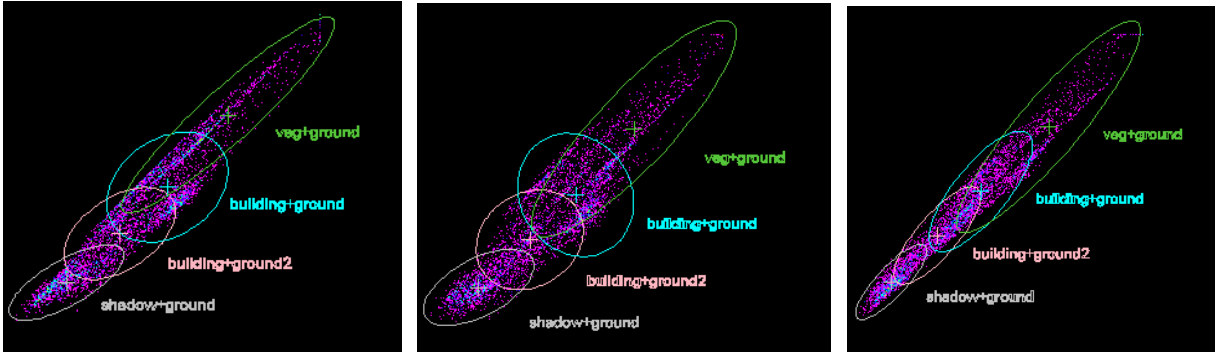
Figure 97 – scenario 7. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The testing results for this image indicate that the roof construction has a crucial affect on the extraction result. Multiple tilted roof sections may produce excess shadow regions on the roof and directly prevent successful segmentation.   No automatic extraction method can adequately extract the house roof and a human intervention is clearly required.

Scenario 8: Two building sections; shadows; trees occlusion; two side illumination
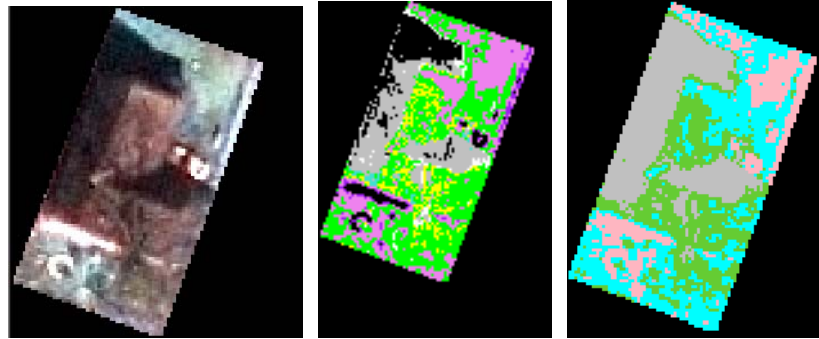
Figure 98– scenario 8 [original image] [Histogram analysis Result] [Isodata Result].
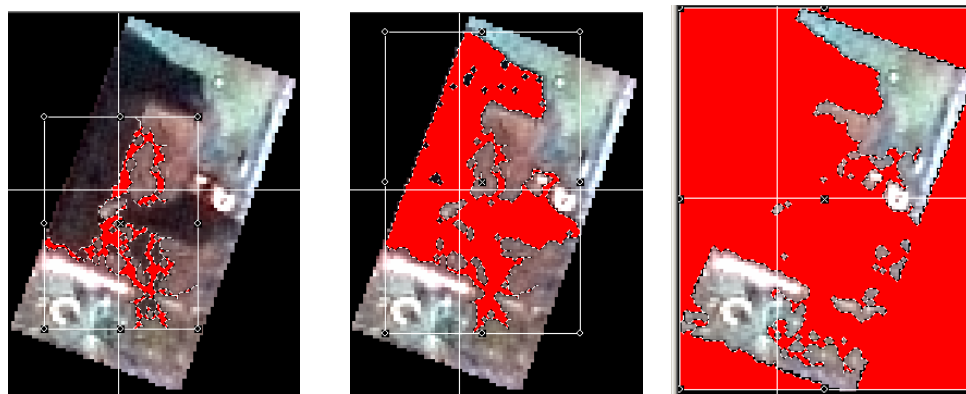
Figure 99- scenario 8: region growing result (spectral distance = 20, 50, 100).

Figure 98 illustrates the original parcel image and the result of the automatic segmentation procedures. The house suffers from tree occlusion and similar to previous example, the tilted roof sections cause shadow regions on the roof. The complexity of the segmentation is evident by the poor classification and segmentation result for all three methods. The Isodata algorithm shows fair segmentation for only the shadow regions in the image. The histogram analysis result does not distinctly segment any class (apart from the shadow area) and can not be used for further feature extraction. The region

growing algorithm fails to segment the roof. The result either covers only parts of the roof (spectral distance 20 and 50) or vastly extends beyond the house area (spectral distance 100).

The histogram plots of the image are shown in figure 100. The concentration of the histogram on the left part of the graph is consistent with the fairly darker pixels in the image and the feature space plots in figure 101. There is an evident higher correlation between band 1 and band 2 and much less correlation with band 3.
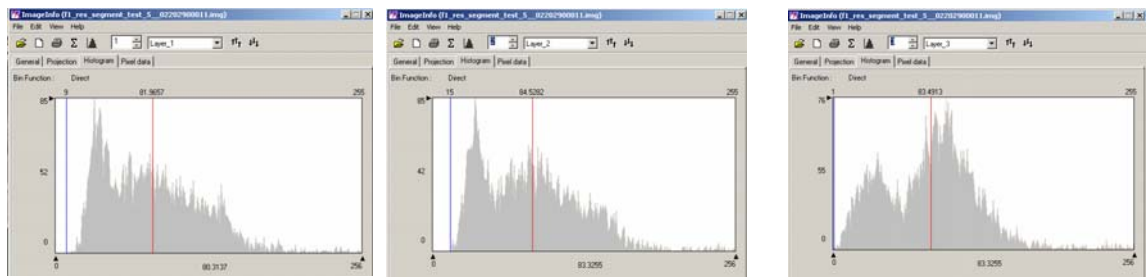
Figure 100 – scenario 8. Histogram plots of the (left to right) Red, Green, Blue bands

The feature space plots in figure 101 are consistent with the high correlation between bands 1 and 2.
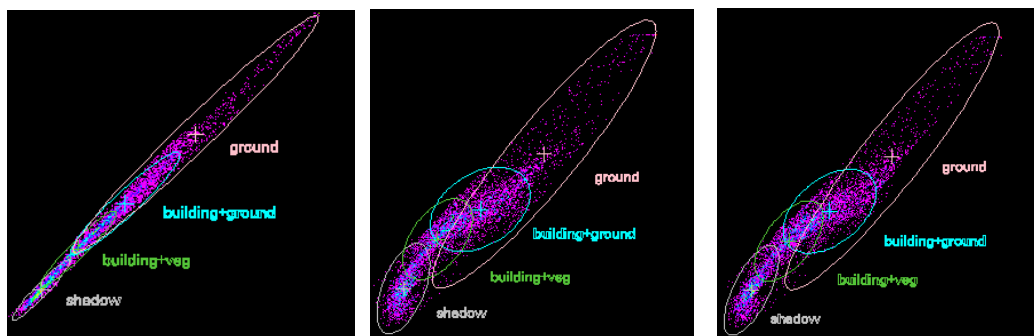
Figure 101 – scenario 8. Feature space plots for bands combinations: 1-2, 1-3, 2-3

The ellipses in figure 101 show pixel concentration on the lower left part of the plot. That indicates a large number of darker pixels in the image. There is significant overlap between the building and ground (pink and cyan) and some overlap of those classes with the building and vegetation class (green). That explains some of the misclassification results between the classes in figure 98. The shadow class, which includes many roof pixels, has a large concentration of pixels and does not have as much overlap with the other classes. As a result, this class is segmented fairly well and distinctly in the image.

The histogram and the feature space plot indicate that the majority of pixels in the image represent the shadows and vegetation (branches). Due to the wide variety of grey levels that represent the entire roof, an automatic algorithm would probably fail to accurately segment the roof. Scenario 8 is another example of a parcel scene that requires manual extraction of the building footprint.

Scenario 9: Several slope building; shadows; no trees; two side illumination
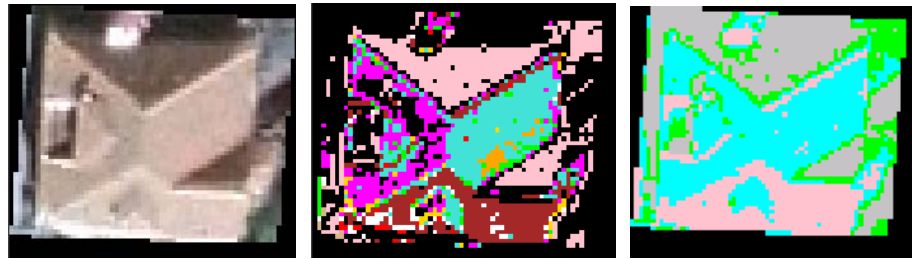
Figure 102– scenario 7 [original image] [Histogram analysis Result] [Isodata Result].
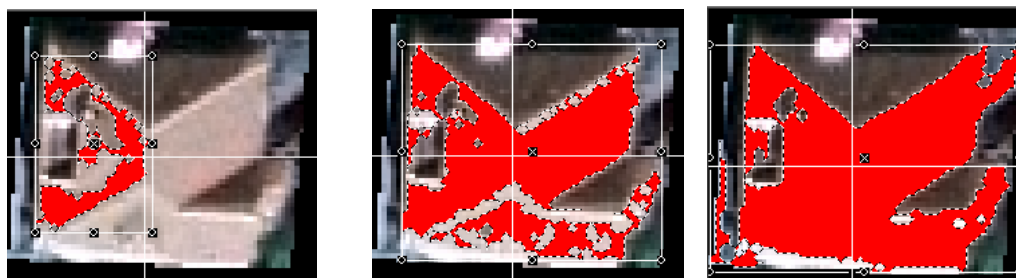

Figure 103- scenario 7: region growing result (spectral distance = 20, 50, 100).

Figure 102 provides the original parcel image and the result of the automatic segmentation procedures. The house in this example covers almost the entire parcel area. The house has no tree occlusion and the only factor that might affect the segmentation result is the multiple tilt roof sections. The tilted roof may derive shadow region, as depicted in scenario 7, and prevent successful segmentation. Similar to scenario 7, the roof is represented by several classes. The histogram analysis and the Isodata result show multiple segments on the roof. Unlike scenario 7, there is a fairly continuous and large section in the middle of the roof that is segmented by the Isodata and the region growing algorithm. This section provides a visual, better representation of the roof. That section, as segmented by the Isodata algorithm, does not adequately represent the entire area of the house. As a result additional manual post-processing is required.

160

The histogram plots in figure 104 provide some explanation to the histogram analysis result. Figure 102 shows four dominant classes that represent the roof (histogram analysis result in the middle). Those are fairly consistent with bands one and two in the histogram plots which appear highly correlated. The third band is not as correlated to bands 1 and 2. That is consistent with the feature class plot in figure 105.
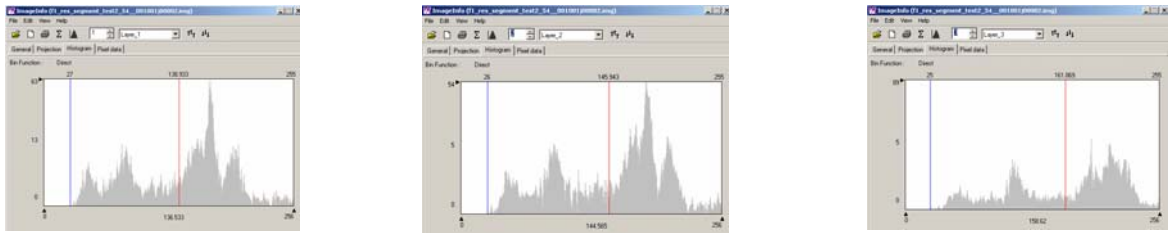
Figure 104 – scenario 7. Histogram plots of the (left to right) Red, Green, Blue bands

The feature space plots in figure 105 show higher degree of correlation between the bands 1 and 2. clearly, the lighter grey level are more dominant in that image as the concentration of most pixels (blue and cyan pixels) are located at the top end of the feature space plots. Since the house covers most of the image, the building class is part of all the classes. That explains the misclassification and confusion in the Isodata classification result. Each of the classes represents pixels on the roof as well as pixels around the house.
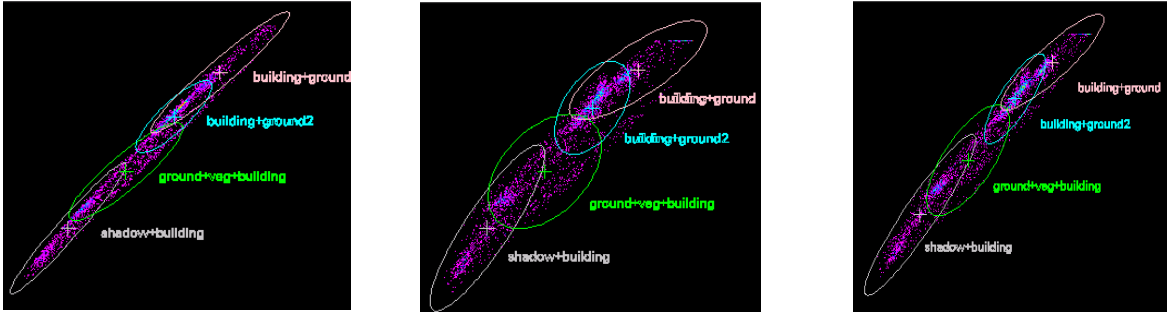
Figure 105 – scenario 7. Feature space plots for bands combinations: 1-2, 1-3, 2-3

Consistent with scenario 7, testing results for this image indicate that the roof construction has a crucial affect on the extraction result. Multiple tilted roof sections may produce shadow regions on the roof and prevent successful segmentation. Automatic extraction methods may only partially extract the roof top area and would require manual post-processing

4.2.3.4   Residential Testing Conclusions

Residential buildings and their environment pose great challenges on any automatic extraction algorithm. The nature of the residential scene emphasizes an attractive living environment.   As a result, residential houses are often characterized by interesting architecture and near by trees. Those may look nice, but make the automatic extraction of the structure complicated or even impossible.

The initial testing of the overall model on residential parcels indicated poor extraction results (4.2.3.1). Since the entire process relies on successful segmentation of the image, further investigation of the segmentation was performed.   The evaluation of the histogram analysis result was performed in comparison with a well known and commonly used classification procedure (ISODATA). The methods were tested in two testing areas

162

that differ by the amount of trees around the houses. No extraction procedure performed well enough to allow further automatic analysis and refinement of the segments. The ISODATA performed slightly better as indicated in section 4.2.3.1 but may still require exhaustive manual post processing. Vegetation was concluded to be one factor that prevents correct segmentation of the roof (table 5, section 4.2.3.1)

Following the automatic extraction procedure, a semi-automatic extraction algorithm was tested. The region-growing algorithm, which has been heavily used in building extraction projects, was examined on 100 parcels (4.2.3.2). The algorithm performance provides an indication to the complexity of the scene. In fairly homogenous regions, the algorithm performs fairly well (figure 69). Hence, failure to segment a roof indicates that the roof signature is either complex or the surroundings of the house cause the segment to extend the house area. In either case, the result is not sufficient for footprint extraction. Initial testing of the algorithm shows that 60 percent of the houses were not adequately extracted and expanded beyond the footprint of the roof.

As indicated by the testing, no method seems to provide an acceptable segmentation result for all residential images. In order to provide better guidelines for extraction, several representative residential parcels were selected and a comprehensive investigation of all the factors was performed (4.2.3.3). The results of the testing suggest that when no trees obscure parts of the roof, there is an insignificant, minor or no shadow presence on the roof, then a fair segmentation of the roof is possible using an automatic or semi automatic procedure. The histogram analysis is relatively more complex in residential parcels since it is characterized by a less smooth graph (abrupt changes in histogram values). As a result, the Isodata classification performed slightly better for the majority

of cases and is recommended for a residential parcel automatic segmentation. Before using the Isodata algorithm, the user should visually scan the residential parcels to verify the four class assumption (house; vegetation; driveway; shadows). The region growing algorithm provides a result that is consistent with or better then the Isodata procedure. However, this algorithm is semiautomatic and requires the user to place one or more seed points on the roof of the building. Moreover, the user has to define a spectral distance prior to the segmentation. In about 50 percent of the testing cases, spectral distance of 50 provides the better segmentation. That value is recommended as an initial attempt, but may probably require adjustments by the user.

Contrary to the simple residential scenes, significant shadows on the roof and major trees occlusions would prevent automatic or semi-automatic extraction of the roof. The presence of either factor was found to be enough to compromise the segmentation. Shadows may be cast by features around the house, but can be created by the multiple tilted sections of the roof. Obviously, when the roof can not be clearly seen from above, an automatic process would fail to extract the structure. In many cases, a human would be able to manually extract the footprint of the building. In some areas, a tree may hide considerable sections of the house or may cast large shadow regions on the roof, and prevent an automatic or manual extraction. It is recommended to visually scan the image prior to extraction. The user may then be able to decide on the most efficient extraction methodology.

4.2.4   High-rise Parcels Testing

For the selection of high-rise buildings in the Memphis area, high-rise was defined as a building with 10 or more stories. An initial selection of the parcels from the tax assessor records identified 60 high-rise buildings that reside within 49 parcels. Prior to extracting the building, a visual scan of the buildings and parcels was performed. The scan revealed several prominent and occasionally unique characteristics of high-rise buildings and high-rise parcels. Those characteristics were identified to pose problems or completely prevent the extraction of the building footprint. A survey of those characteristics with examples is provided in the following section.

4.2.4.1 High-rise buildings characteristics

This section provides several examples of common problems that arise during the extraction of high-rise buildings. Several of those characteristics may be unique to high-rise buildings or represented in a unique manner in those types of buildings.

Figure 106 provides an example of a high-rise building with multiple sections. One part of the building appears much higher then the other part. As a result, the high section cast a fairly large shadow and hides a significant section of the building. Shadow cast on a building can definitely compromise the footprint extraction process.

Figure 106 – multi-section high-rise building. Building footprint in green.

Compound buildings are not unique for high-rise buildings, but pose a great challenge to the extraction process. Figure 107 illustrates the difficulty of extracting such a compound. As seen in the figure, each section of the compound appears differently and would be extracted as a separate building.

Figure 107 – High-rise: Compound of high-rise buildings. Building footprint in green.

Parcel boundary in yellow.

One phenomenon that seems to be fairly common for high-rise buildings is the inconsistency between the building outline and parcel boundary. As seen in figure 108, the high-rise buildings (green polygons) do not fully reside inside the parcel area (yellow polygons).   In those examples, the majority of the building area is located outside the

parcel boundary. In the Memphis area, when the building is located inside the parcel, it

seems to cover the majority or the entire area of the parcel.
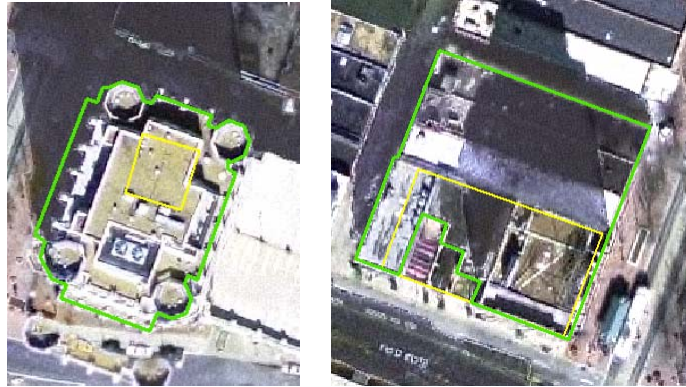
Figure 108 – High-rise building (green) outside the parcel boundary (yellow)

As already seen in the above examples, figure 109 provides multiple examples of

high-rise buildings with complex roof signatures. The presence of multiple objects on the

roofs, multiple levels and multiple sections, create a complex and complicated roof

signature. High-rises are commonly used by companies to install features such as

telecom switches and antennas due to the relative height above the environment. The

extreme difference in signature between the different sections and variety within the

sections adds to the complexity of the extraction procedure.

Figure 109 – High-rise: Complex roof signature

An interesting example that emphasizes the difficulty of extracting high-rise buildings, even manually, is provided in figure 110. Figure 110 includes two images of the same building. The building in the Memphis orthophoto on the left and the same building, taken from google-earth, on the right. On the orthophoto, it is very hard to identify the two sections of the building. The building appears to be one tower rather then one tower and a lower level section.

Figure 110 – High-rise building. (Left) Memphis orthophoto (Right) Image taken from google-earth

As predicted for high-rise buildings, shadow occlusion is one major obstacle for the extraction process. Since the size of the shadow depends on the height of the building, the shadow cast may affect buildings around the high-rise structure as well as lower section of the building.
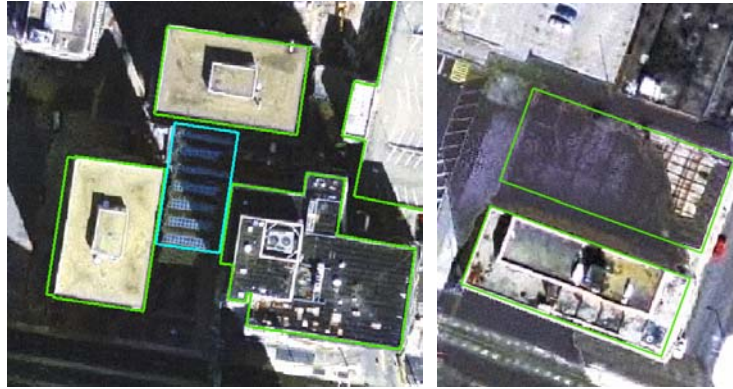
Figure 111 – High-rise: shadow occlusion

Prior to extracting the building, a visual scan of the high-rise buildings in the Memphis area was performed. Buildings were categorized based on the above parameters as follows (buildings are not restricted to one category):

- 21 buildings outlines intersect the parcel outline

- 19 buildings have multiple sections or are part of a larger compound

- 44 buildings have complex roof signature.

- 23 buildings have multiple levels. All have mild or severe (obscure large sections) shadow occlusion.

- 10 buildings have a roof signature that may allow automatic segmentation.

The above categories assert that the majority of the buildings (44 out of 60) have a complex signature that may prevent adequate segmentation and extraction of the footprint. About 30 percent of the buildings intersect the parcel outline and the same

percentage of buildings has significant shadow occlusion. Out of the 60 test buildings, only 10 were found to be potentially extracted by an automatic process.

4.2.4.2 High-rise testing results

Similar to the commercial buildings testing, section 4.2.2.1, the 49 high-rise parcels were used to produce parcel-sized images. The result of the extraction process is a vector, polygon layer. The polygon layer was compared to the ground truth, digitized building dataset. A quantitative evaluation of the result compared the vector layer with the digitized building dataset. Table 6 presents the results for the high-rise buildings testing. The testing parcel subset includes 49 parcels that contain 60 buildings. The "Method" column represents the sequence of processes as used for the testing. "All eliminate" means a full run that includes all three elimination steps: elimination by size, shadow and geometry characteristics. "No size" indicates that the elimination by size was not part of the process. Similarly, the methods "No Shadow" and "No Geometry" were tested. Type 1 error represents the number of segments that were extracted but are not buildings. Type 2 error represents the number of buildings that were not identified. Those buildings were either not extracted or extracted and erroneously eliminated. Partial extraction represents segments that represent buildings, but were extracted with an error of more then 50 percent of the building (for example, 45 percent of the building was extracted). Success represents the remainder of the segments, which cover 50 or more percent of the buildings. Three additional values were calculated in order to better evaluate the degree of error and discrepancy between the result and the ground truth building layer. The average of the over/undershoot for the buildings is calculated. The

average value is the calculated mean of undershoot and overshoot for all the buildings. Last, the number of buildings with an area error less the 10 and 15 percent is provided.

Table 6 - High-rise testing results

Method	Type 1	Type 2	Partial	Success	Average Over/under	<10%	<15%
All eliminate	11	27	23	10	26.72	2	4
No size elimination	13	18	32	10	26.71	2	3
No Shadow elimination	16	21	25	14	26.95	2	4
No Geometry elimination	74	19	35	10	36.77	2	4

For the full run, 11 (about 18%) segments were flagged as Type 1 error, "false positive" identification, where a building was extracted but does not exist. 27 (about 45%) segments were flagged as Type 2 error, "false negative" identification, where the building exists but was not identified. Out of the 60 buildings, 2 (about 3%) were extracted with under/overshoot percentage of not more then 10% and 4 (about 6%) with under/overshoot percentage of not more then 15%. In the second scenario ("No Size"), the process added 9 legitimate parts of buildings, and Type 1 errors increased to 13. In the third scenario ("No Shadow"), the process identified 6 additional legitimate parts of buildings, but Type 1 errors increased to 16 segments. The final scenario ("No

171

Geometry") identified 8 additional legitimate parts but received a significant increase in Type 1 errors.

(a)   (b)

Figure 112 – High-rise: successful building extraction.

Figure 112 provides two examples of successful building extraction results. The left image (section a) contains one parcel and one building. The green polygon represents the digitized building and the red polygon represents the result of the automatic process. The right image (section b) illustrates three results: successful extraction of the corner building; type 2 error, where the building was not identified (the left most building); only parts of the building were identified separately (top building). Figure 113 provides similar examples of one successful footprint extraction and two partial extractions of sections within the building.

Figure 113 – High-rise extraction result

The one overwhelming conclusion from the above results is that the automatic process fails to successfully extract large percentage of high-rise buildings. The process manages to identify and extract simple buildings, such as the examples in figure 112 and 120, but those are not representative of high-rise parcels. The characteristics of the high-rise building and the high-rise parcels, as indicated in section 4.2.4.1, prevent a successful footprint extraction using any automatic process. The complexity of the signature on the roof of the building and the shadow casts would prevent an adequate segmentation using an automatic or semi-automatic process such as the region growing (4.2.3.2).

One phenomenon that was not apparent in the Memphis orthophoto, but may have a significant affect on any extraction result for high-rise buildings is "relief displacement". The phenomenon and the possible artifacts for the extraction process are discussed in the following section.

4.2.4.3 Relief Displacement

In aerial imagery, one of the inherent distortions of the perspective projection is that objects are displaced from their accurate planimetric location due to a phenomenon called "relief displacement". Based on a definition from ASPRS (American Society for Photogrammetry and Remote Sensing), relief displacement is "… displacement of an image-point toward or away from the nadir according as the corresponding ground point is below or above the ground nadir Also called height displacement ". The severity of the relief displacement artifact is determined by the location of the object in the image, the height of the object and the altitude of the sensor. In order to overlay GIS data sets and measure distances on the image, un-rectified images are "orthorectifid". Orthorectification changes the geometry of the image such that distortions due to terrain relief displacement, tilt and camera attitude variation are corrected (Jensen, 2000), and the result image has uniform scale. The orthorectification process is usually performed using DTM (Digital Terrain Model) of the area which does not account for structures above the terrain. As a result, the process does not eliminate displacement of tall objects such as high-rise buildings. In figure 114 the high-rise buildings are "leaning" and show the side of the structure. In a "true" orthophoto that problem is corrected, but it would still suffer from issues such as occluded area. True orthophoto generation that uses DSM (Digital Surface Model) rather then a DTM, has been in research (Ayman et al, 2007; Gunay et al, 2007) but has not been widely adapted by the industry.

Figure 114 – High-rise buildings in mid-town Atlanta
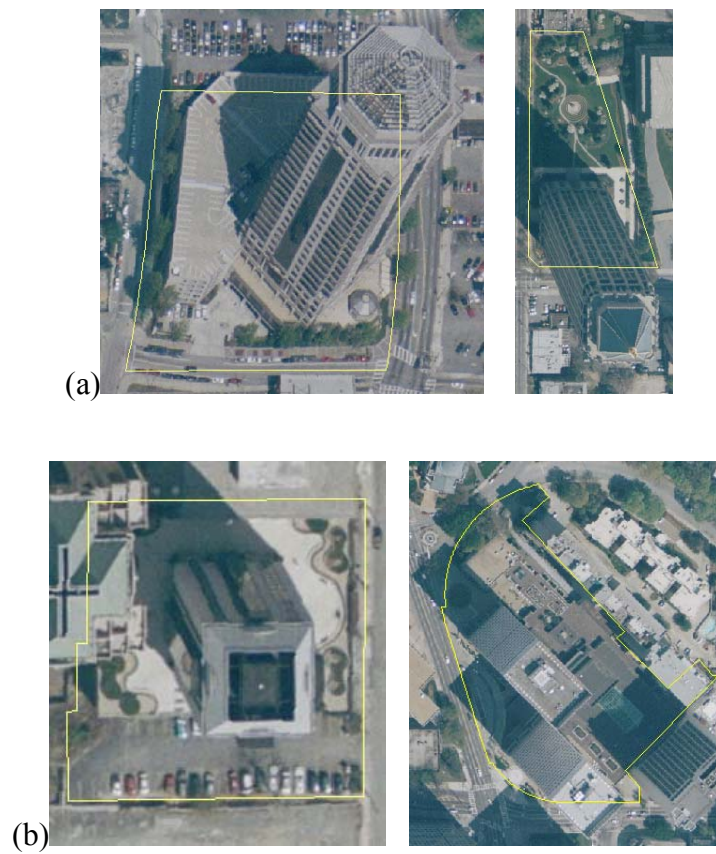
Figure 115 – High-rise buildings in downtown Atlanta. Parcel lines in yellow

Figure 115 provides several examples of the artifact caused by the "leaning" of the buildings. While the bottom of the building is located inside the parcel, the top of the building seems to be located outside the parcel boundaries (section a). Depending on the severity of the relief displacement affect, the top of the building may stay within the parcel boundary (section b). The side of the building can complicate the extraction procedure since it may be extracted as a valid section of the building roof. The texture of the side has been used in research to locate high-rise buildings (Liang, Weixin and Jianjun, 2004).

Prior to any extraction procedure, the user has to evaluate the severity of the relief displacement phenomenon. A minor affect may allow automatic or semi-automatic extraction, while more severe artifacts may prevent the extraction. As a result, generation of a "true" orthophoto should be considered on a project basis.

4.2.5   Number of Peaks Evaluation

The initial step of the automatic building extraction process is the histogram analysis (4.1.2.1). The peaks in the histogram are located and a range around the peak is defined as the peak area. The values of the grey levels (colors) within each range are then segmented back into the image to create a classification (segmentation) result. The result is further analyzed to identify the buildings and extract the footprint.

In the current testing, the search was limited to four peaks within the histogram (4.1.2.1). The initial assumption of the process is that the building area is fairly dominant within a localized region such as a parcel. As a result, the range of colors that represent the building should manifest itself as a peak within the histogram. In cases where the

building is large enough, the highest peak would be representing the building. In other cases it may be the second, third or even forth peak. The purpose of this section is to evaluate how many peaks should be selected to extract buildings from aerial parcel-sized images. For example, when selecting four peaks, what is the benefit of the forth peak compared with the number of type 1 errors that are added to the analysis.

In order to evaluate the contribution of each peak to the overall extraction process, the segmentation result of the peaks in the image space is examined. The location of the pixels, representing each of the peaks, is determined as "inside" or "outside" a building polygon. Figure 116 and figure 117 illustrate the steps to evaluate the location of the peaks compared to the buildings.
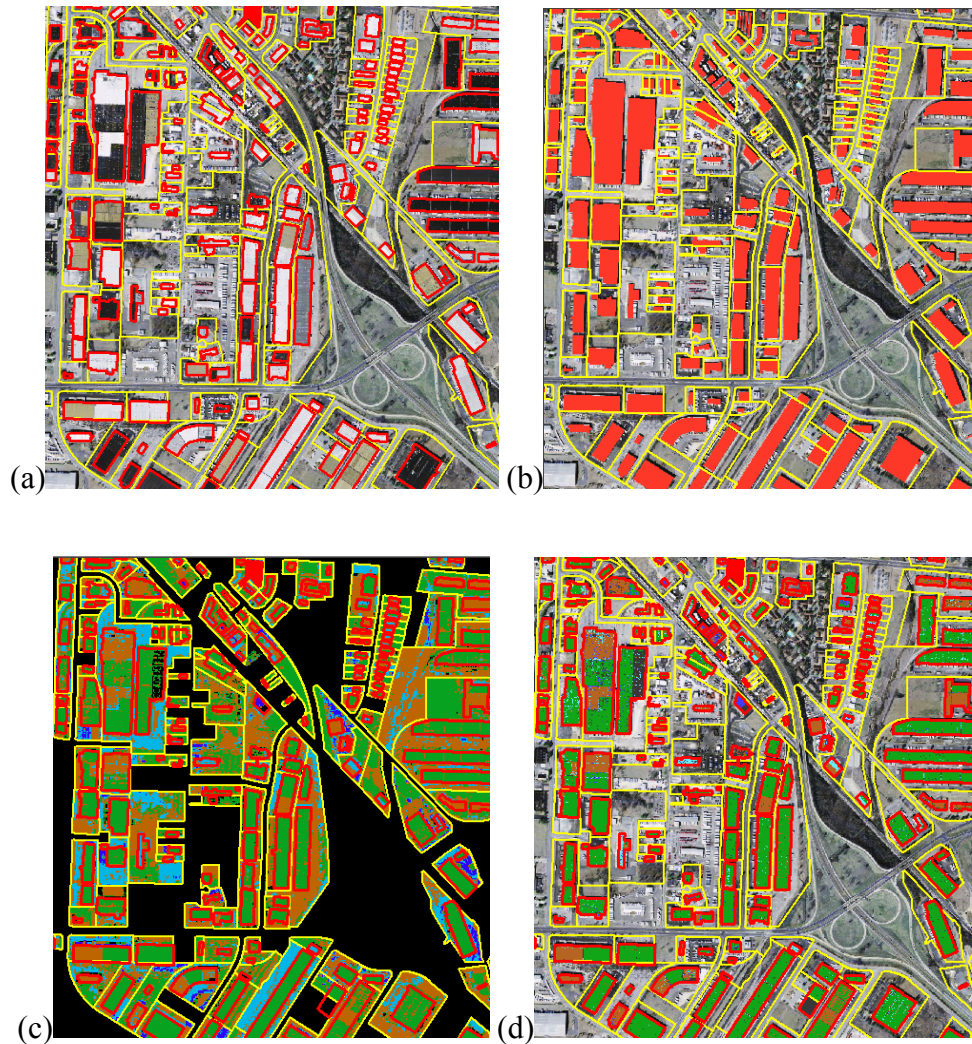
Figure 116 – Peak analysis of a 1.25 square mile in downtown Memphis

Figure 116 and figure 117 provide two examples for the peaks analysis. Every reference to figure 116 reflects directly to figure 117. Section (a) in figure 116 shows part of the testing area. The yellow polygons represent parcels boundaries and the red polygons represent digitized buildings. The digitized buildings polygon layer was converted to raster: buildings coded as value 1 and non-buildings coded as 0 (Figure 116, section b). Section c shows a mosaic of all the parcel-sized images after the peak

segmentation. Green pixels represent peak number1; brown pixels represent peak number 2; light blue pixels represent peak number 3; blue pixels represent peak number 4. The image in section b (buildings=1, non_building=0) and the segmentation result in section c were multiplied to create the image in section d. This image contains the segmentation pixels inside the buildings. Each pixel in the image that represents a building has a value between 1-4. All the other pixels have a value of 0.
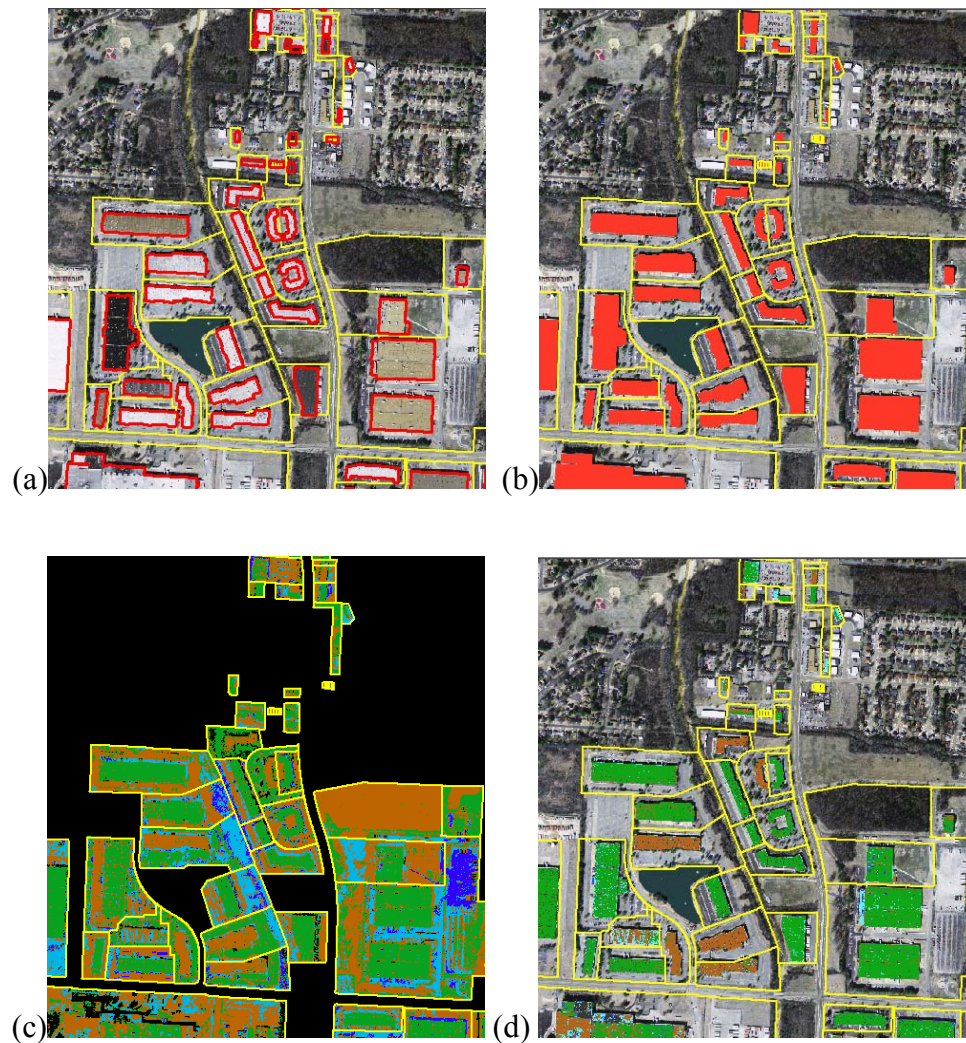
Figure 117 – Peaks analysis of 0.9 square mile in downtown Memphis.

179

Each building (or the symmetric difference between the parcel and the building) was given a zone number to create a zone layer. The zone layer was overlaid on the segment image and summarized for statistics.
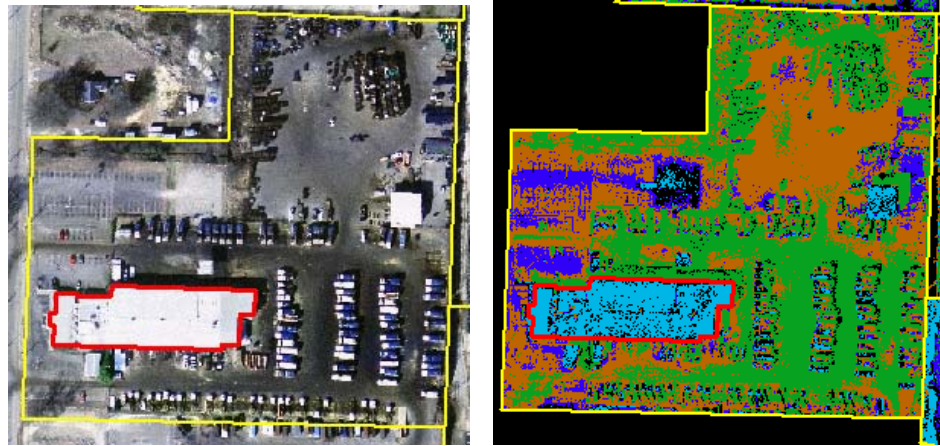
Figure 118 – Peaks analysis – building represented by peak number 3

Figures 118 and 119 provide examples of two buildings that are small relative to the parcel area. As a result, the third (figure 118) and the forth (figure 119) peaks represent the buildings. Figure 118 illustrates the intermediate result that includes all the segmentation of all peaks in the image space overlaid by the building footprint (red).
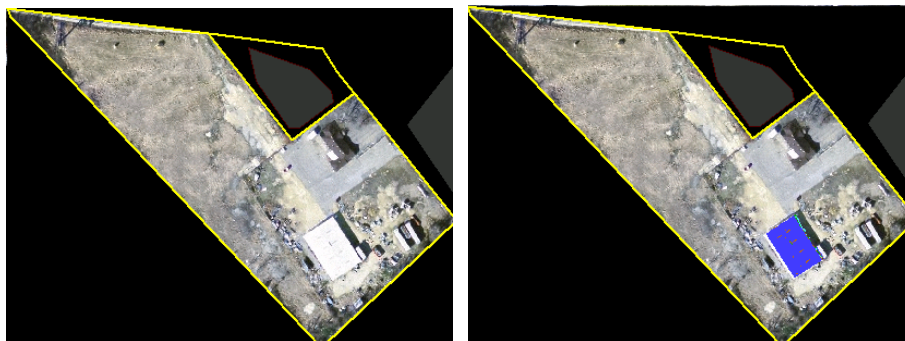
Figure 119 – Peaks analysis – building represented by peak number 4

Figure 119 illustrates the final step of the analysis that includes only the segmentation result inside the building.

The zones define the building area as well as the area around the building. Both areas were analyzed and summarized for statistics. Two detailed examples of the summary are provided below and illustrated in figure 120 and figure 121.
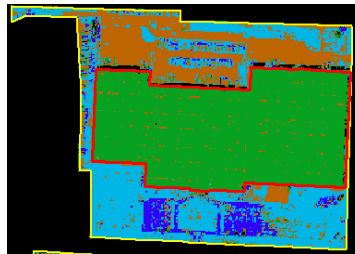
Figure 120 - Peaks analysis: peak segmentation into the image. Building polygon in red.

Figure 120 shows a segmentation result of a parcel with one large building. All four peaks are segmented into the image. The building is mainly represented by peak 1 to create a large continuous segment. The summary result for this parcel is provided in table 6. Each row in the table represents a zone (the building or the area around the building). The area that each peak covers within a specific zone is divided by the total number of segmented pixels within the zone (does not include 0 value pixels – not segmented!), to calculate the percent values.

Table 7 – Peak analysis for building in figure 127. Percent of the peak area within the image

	peak 1	peak2	peak3	peak4
The building	95.5%	3.6%	0.67%	0.2%
Around the building	1.3%	31.9%	45.4%	12.8%

Clearly, peak 1 represents the building and does not add significant noise pixels around the building (pixels that do not represent features around the building). Peaks 2, 3 and 4 do not contribute significantly to the building segmentation (mostly peaks 3 and 4), but segment the area around the building.
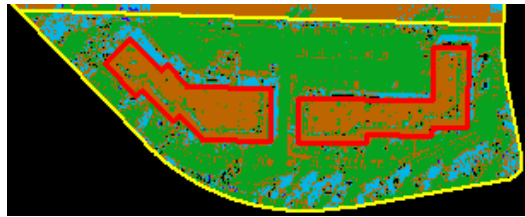
Figure 121 – Peaks analysis: 2 buildings peak segmentation into the image. Building polygon in red. Percent of the peak area within the image

Figure 121 depict a parcel with two buildings. Unlike the building in figure 120 that was represented by peak number 1 (green pixels), the buildings in figure 121 are represented mostly by peak number 2 (brown pixels).   The peak analysis statistics calculations are provided in table7.

Table 8 – peak analysis for buildings in figure 128

	peak 1	peak2	peak3	peak4
The left building	8%	85.1%	3.9%	2.9%
The right building	9.1%	85.9%	1.6%	3.3%
Around the building	59.6%	5.5%	19.3%	5.3%

The statistics calculations in table 7 show that peak 1 and peak 2 represent the majority of the pixels inside and outside the building. Peak number 1 provides 60 percent of the area around the building, while peak 2 represents the majority (about 85%) of the pixels inside the buildings. In this case, peak number 1 provides a more considerable addition to the building (about 9%) compared with peak 3 and peak 4 combined (about 5%).

Similar peak analysis was performed for the commercial, residential and high-rise parcels. Section 4.2.5.1 provides the calculation results for the commercial parcel, and section 4.2.5.2 and 4.2.5.3 provide the results for the residential and high-rise parcels respectively.

4.2.5.1 *Number of peaks evaluation for commercial parcels*

This section presents the statistics results as calculated for 1079 commercial parcels and 1128 buildings in the Memphis, TN downtown area. First, the summary results for the pixels inside the buildings, is provided:

- Within 667 buildings (59%) peak1 is the majority of the pixels (over 50%)

183

- Within 240 buildings (21%) peak 2 is the majority of the pixels

- Within 76 buildings (6.7%) peak 3 is the majority of the pixels

- Within 26 buildings (2.3%) peak 4 is the majority of the pixels

- Within 119 buildings (10.5%) no peak has a majority

Clearly, peak 1 and peak 2 represent the majority (80%) of the buildings. The contribution of each peak to the buildings polygons was calculated as the average percentage of the peak within all buildings polygons. The average values are:

- Peak 1- average 56.6% within every building

- Peak2 – average 25.4% within every building

- Peak3 – average 11.8% within every building

- Peak4 – average 6.1% within every building

The average values emphasize the contribution of the first two peaks to the buildings polygons. A count of the number of pixels that each peak contributes to the building region is provided. The percentage of the pixels from the entire buildings area is also calculated:

- No data - 469269

- Peak 1 – 48981098 pixels (76.1%)

- Peak2 – 9514220 (14.8%)

- Peak3 – 3828310 (5.9%)

- Peak4 – 1942333 (3%)

It is clear from the statistics that the first and the second peaks contribute the most to the building area (90% of the pixels). Besides the total number of pixels, those peaks constitute to the majority of pixels within 80% of the buildings and represent a total of about 80% of the pixels within each building (on average).

The statistics was calculated for the area around the building within the 1079 parcels:

- Within 328 parcels (30.4%) peak1 is the majority of the pixels (over 50%)

- Within 190 parcels (17.6%) peak 2 is the majority of the pixels

- Within 35 parcels (3.2%) peak 3 is the majority of the pixels

- Within 14 parcels (1.3%) peak 4 is the majority of the pixels

- Within 512 parcels (47.4%) no peak is the majority

Clearly, peak 1 and peak 2 represent the majority of pixels in about half of the parcels. The contribution of each peak to the region around the buildings was calculated as the average percentage of the peak within the polygons. The average values are:

- Peak 1- average 33.9% within every parcel

- Peak2 – average 27.1% within every parcel

- Peak3 – average 15.1% within every parcel

- Peak4 – average 7.6% within every parcel

Peak 1 and peak 2 still show a considerable contribution to the regions around the buildings. The average values, however, are not as large as inside the buildings polygons. About half (47.7%) of the regions around the buildings are not represented by a particular peak. Still, based on the average values, it can be asserted that peak 1 and peak 2 provide significant data while peak 4, in particular, does not.

The total number of pixels within each parcel for the region outside the buildings polygons was calculated:

- No data – 16745093 (11.1%)

- Peak 1 –49259386 pixels (32.7%)

- Peak2 – 46375369 (30.8%)

- Peak3 – 26106998 (17.3%)

- Peak4 – 11969122 (7.9%)

Compared to the buildings polygons, there is a larger number of "No data" pixels, that represent pixels that were not segmented as part of any peak. Peak 1 provides a lower percentage of pixels around the buildings compared to inside the buildings polygons. Peak 4, in particular does not add significant percentage of pixels to the region.

The above calculations show that the commercial buildings are dominant features within the parcel. For the majority of parcels (80%), the first or second peaks represent the pixels of the building roof. Those two peaks also provide 60 percent of the "noise" pixels around the buildings. Peak 3 and peak 4 contribute significantly to about 10% of

the buildings and add about 25% of the "noise" around the buildings. It may be beneficial to use only peak 1 and peak2 for the extraction of commercial buildings. Removing peak 3 and peak 4 reduces the extraction rate by about 10% but also reduces the number of candidate segments and the significant effort and time of manual post-processing. The use of peak number 4 in particular should be considered due to the low contribution of the peak to the overall extraction process.

4.2.5.2 Number of peaks evaluation for residential parcels

The peak analysis was performed for similar number of residential parcels and buildings. The statistics for 1138 buildings was calculated as follows:

- Within 290 buildings (25.5%) peak1 is the majority of the pixels (over 50%)

- Within 64 buildings (5.6%) peak 2 is the majority of the pixels

- Within 27 buildings (2.3%) peak 3 is the majority of the pixels

- Within 2 buildings (0.2%) peak 4 is the majority of the pixels

- Within 755 buildings (66%) no peak is the majority of the pixels

Two of the above values may explain the difficulty of automatically extract residential buildings. Peak 1, that was shown to provide a majority in almost 60% of the commercial buildings, has a majority of pixels in only 25% of the buildings. More over, peak 1 and peak 2 combined, have a majority of pixels in only 30% of the buildings. In 66% of the buildings there is no clear majority to any class. That indicates a high degree

of confusion within the building area segmentation and explains the low extraction rate of residential buildings. That finding is consistent with the result of the residential buildings classification analysis in section 4.2.3.3.

The contribution of each peak to the region inside buildings was calculated as the average percentage of the peak within the polygons. The average values are:

- Peak 1- average 36.1% within every building

- Peak2 – average 18.4% within every building

- Peak3 – average 12.2% within every building

- Peak4 – average 7.2% within every building

The total number of pixels for each peak inside the buildings was also calculated:

- No data – 549828 (26%)

- Peak 1 –783871pixels (36.8%)

- Peak2 – 394679 (18.5%)

- Peak3 – 249794 (11.7%)

- Peak4 – 150768(7%)

The calculated statistics for the residential buildings suggest that peak1 is not as dominant compared to the commercial buildings (36.8% vs 76.1%). Peak 3 and peak 4 maintain the low contribution to the buildings polygons and probably add to the confusion during the extraction. About 25% of the pixels within the buildings polygons

are flagged as "No data", which suggests that about a quarter of the pixels inside each building polygon are not classified. That significant lack of segmentation emphasizes high degree of signature confusion of the roof of residential buildings.

Statistics was calculated for the regions around the buildings:

- Within 219 parcels (21.1%) peak1 is the majority of the pixels (over 50%)

- Within 23 parcels (2.2%) peak 2 is the majority of the pixels

- Within 5 parcels (0.5%) peak 3 is the majority of the pixels

- Within 1 parcels (0.1%) peak 4 is the majority of the pixels

- Within 789 parcels (76%) no peak is the majority of the pixels

The contribution of each peak to the region around the buildings was calculated as the average percentage of the peak within the polygons. The average values are

- Peak 1- average 30.8% within every parcel

- Peak2 – average 15.7%   within every parcel

- Peak3 – average 12%   within every parcel

- Peak4 – average 8.5%   within every parcel

The total number of pixels for each peak around the buildings was also calculated:

- No data – 3307899 (53.8%)

- Peak 1 –1476607 (24%)

- Peak2 – 656926 (10.6%)

- Peak3 – 422850   (6.8%)

- Peak4 – 280783 (4.6%)

The large number of pixels that are not segmented ("No data") may indicate that there are many small patches and island segments within the parcel. Segments around the buildings are classified as noise, and there is no reason to expand the number of peaks in order to include them. The statistics suggest a great degree of confusion within the residential extraction process, which directly affects the extraction rate. A building polygon that does not have a majority of pixels with the same value (same peak) has a low probability for successful extraction. As indicated in section 4.2.3.3., automatic extraction of residential buildings poses a great challenge. The findings in this section help to illustrate the nature of the residential scene and help to explain the analysis challenges and low extraction rate.

4.2.5.3   Number of peaks evaluation for high-rise parcels

Similar to the commercial and residential peak analysis, statistics was calculated for the 49 high-rise parcels and 60 buildings.

- Within 12 buildings (20%) peak1 is the majority of the pixels (over 50%)

- Within 2 buildings (3.3%) peak 2 is the majority of the pixels

- Within 6 buildings (10%) peak 3 is the majority of the pixels

- Within 2 buildings (3.3%) peak 4 is the majority of the pixels

- Within 38 buildings (63.3%) no peak is the majority of the pixels

Clearly, for a large percent of the high-rise building (over 63%), no peak dominates the roof top area. That value explains the high degree of confusion during the segmentation of high-rise buildings. It can explain the low extraction success rate for high-rise structures.

The contribution of each peak to the regions inside buildings was calculated as the average percentage of the peak within the polygons. The average values are:

- Peak 1- average 27.3% within every building

- Peak2 – average 15.2% within every building

- Peak3 – average 15.5% within every building

- Peak4 – average 7.6% within every building

Total number of pixels representing each peak, and their percent within the building, were calculated:

- No data – 482439.000 (27.3%)

- Peak 1 –424112 (23.9%)

- Peak2 – 239765 (13.5%)

- Peak3 – 466207 (26.3%)

- Peak4 – 156721 (8.9%)

The statistics calculations for the buildings polygons illustrate the complicated building extraction task. There is a similar average contribution of peak 1 (24%), peak 3 (26%) as well as the not-segmented pixels within the buildings. This mixture of pixels within a building prevents a successful identification and extraction of the building footprint. The statistics emphasize and are consistent with the low extraction rate for the high-rise buildings, as presented in section 4.2.4.

The same statistics were calculated for the area around the building for all 49 high-rise parcels:

- Within 9 parcels (18.4%) peak1 is the majority of the pixels (over 50%)

- Within 7 parcels (14.3%) peak 2 is the majority of the pixels

- Within 2 parcels (4%) peak 3 is the majority of the pixels

- Within 0 parcels (0%) peak 4 is the majority of the pixels

- Within 31 parcels (63.2%) no peak is the majority of the pixels

The contribution of each peak to the regions around the buildings was calculated as the average percentage of the peak within the polygons. The average values are:

- Peak 1- average 34.6% within every parcel

- Peak2 – average 25.8% within every parcel

- Peak3 – average 12.4% within every parcel

- Peak4 – average 5%   within every parcel

Finally, the number of pixels and their percentage was calculated:

- No data – 389972 (14.8.8%)

- Peak 1 –1476607 (37.7%)

- Peak2 – 656926 (23.4%)

- Peak3 – 422850   (17.4%)

- Peak4 – 280783 (6.7%)

The statistics for the high-rise buildings and the regions around the buildings emphasizes the great challenges of extracting high-rise buildings from aerial images. In order to achieve successful extraction result, the desirable result of the segmentation should be continuous, large segments that represent the buildings, and small insignificant, easy to eliminate segments around the buildings. The significant contribution of multiple peaks to the building polygons suggests that the segmentation result would not produce clear and obvious building footprints. The building region are represented as multiple smaller segments and require manual post processing.

Consistent with the findings for the residential and commercial parcels, the forth peak contributes the least information to the extraction process. In the case of high-rise buildings, the average contribution of the peak is minimal (5% on an average), and the peak pixels do not represent a majority for even one building polygon. Elimination of the forth peak would probably not considerably change the extraction results.

4.2.6   Using Parcel Setbacks in the Analysis

The following section examines the use of parcels setbacks to isolate the built area within the parcel. Setbacks identify the approved building area within a parcel by defining distances from the parcel boundaries. Setbacks may vary between zoning categories and subcategories (different residential zoning areas). Moreover, the same zoning category might have different setbacks in different regions. For example, Commercial parcel within the Central Business District (CBD) in Memphis has no setbacks, while commercial parcel within a local commercial district has defined setbacks (See ordinance Appendix J). Due to the large number of categories and subcategories, one residential and one commercial zoning categories were tested to evaluate the possible affect of setbacks on the extraction process.

4.2.6.1   Setback analysis – Residential parcels

The Memphis region ordinance includes 4 different zoning categories for single-family houses parcels (R-S15, R-S10, R-S8, R-S6). The minimum setback required for those categories is 5 ft setback on the side of the parcel. In order to evaluate the feasibility of using setback to better isolate the building prior to the extraction, a 5ft buffer was created for all selected residential parcels (OCC_CLASS column in the database equals "RES1").

1761 residential parcels were extracted from a mosaic orthophoto image. The parcels layer original boundaries intersect 238 (13.5%) houses. Figure 122 illustrates four

residential parcels (yellow polygons) with four houses (green polygons). A 5 ft setback inner buffer was defined within each parcel. The buildings polygons that intersect the parcel and setback outlines are highlighted.

Figure 122 – Setbacks analysis –single family houses. (Green polygon) digitized house outline. (Yellow) parcel. (Pink) 5ft buffer inside the parcel. Building crossed by the parcel line is highlighted.

A total of 364 houses are crossed by the outline of the 5 ft buffer. This is an increase of 126 houses or 53% over the 238 houses crossed by the original outline. Besides the large addition of houses crossed by the minimum-setback buffer, the minimum-setback buffer does not cover a significant portion of the parcels, hence would probably not have a significant affect on the result of the automatic process. A methodology that takes into consideration the different setbacks on the front/side and rear of the parcel by possibly including a road layer may be implemented and better isolate the building. Since three of the residential categories have 5ft setback for the side and 20 ft setback for the rear and front, a 20 ft setback buffer was also applied to the parcels (Figure 123).

Figure 123 – Setback analysis. 20 ft inner buffer (orange)

The residential parcels dataset has an average parcel area of 1961.8 sqft. A manual visual observation of the 5ft and 20 ft buffer reveals that the buffers do not significantly reduce the number of objects that would be segmented around the building. As a result, the setbacks do not decrease the processing or the post processing time and effort. The setbacks were shown to increase the number of buildings that intersect with the boundary (5ft inner buffer provides an increase of 50%). An intersection between a building and a boundary indicates that sections of the building would not be segmented or extracted. That is a major limitation of using setbacks, and the degree of the phenomenon has to be evaluated on a project (region) basis. This approach may be re-evaluated in regions where the setback area constitutes a larger percentage of the parcel and does not increase the number of buildings that intersect the boundary of the parcel.

4.2.6.2   Setback analysis- commercial parcels

The Memphis ordinance includes several commercial and office zoning categories. The C-P and C-L required zoning setbacks were used to define the buffers for the testing. A10 ft minimum setback for the side, rear and front of the parcels was defined. From the parcel layer, commercial parcels with occupancy of "COM2", which represents warehouses in the Memphis area, were extracted. That category includes 399 parcels.

105 buildings were crossed by the outlines of the 399 commercial parcels boundaries before the buffer was incorporated. 200 buildings were crossed by the outline of the 10ft buffered parcels (pink polygons, figure 124). This indicates a significant increase (90%) in the number of buildings that intersect with the polygon boundary. For that parcels category, the front setback is defined as 30ft and the rear setback as 15ft. In order to evaluate the isolation of the building using setbacks, a 30 ft setback was also defined for the parcels. Similar to the residential parcels, the setback buffers, even the maximum setback values buffers, do not significantly reduce the number of possible extracted objects within the parcel. However, the setbacks inner buffers add extensively to the number of buildings that intersect the buffered outline. 328 buildings were crossed by the outline of the 30 ft buffer, which indicate an increase of 223 buildings (or 212%) compared to the buildings that intersect the original parcels layer. The uniform buffer that was used in the testing may be replaced by the specific setbacks for the side/front/rear of each parcel. Based on the commercial parcels setback analysis, setbacks do not seem to better isolate the built area in the parcel without increasing the probability of eliminating sections of the building.

Figure 124- setback analysis – commercial buildings. (yellow) parcels (pink) 10 ft buffer

(orange) 30 ft buffer

4.2.6.3   Setback analysis – high-rise parcels

High-rise buildings, particularly due to relief displacement, extend the boundary of the parcel. 34 out of the 60 high-rise buildings are crossed by the outline of the parcels boundary. The majority of the buildings are office buildings, and as such, a minimum setback of 5ft has been used for the inner buffer.
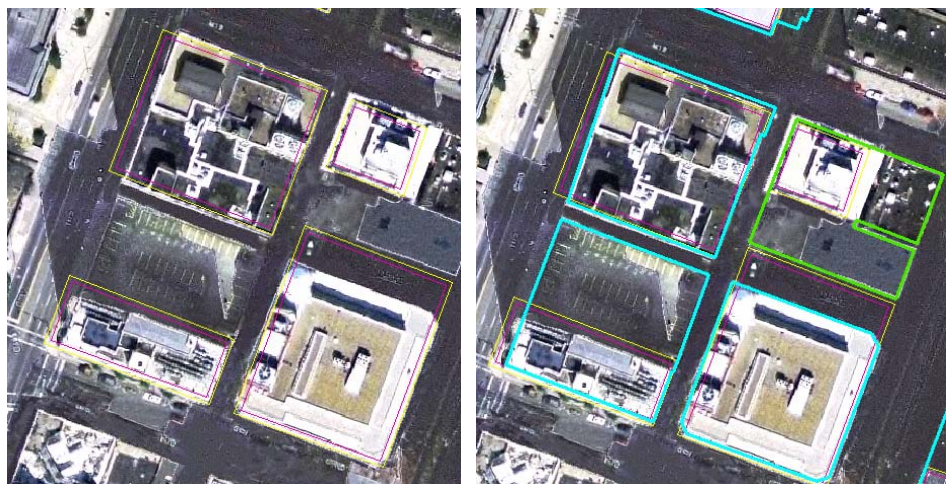
Figure 125 - Setback analysis – high-rise parcels. (yellow) parcels (pink) 5ft buffer

The 5ft inner buffer within the high-rise parcels did not change the number of buildings that intersect the boundary. As can be visually seen in the examples in figure 125, the buffer do not reduce the number of potential segmented objects. For many high-rise parcels, the building area covers most of the parcel area. As a result, the setbacks do not seem to isolate better the built area within the parcel.

<u>4.2.7</u>   <u>Ratio of Building Area to Parcel Area Evaluation</u>

The following section examines the ratio between the building area and the parcel area. The ratio of buildings that were successfully extracted as well as building that were not extracted is presented in order to identify a possible connection between the ratio and the probability to extract the building.

4.2.7.1 Building to parcel ratio - commercial

The ratio between the building area and the parcel area was calculated for all 1128 commercial buildings.

- 353 have a ratio less then 10%

- 235 have a ratio between 10-20%

- 449 have a ratio between 20-50%

- 70 have a ratio between 50-100%

- 21 buildings have a ratio greater then 100%

For the buildings that were not extracted, the average ratio is 11.6%, ranging from 0.47% (3590/671234.9) to 247% (32566/13142.4).

- 109 out of 169 (64.5%) undetected buildings have a ratio less than 10%

- 143 out of 169 (84.6%) undetected buildings have a ratio less than 20%

- 164 (97%)– less than 50% and 165 less than 56

- The rest (3) are over 100%

Office complexes that were tested within the commercial category tend to expand over multiple parcels, have a ratio greater then 100% and were complicated to extract.

For the successfully extracted buildings, the average ratio was calculated as 32.1% ranging from 0.6% (9217.8/1542941.8) to over 100% (4142.4/2110). The building with the ratio over 100% was extracted in two different parcels. Clearly, the average ratio of the extracted buildings (~32%) is higher than the average for the buildings that were not extracted (~11%).

- 244 out of 936 (26.1%) detected buildings have a ratio less then 10%

- 445 out of 936 (47.5%) detected buildings have a ratio less then 20%

- 870 out of 936 (92.9%) detected buildings have a ratio less then 50%

- 915 out of 936 (97.8%) detected buildings have a ratio less then 100%

- The rest were extracted as section within more then one parcel.

The ratio calculations show that the majority of the extracted buildings (92.9%), as well as the majority of the un-extracted buildings (97%), have a ratio less than 50%. However, 64.5% of the buildings that were not extracted and only 25.5% of the extracted buildings have a ratio less than 10%. Moreover, almost 85% of the buildings that were not extracted have a ratio less then 20% compared to 47.5% of the extracted buildings. It can be inferred that buildings with "building to parcel ratio" lower then 10% (hence, not dominant within the parcel) have a lower probability of being extracted. That conclusion is directly correlated to one of the initial project assumption. The assumption suggests that within a parcel region, a building is dominant enough to be extracted. Of course, the extraction result also depends on the number of other distinct features in the parcel (each feature type represented as a peak in the histogram).

4.2.7.2 *Building to parcel ratio - residential*

For the residential houses ratio calculation, 949 houses within the same image mosaic were selected. As concluded in section 4.2.3, the automatic extraction process shows insufficient results for residential parcels. The evident factors that prevent the extraction were identified as trees, shadows and roof construction. This section examines the ratio between the house area and the parcels area as follows:

- 224 (23.6%) buildings have a ratio less then 10%

- 527 (55.5%) buildings have a ratio less then 20%

- 799 (84.1%) buildings have a ratio less then 50%

- 816 (85.9%) buildings have a ratio less then 100%

- 133 (14%) buildings have a ratio over 100%

The mean ratio for the residential buildings is 97.9% with range of 0.09% – 2719.2%. It should be noted that all the buildings with a ratio over 100% are either townhomes or condominiums. As expected, the ratio of these structures has a significant affect on the mean ratio. Those structures expand over multiple parcels and are very difficult to detect. It is recommended to merge all adjacent condominium and townhome parcels prior to any extraction attempt.

The percent of residential buildings that have a ratio less then 10% and less then 20% is similar to the percent of the commercial buildings that were extracted (23.6 vs 26.1/ 55.5 vs 47.5). These ratio percentages were shown previously to successfully differentiate between the extracted and un-extracted commercial buildings. Hence, for residential buildings it is suggested that the ratio between the buildings/house and the parcel is not a significant factor that can prevent the extraction. The environmental and architectural characteristics have a greater affect on the extraction outcome.

4.2.7.3 Building to parcel ratio – high-rise

The ratio between the building area and the parcel area was calculated for the 60 high-rise buildings as follows:

- 15 buildings have a ratio greater then a 100%

- 13 buildings have a ratio between 50-100%

- 32 buildings have a ratio less then 50%

The 10 buildings that were successfully extracted range between 4-88% ratio of building to parcel area, with an average of 35%. Buildings that expand beyond the boundary of the parcel in a rate of 200 to even 1000 percent (building 10 times bigger then the parcel), obviously, can not be successfully extracted. There were 9 buildings with that extreme ratio within the 60 building sample. Similar to the residential buildings conclusion, the building to parcel ratio does not appear to be a significant factor that accounts for the poor extraction. 45 of the 60 buildings cover more then 50% of the parcel, while only 10 were extracted successfully. Factor such as shadows, multi-levels and complex roof signatures appear to have the most considerable affect on the extraction result (section 4.2.4)

4.2.8   Testing Manual Digitization

The manual portion of the testing entails manually identifying and digitizing buildings from 1ft imagery. The goal is to compare the automatic and manual results as well as two different manual digitization methodologies. The testing comprised of three steps: (1) complete manual digitization of buildings from a full aerial image scene; (2)"cleaning" the automatic extraction result; (3)complete manual digitization of buildings from pre-cut parcel-sized images. The steps were evaluated for commercial and residential areas separately and do not include the second step for residential buildings (due to the poor performance of the automatic process for those buildings). The user logs the time it takes to perform each of the steps.

The first step provides the reference work, as currently performed when building footprints are manually digitized from images. The result is compared with the "clean"

result of the automatic process and includes a quantitative evaluation. In addition, the result of the first step is compared to a manual digitization from parcel-sized images. The comparison allows evaluating the potential savings of time and effort by digitizing from parcel-sized images rather then the entire full image scene.

For the manual testing, 50 commercial and residential parcels were selected (figure 59 – selected parcels are highlighted). In order to compare the manual effort with the automatically extracted result, the subset of selected parcels has to account for the types of parcels used in the automatic process. As a result, the selected parcels comprised of multi-building parcels (12 parcels); parcels in which buildings were not automatically extracted (7 parcels); non-rectangular buildings (20 parcels); small parcels (14 parcels) and a variety of grey level color buildings. The residential parcels were selected to include houses that can be easily and clearly identified and houses that have interfering objects and shadows in their surroundings.
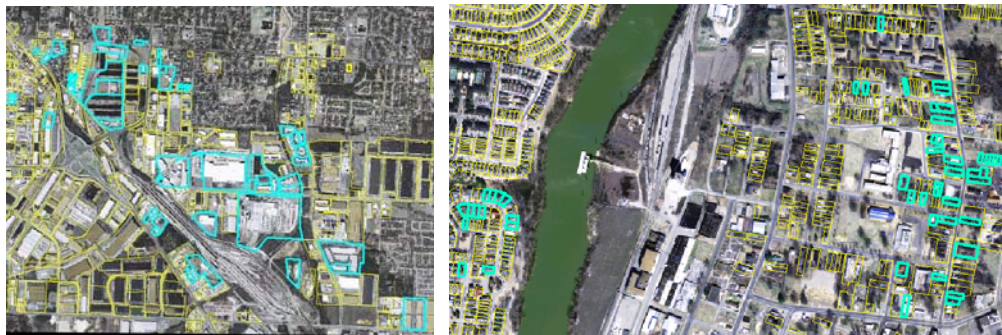
Figure 126 – 50 commercial (left) and 50 residential (right) parcels selected for manual digitization

During the digitizing process, the parcel layer, as illustrated in figure 126, is overlaid on the image. The results are presented in table 8 below. The commercial parcels

include 50 parcels with 84 digitized structures, comprised of 825 points (built structures range from 4 corner buildings to 29 corner compound building). The residential buildings include 50 structures comprised of 334 points (houses range from 4 corners to 12 corners structures).

Table 9 - Results of manually digitizing building within parcels

Method	Time	Time Per building	Time per corner	Average Difference in Area (%)
Commercial – manually digitizing a full image	42 min	30sec	3.1sec	N/A. this is the reference layer
Commercial – "cleaning" automatic extraction result	18.5min	13.2sec	1.3sec	3.35%
Commercial – manually digitizing parcel-sized images	26min	18.5sec	1.9sec	2.01%
Residential - manually digitizing a full image	16min	19.2sec	2.9sec	N/A – this is the reference layer
Residential - manually digitizing parcel-sized images	11min	13.2sec	2.0sec	7.01%

Table 9 presents the results of the manually digitized buildings for the scenarios detailed above. The table includes several quantitative comparisons between the methods. First, three different time measures are provided: the overall time for accomplishing the task is provided; the time is divided by the number of buildings to allow better evaluation of the average time it takes to complete the task per building; time is divided by the number of corners. The last column provides a measure of discrepancy between the result of a method and the digitized buildings layer. The reference layer for the comparison is the result of digitizing buildings from the full image scene. The layer was selected as the reference to allow consistency for user decision making during the

process as well as a measure of reference to the currently used method by the industry to digitize buildings. The selected method is not confined or restricted by the parcel layer. Hence, it allows an analysis of the advantages and disadvantages of using parcels for the manual building extraction. The area discrepancy is measured in percent and is calculated by adding all the "symmetric differences" in area between the two layers, and then dividing by the total building area in the reference layer.

The first three rows in the table compare three different methods for extracting commercial buildings. The first row provides the results for manually digitizing buildings on an entire image scene. The second row provides the results for "cleaning" the automatic process result. "Cleaning" refers to eliminating segments that are not buildings, merging segments where appropriate, moving or deleting vertices and fully digitizing buildings that were not extracted. The cleaning is performed on the vector polygon layer overlaid on the full image. The third row provides the results for manually digitizing buildings from parcel-sized images.

Based on the testing, the most time consuming method (42 min) to manually extract buildings is to digitize the structures from an entire image scene. There is almost a 40 percent reduction of time when the digitizing is performed on parcel sized images (26min versus 42 min). There is about 56 percent reduction in time for "cleaning" the automatic result (18.5min versus 42min). The significant reduction in time between manually digitizing from a full image versus parcel-sized images can be attributed to the constant miscellaneous zooming (in and out) and panning through the image. Figure 127 portraits the substantial differences between the commercial structures. The image on the left illustrates the great variety of sizes that represents commercial buildings. Moreover, the

206

residential image on the right emphasizes the difference between commercial and residential scenes. While the commercial parcels and buildings deviate greatly in size, the residential buildings are fairly uniform. As a result, while digitizing buildings from a commercial scene, the user has to constantly zoom in and out and pan throughout the image. Those actions contribute considerable time to the overall digitizing task.

Figure 127 – (Left) commercial buildings and parcels show a wide variety of building sized; (Right) residential buildings and parcels fairly uniform in size.

The least amount of time was attributed to "cleaning" the automatic extraction process. The considerable time difference is due to the fact that many of the buildings were already extracted and some did not need any correction. As pointed by the user, the actions of merging segments and moving or deleting vertices require more time then simple digitizing. Hence, even though not all the points had to manually extracted, there is only about 50 percent savings in time between fully manually digitizing and cleaning the result. That difference, however, can become significant when extrapolating to larger regions. The 56 percent difference for cleaning the automatic result and 38 percent for digitizing from parcel sized images, interpret into 78.3 and 53.3 working hours

respectively for 10,000 buildings. That means that in order to extract 10,000 buildings from aerial images, we may save 78.3 hours by cleaning the result of an automatic process and 53.3 hours by digitizing from the parcel-sized images. The minor discrepancies in areas between the different methods are largely insignificant as depicted in figure 128. The result of manually digitizing the parcel-sized images is better (2.01%) then the clean result of the automatic process (3.35%), since some artifacts of the automatic process and generalization were not corrected (see right image on figure 128).
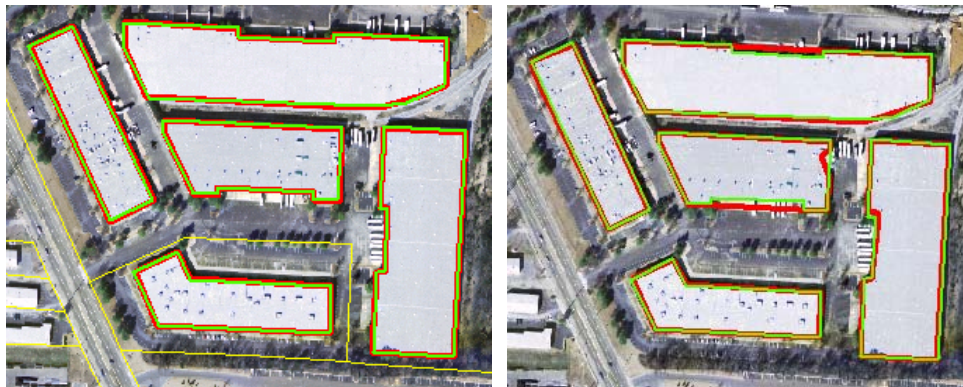
Figure 128 – (Left) green - buildings digitized on a full image; red – buildings digitized on parcel-sized images. (Right) green - buildings digitized on a full image; red – "clean" result of an automatic process

The 2.01% difference is attributed to two buildings that were cut by the parcel boundary. Figure 129 illustrates one of the two buildings with 25.2% discrepancy in area. The user was able to take advantage of sections of the building within the parcel boundary to assess the location of sections that were not available outside the parcel boundary. Without the buildings that had sections outside the parcels boundaries, the method of digitizing buildings from parcel sized images yields an area discrepancy of 1.5 percent.

Figure 129 – (green) building footprint as digitized on a full image (red) building footprint as digitized on a parcel sized image (yellow) parcel boundary

The forth and fifth rows in table 8 provide the results for the residential buildings. The 50 residential parcels include only 50 buildings with significantly smaller number of corners compared to the commercial parcels. Hence, the considerable difference in manually digitizing the commercial and residential buildings. Moreover, as depicted in figure 60, the size of residential parcels is fairly uniform and does not require as much zooming in and out during the digitizing process. The results show a consistent significant reduction in time (32%) between digitizing from a full image (16min) versus digitizing from parcel sized images (11min). The major difference between a commercial and residential scene is the degree of decision making required by the user. In a residential scene there are many interfering objects and phenomena, such as trees and shadows that obstruct the building footprint. The user has to invest more time to assess the actual location of the building corner. As a result, even the same user may extract the footprint of the building with great discrepancy. Figure 130 illustrates two such examples, where the same buildings were extracted manually with differences of 46.5 (left image) and 15 (right image) percent in area. Unlike the left image (with the 46.5%

discrepancy), the user was able to take advantage of a more clear shadow area and tree branches to better locate the building footprint on the right image.

Figure 130 – Footprint discrepancy between two manually digitized residential buildings. Green and red represent the building footprints. Yellow represents the parcel boundary.

Figure 131 illustrates the result of manually digitizing residential houses in a clear, free from trees parcels.

Figure 131 – discrepancy between manually digitized residential buildings. (Green) – digitizing on a full image (Red) – digitizing on parcel-sized images.

The manual testing provides several insights regarding the use of parcels for the building extraction process. Parcels can assist to significantly reduce the time and effort required to extract buildings. They may be utilized as part of an automatic process as well as part of a manual extraction procedure. As presented above, digitizing buildings from parcel sized images rather then the full image scene may dramatically reduce the extraction time. To avoid cases where the parcel cuts through a building outline, it is recommended to apply a buffer around the parcel prior to cutting the images.

Chapter 5   CONCLUSIONS

5.1   Recap of the process

The research as described in this document presents a new approach to building extraction from aerial imagery. The integration of several available GIS data sources including both the spatial and attribute data, provide an innovative methodology to solve or better meet the challenges of building extraction.   The research presents a methodology that includes several new approaches to address the complexity of the problem. Those include partitioning the image, adding tax assessor attribute data and including user-one-time-intervention for shadow direction definition.   The analysis procedure entails multiple stages including geometric evaluation of the extracted segments. The importance of the geometric analysis is demonstrated and emphasized in the geometric analysis sections of this document.

The process is tested and evaluated on three distinct and different types of buildings including commercial, residential and high-rise. The evaluation provides a comparison between the three types of buildings as relates to the extraction process. The automatic extraction was found to be more successful in commercial parcels compared with residential and high-rise parcels. Various factors that may affect the result of the extraction process were identified and investigated. The nature of the roof signature has a significant and direct affect on the segmentation, and as a result, on the entire extraction process. Other contributing factors may be the size of the building relative to the parcel, shadow analysis, the slope of the roof, number of buildings within the parcel (multi-

building parcels), building compound architecture and building-specific phenomenon such as relief displacement for high-rise structures (4.2.4.3). Some factors may be possibly resolved (size, shadow, compound) while some (complex roof signature) will remain an obstacle to any automatic extraction process. The unique nature of each type of structure poses specific challenges on the extraction procedure that need to be tackled and further investigated.

The extraction process design, implementation and results are presented in the previous chapters. Initial parcel-based partitioning of the image is performed. The simplicity of the partitioned image provides an easy and efficient way to improve the extraction process. The benefit of the image partitioning as well as the drawbacks are presented. Image partitioning localizes the search to a significantly smaller and limited area and provides the opportunity to make certain simplifying assumptions. It allows a relatively simple histogram analysis that correlates section in the histogram to features in the image space. The image partitioning was examined to evaluate the potential benefit as part of a manual building extraction. The drawbacks of the partitioning, mainly parcels outlines intersecting structures, were evaluated for the three different types of buildings (commercial, residential, high-rise). Recommendations were provided according to the testing and evaluation findings. It was found that the vast majority (about 85%) of commercial and residential buildings resides within the parcels boundaries, and most non-compound buildings that intersect the parcel boundary, maintain a significant portion of the building within the parcel area. As a result, it can be concluded that the parcels layer can be used to localize the search for the buildings by dividing the image. The loss of information is secondary to the obvious benefit to the

extraction process and can be further reduced by applying a buffer around the parcels. On the contrary, the majority of high-rise buildings, particularly due to relief displacement, extend the boundary of the parcel. Successful extraction of high-rise buildings would require applying a buffer around the parcel. Image partitioning was also found to significantly reduce the time and effort required to manually extract buildings. It is shown that digitizing buildings from parcel sized images rather then the full image scene may dramatically reduce the extraction time. It is recommended to apply a buffer around the parcel prior to cutting the images in order to provide the user with the entire building area.

Following the image partitioning, histogram analysis and image segmentation is performed. The histogram analysis and segmentation is a crucial stage within the extraction process. The results establish that a poor segmentation results are correlated with poor building extraction. Different factors may affect the segmentation result and they may vary by type of structure. The significant factors that can affect the segmentation are identified for each type of structure. Moreover, the project evaluates how prevalent the different factors are in preventing the building extraction.

One of the main assumptions of the process suggests that the building is a dominant feature within a localized parcel area and as a result should be manifested as a significant peak within the image histogram. The correlation between the histogram peaks and the features in the image is evaluated to determine the number of peaks that should be considered for the histogram analysis.

The feature segmentation is followed by a series of analyses. The analyses include size analysis, shadow analysis and geometry analysis of the segments. Each step is tested

214

and evaluated for the contribution it provides to the extraction process. The size analysis is based on tax assessor attribute data and was shown to benefit the extraction but appears to be sensitive to extraction artifacts such as segmentation overshoot/undershoot. The shadow analysis requires user input. The analysis verifies the legitimacy of a segment as a 3D, shadow casting feature, such as a building. Shadow segmentation in the image space and the analysis limitations are presented and discussed. The geometry analysis incorporates multiple common geometric shape characteristics. These parameters are individually examined by comparing the behavior of the parameter between building and non-building extracted segments. The value of each parameter is then used to calculate the probability of a segment to be a building. The remaining segments are converted to a vector polygon layer for possible manual post-processing.

The process was tested for different scenarios: full run including all three analysis steps; excluding the size analysis; excluding the shadow analysis; excluding the geometry analysis. Each run provides a way to isolate the affect of certain analysis on the extraction process, hence, evaluate the benefit to the overall process. Results and recommendations are presented and discussed. The commercial building analysis found that the full run, that includes all elimination steps, increases the overall accuracy of extraction and reduces the amount of manual post-processing. The geometry analysis appears to be a significant step in the elimination analysis with considerable affect on the final result. The "no size" and "no shadow" analysis, maintain an important role in the overall analysis and should be integrated in the process to reduce the post processing effort   A detailed examination of residential parcels was performed to evaluate and better understand the unique nature of the residential environment. Two well known and

accepted feature classification methodologies were tested and compared to the histogram-based segmentation process. The complex environment and the factors that prevent successful extraction are discussed at length. It was noted that residential buildings and their environment pose great challenges on any automatic extraction algorithm. The nature of the residential scene emphasizes an attractive living environment. As a result, residential houses are often characterized by interesting architecture and near by trees. Those may appear aesthetic, but make the automatic extraction of the structure complicated or even impossible. When no trees obscure parts of the roof, there is an insignificant, minor or no shadow presence on the roof, then a fair segmentation of the roof is possible using an automatic or semi automatic procedure. As a result, the initial testing of the overall model (automatic) on residential parcels indicated poor extraction results (4.2.3.1). Other extraction procedures (ISODATA, Region growing) did not performed well enough for all buildings to allow further automatic analysis and refinement of the segments and may still require exhaustive manual post processing.

Since discovered by Hu in 1962, the moment theorem was examined, improved and expanded in numerous research projects. The nature of buildings (as man made features that commonly appear in certain shapes) provides the opportunity to use moments to classify the shape of the building. Previous projects developed indices for simple geometric shapes such as ellipse, rectangle and triangle based on moments. This project presents a method to identify and distinguish between "I" and "O" shape features and can be used for building shape identification. The method is tested and evaluated on multiple "clean" and "noisy" segments and proved to be successful.

The research identifies distinct as well as similar properties between the different types of building. A separate section is devoted for each type of building to discuss the distinct characteristics of the building and the environment, as relates to the extraction process. General factors are examined within the same chapter for all types of buildings. Those include building and parcel outline intersection, parcel setbacks and investigation of the number of peaks in the histogram.

The contribution of the research to the advancement of science within the GIS, building extraction and shape recognition disciplines is discussed in the following section.

5.2   Contribution to the Domain

This section evaluates the contribution of the project to several research domains. The evaluation includes contribution to the industry and academic research that demands building inventory, contribution to the automated feature extraction effort within the photogrammetry and remote sensing discipline and contribution to both image processing and photogrammetry by introducing the moment theorem towards building extraction and building shape identification. Each of the following section investigates the contribution of a specific section to different domains.

5.2.1   GIS and Imagery Integration

The research methodology as presented introduces a new overall approach to building extraction. The integration of GIS and remote sensing sources as presented has not been implemented and tested as an entire approach. The research integrates several simplification algorithms to alleviate the complex extraction process. Image partition

217

techniques have been researched and developed within the computer science community for diverse applications, including feature extraction. Clearly, an easy and effective method to partition the image prior to the extraction can dramatically simplify the entire procedure. The research project tests and evaluates image partitioning using a parcel layer. Parcels were shown to simplify the extraction process while maintaining minimum loss of information. About 15% of the commercial and residential houses intersect the parcel boundaries as well as about 50% of the high-rise buildings. Buildings that cross parcels lines were found to maintain a significant portion of the buildings within the parcel area. Specific pre-processing recommendation for certain types of structures such as office compounds and townhouses are discussed as a result of the testing. General recommendations for reducing the loss of information such as applying a pre-processing buffer are also discussed. The research provides initial investigation of parcel setbacks as a method to isolate the built structure better within the parcel area. Using setbacks did not appear to efficiently isolate the building in the parcel without increasing the probability of eliminating sections of the building. The research concludes that the parcels layer can be used to localize the search for the buildings by dividing the image. The loss of information is secondary to the obvious benefit to the extraction process and can be further reduced by applying a buffer around the parcels. The research project acknowledges the unique nature of spatial data by applying a-priori GIS data to the image partitioning task. Hence, the project clearly contributes to the computer science work aimed at image partitioning for feature extraction (Berretti et al, 2000; Carson et al, 2002; Yixin and Wang, 2002; Jia and Wang, 2003; Chen and Wang, 2004) and specifically for

building extraction (Sahar and Krupnik, 1999; Sohn and Dowman, 2003; Jiang, et al, 2008; Zhengjun, et al, 2008).

5.2.2   Reducing Manual Digitizing Effort

The photogrammetry and remote sensing field has been attempting to develop automated and semi-automated approaches for feature extraction and in particular building extraction in the last 20 years. Today, we still do not have an "accepted" methodology to extract buildings from aerial imagery and therefore digitize those features manually. The research project presents a detailed evaluation of the testing results and includes a comparison between the automated and manual extraction procedures. The results suggest an advantage to automatically extracting commercial buildings while maintaining manual digitizing for other structures. The research offers a significant contribution to manual extraction effort by presenting a method that can reduce time and cost for the manual digitizing process. Using parcels to localize the extraction area and eliminate user extraneous operations was shown to be extremely efficient. This contribution is significant as efficient building extraction procedures are required to inventory development, day to day management of cities and counties and for more complex application such as evaluating damage during an earthquake. All those applications can greatly benefit from a methodology or procedure that can produce large percentage of the building inventory or at least considerably reduce the effort.

5.2.3 Multiple technique integration

The research methodology integrates readily available sources and image analysis techniques in an innovative manner. The research focuses on extraction from an image product (1ft orthophoto) that is widely acceptable as a standard imagery source. The integration of GIS data (parcel attributes), shadows and geometry analysis expands many research efforts that encompass some of those techniques (Huertas and Nevatia, 1988; Irvin and McKeown, 1989; Kolhe, Plumer and Cremers, 2000; Duan, Prinet and Lu, 2004; Khoshelham, 2004;Wei, Zhao and Song, 2004; Müller and Zaum, 2005). The research testing is geared towards evaluating the individual benefit of each analysis to the overall extraction procedure. The suggested benefit of the attribute data, shadow analysis and geometrical analysis is provided in details. The geometry analysis appears to be a significant step in the elimination analysis with considerable effect on the final result. The attribute data and shadow analysis sustain an important role in the overall analysis and should be integrated in the process to reduce the post processing effort.

5.2.4 Testing Different Structure Types (Multiple Land Use)

Past research projects have tended to concentrate their effort on specific building structure types or geometric properties (Kim et al, 2004; Tang et al, 2004). The research project presented was not designed to extract specific types of buildings such as high-rise or commercial and was not constrained to a specific geometric shape. Several geometric parameters were used to differentiate between building and non-building structures. Each parameter was evaluated and possible shape identification tools provided for future research. The methodology was tested on three types of parcels and buildings including

220

commercial, high-rise and residential. The testing suggests successful results for commercial buildings. A 79% success rate for commercial buildings is comparable with previous research that reported 72.7% extraction rate with 58.8% accuracy (Jin and Davis, 2004) and 64.4% (Lee et al, 2003). Jin and Davis (2004) suggest that any extraction rate over 70% is considered a success and Fraser et al (2002) concludes that 15% of the buildings can not be manually identified in the (IKONOS) image. The commercial buildings testing, includes investigation of factors that may complicate the analysis. Those factors include roof construction, roof signature complexity, building size, shadows and building compounds. Each factor is quantitatively examined and evaluated. The automatic procedure did not extract as successfully residential and high-rise buildings. The results provided an opportunity to examine the distinct nature of each structure type and identify the factors that prevent the extraction. The environmental and architectural aspects of the buildings are investigated including a quantitative assessment, in a manner that was not previously attempted. The extraction of residential structure in particular, was examined by comparing three different extraction approaches (automatic procedure, region growing, Isodata classification). The complexity of the residential scene was established and possible automatic extraction discussed. The results are consistent with current research projects that concentrate on extracting the roof as multiple different polygons (Avrahami et al, 2008). High-rise structures pose distinct challenges on the extraction process and have not been the target of many previous research projects. This research provides specific assessment with regards to automatically extracting high-rise buildings from aerial images. The unique

characteristics of high-rise buildings as portrayed in aerial imagery and complicate the extraction process are identified and discussed.

5.2.5   Building Shape Recognition

The work of Hu (1962), Rosin (1999) and Rosin (2003) is used and implemented in the proposed methodology. Moments are used to evaluate the rectangularity of the extracted shape and can further assist to differentiate between buildings and non-buildings features. The research successfully attempts to extend the work of Rosin (2003), Reiss (1991) and Schweitzer and Straach (1998) that evaluate properties of specific shapes based on moment invariants (Appendix D). We present a methodology to describe and differentiate between buildings shapes using image moments. Indices based on moments were developed and can be used at the end of any feature extraction process to identify a shape with a calculated probability.

5.2.6   Conclusions

The contribution of this research project can be evaluated through the success and failure of the procedure. The successful extraction of the commercial buildings extraction provides a methodology that can be easily replicated. The failure of the procedure for the residential and high-rise buildings provides a contribution to the understanding of the complexity of the problem. When a methodology fails, the circumstances and mitigating factors have to be carefully examined. This research makes an effort to identify the factor that prevent the successful extraction of the buildings and suggests recommendations when possible. Clearly, an automatic process can not always

imitate and completely replace the human logic; hence, this research recognizes the situations that require a manual effort. Since manual digitizing is suggested for multiple image scenes and structure types, a method to reduce the manual effort is tested and evaluated.

The research provides an important analysis of the three very distinct types of buildings. The environmental and architectural aspects of each building type is carefully examined and evaluated as relates to the building extraction effort. The factors that may prevent a successful extraction are identified for each building type and mitigation measures are recommended where possible. Additional extraction procedures including semi-automatic procedures and manual extraction are presented as well. As the building extraction research community concludes two decades of research, a discussion should be initiated regarding the possibility of achieving an automated process. Research can not focus on a specific type of building or specific geometry; there should be a broad perspective that includes careful examination of common building types such as commercial, high-rise and residential. The solution can be an integration of selective methodologies, each aimed at specific building type. This research provides an initial examination of the unique complexity of the problem for each building type and possible solutions and recommendations.

5.3   Limitations of the Approach

The limitations of the approach can be divided into limitation rooted in the imagery source, limitation of the overall approach, and specific limitations within the methodology.

The project concentrates on one of the most readily available and commonly used imagery source. Orthophotos are used in every organization from government to the private industry for many applications, including feature extraction. Using an accepted image source is a clear advantage. However, orthophotos pose obvious limitations to any extraction process. First, any mono image provides only two dimension data and lacks height information. Shadows extraction and analysis is one approach to overcome that limitation. Current research in shadow analysis can not completely replace stereo information. As suggested for future research, LIDAR data may also be integrated into the procedure to provide height information. Second, the common orthophoto generation process does not correct relief displacement artifacts for structures above the datum. This phenomenon is more severe and evident in high-rise buildings and requires special attention prior to extraction.

The design of the methodology formulates a series of analyses that rely heavily on the initial partition of the image as well as the image segmentation results. The initial partitioning of the image is achieved using a parcel layer. The time gap between acquiring the image and the compilation of the parcel layer should be minimal to avoid discrepancies. The image and the parcel layers for this project were created in the same year. Some discrepancies (about 15% of the parcels) still occur due to data errors. Inconsistencies between the two sources create inevitable loss of information. The extraction process can not identify sections of a building that lie outside the parcel boundary. A recommendation to mitigate this limitation is to create a buffer around the parcel prior to partitioning the image in order to capture as much of the building area as

possible. Special attention should be given to high-rise buildings due to relief displacement artifacts.

The segmentation of the image follows the image partition and is the first step within the image analysis. As presented in section 4.2, the final results of the building extraction rely heavily on the segmentation, for all building types. Complex roof signature that provides a scattered segmentation result was shown to prevent successful extraction. This factor was found to be extremely significant for residential and high-rise buildings due to the nature of the environment and roof construction. Since each type of building has unique characteristics, a segmentation methodology should be tailored to each type individually. The methodology can be automatic or semi-automatic to provide the least amount of post processing.

As indicated by the testing, the methodology does not provide successful extraction results for residential and high-rise buildings. Simple scenarios that include minor shadows inclusion, insignificant trees around the buildings and simple roof construction allow reasonable building extraction. On the contrary, when the roof can not be clearly seen from above, an automatic process would fail to extract the structure outline. In many cases, a human would be able to manually extract the footprint of the building, but severe tree and shadow presence would prevent even manual extraction. Recommendations for mitigating measures prior to extraction were provided.

The testing concludes that the automatic process fails to extract high-rise buildings. The process manages to identify and extract simple buildings but fails to extract the common high-rise structures. The characteristics of the high-rise buildings and the high-rise parcels, as indicated in section 4.2.4.1, prevent the successful footprint extraction

using any automatic process. The complexity of the signature on the roof and the shadow casts prevent an adequate segmentation, hence fails to extract the building.

The methodology encompasses a set of analysis techniques and tools. Those techniques entail certain limitations. Some limitations may be overcome by an alternate design, while some are the result of the extraction procedure artifacts and may be very complicated to resolve. As already mentioned above, the process heavily relies on the segmentation result. A segment may define an overshoot or undershoot of the actual building, may include islands and gaps and even exclude an entire section. Compound buildings may be extracted as multiple sections and merged during the post processing. All the artifacts greatly affect the analysis procedure following the segmentation. The size elimination essentially compares the size of the segment to an expected structure size in the database. In case of a severe overshoot or undershoot, we expect a significant discrepancy between the values. Moreover, the geometry analysis of a segment depends on the actual shape of the segment. A segment that deviates from the actual shape would be erroneously analyzed. As a result, a building segment may be categorized as a non-building with high probability and vice versa. The methodology attempts to mitigate the affect of the segmentation by defining a threshold for the size comparison, and multiple geometric measures that define the probability of a segment. The multiple geometric measures were designed to identify all buildings and not confine the process to specific shapes. However, there seem to be some bias towards rectangular buildings and non-rectangular may receive lower probability of representing a building. As a result, the final building vector layer includes segments with a probability greater then 50 (more then 50 percent probability of representing a building). An improvement to this approach

226

should include a library of known shapes with pre-defined shape indices (as defined got the "O" and "I" shapes using moments).

The generalization procedure result relies on the convex hull generation. Artifacts of the automatic process, may add or remove points to the convex hull location, hence, may have a direct impact on the final result. For example, an unexpected spike on one side of the building may significantly change the geometry of the convex hull. Moreover, the angular and distance generalization criteria may eliminate or keep parts of the buildings and should be re-considered on a project-purpose base (the user should decide whether to maintain the essence of the building; exactly the outline; eliminate small intrusion etc).

Shadow analysis can assists to identify three dimensional structures. This project takes advantage of shadow information in a limited manner that includes verifying the legitimacy of a segment as a building, based on the location of the shadow. This methodology may fail when there is a gap between the building roof and the shadow (for example when there is a belt around the building), when objects such as trees around the building cast shadows on the wrong side of the building, when shadows of several buildings interleave or when the color of the building roof resemble shadows. Shadow analysis is important and should be broaden to identify possible building locations, not only to distinguish between 3D and 2D features.

Other limitations can be attributed to the histogram analysis. The histogram analysis is based on identifying the peaks within the histogram and their respective ranges. As emphasized in the testing results, the nature of the histogram graph may have a considerable affect on the histogram analysis result. It is fairly easy to identify the peaks in "smooth" histogram graphs (for the commercial buildings), while "jagged" graphs with

abrupt changes (for the residential) clearly pose a greater challenge on the analysis. Moreover, the histogram analysis requires the number of peaks to identify. The balance between loosing information and extracting too much un-relevant (noise) information should be evaluated. More peaks usually mean more segments to analyze.

The following section discusses future work as relates to the limitation of the research and the testing results.

5.4 Future Research

Future research work should be derived from the unique nature of each type of building. Histogram analysis may be sufficient for commercial buildings extraction but not sufficient for residential structures. Residential roofs, as mentioned before, are designed to appear attractive. As a result, the material and the set up create a rough surface. Moreover, shadow casts may generate visually different sections. A potential alternative to the histogram analysis for residential parcels is based on the fact that the colors may appear different, but the texture of the roof remains the same. Texture, although easy for a human to comprehend, is still a fairly complex task for a computer. Haralick et al. (1973) introduced the "Gray-Tone Spatial-Dependence Matrices" to extract texture of surfaces from images, which are referred to as GLCM - Grey Level Co-occurrence Matrix. The GLCM examines combination of pixels in the image based on adjacency and the frequency of their occurrence. GLCM is computation intensive procedure that is not easy to implement but the methodology is used to examine texture of objects (Soh and Tsatsoulis, 1999; Kuplich et al, 2003) and texture classification (Jing Yi, et al., 2008).

Features can be described using different characteristics such as geometry, spectral signatures and texture. Regardless of the segmentation methodology, geometry analysis can definitely assist during the analysis and shape recognition phases. Improved geometric analysis would consider developing shape indices for a library of shapes. The testing for the "O" and "I" shapes proved a plausible shape differentiation and shape recognition using the moment theorem. The library can be used to enhance the probability calculation of a segment within the extraction process. Besides rectangularity, a segment would be examined for other pre-defined shapes. The highest value would be used to evaluate whether the shape represents a legitimate building shape.

Shadows may provide significant information about the location and shape of the building. The use of shadows in this project is somewhat limited and can be broaden. Shadows may assist to identify the actual location of the building. Moreover, geometry analysis of the shadow cast should enable to differentiate between actual building shadows that have straight lines adjacent to the building, and objects like trees, that have less specific characteristics.

Shadows help to capture the three dimensional aspect of the buildings. An optional data resource that provides three dimensional data is a LIDAR cloud. LIDAR image integrated with an orthophoto can dramatically improve the extraction result, especially for high-rise buildings. Those structures are large enough to maintain a sufficient number of LIDAR points. Moreover, the affect of relief displacement would not continue to be a factor that prevents the extraction. LIDAR may significantly improve residential area analysis by eliminating ground level features as well as identifying tree canopies. Due to the nature of the LIDAR point cloud, building sides would appear as an abrupt change,

while tree sides appear as a gradual change. Multiple return LIDAR points may be used to extract structure hidden under trees.

The project initiated an investigation of factors that may prevent automatic extraction of buildings. Future work should carefully investigate those factors as well as expand the examination to mitigating factors. For example, "is there a preferred time of day for acquiring the image?". Shadows information should be exhausted during the extraction process, but may potentially compromise the procedure. It may be better to acquire the image with limited shadow affects to achieve the best extraction result.

A successful methodology would apply a specific set of analysis tools for a specific type of building. One methodology approach can not resolve the extraction problem and the cost-benefit of an automatic extraction should be evaluated on a project basis. However, future work should integrate multiple methodologies and procedures and allow the user to initiate the proper methodology in his/her discretion. Any such system should employ the parcel based image partitioning with a buffer. The user would be able to evaluate whether a fully automated, semi automated or manual extraction should be used. Testing would compare the system to a fully manual extraction. As the need for building extraction grows within GIS departments in the public and private sector, a solution that alleviates the tasks and allows efficient extraction, is required. The described system may be the anticipated solution for a more efficient, mass extraction solution.

An important aspect of the project is the integration of image sources with GIS data. GIS data was proved to assist and significantly alleviate the extraction procedure as well as provide important segment analysis information (the building area elimination phase). Future research should exhaust the use of available GIS information such as parcel data,
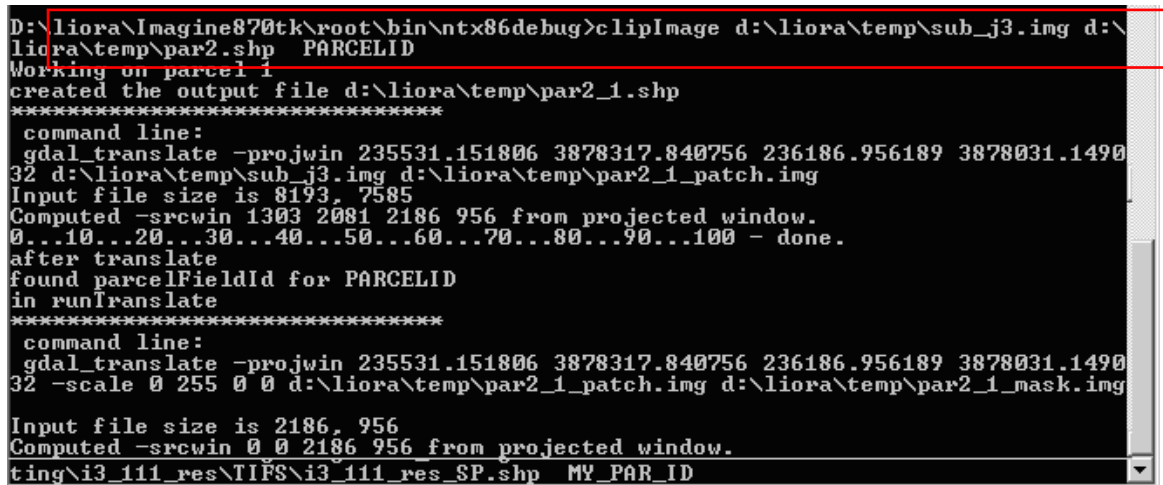
building blue prints and road layers to simplify the analysis, provide clues for the location of the building and update changes in the structure.

Finally, the entire approach should be tested in other regions. The Memphis test-bed provides a wide variety of building types, sized and colors. Testing in a different environment would allow to substantiate the results of the testing, and generalize the conclusions. Additional testing may also include image source at different resolution to evaluate the validity of the assumptions and methodology for other images (either higher resolution Quickbird image or lower resolution IKONOS image).

APPENDIX A . Image Sub-Setting Procedure

The Image Sub-Setting procedure was implemented as a java program, using GDAL (Geospatial Data Abstraction Library) libraries. GDAL libraries are open source libraries that can be found in http://www.gdal.org/. The program runs as a command line and requires an image and polygon layer as an input. The image format should be "img" (ERDAS-IMAGINE native format) and the polygon layer is expected as a "shp" file (ESRI vector format). Each new image is identified using an attribute in the shp file as defined by the user (See Figure 132)


```
D:\liora\Imagine870tk\root\bin\ntx86debug>clipImage d:\liora\temp\sub_j3.img d:\
liora\temp\par2.shp  PARCELID
Working on parcel 1
created the output file d:\liora\temp\par2_1.shp
*******************************
 command line:
 gdal_translate -projwin 235531.151806 3878317.840756 236186.956189 3878031.1490
32 d:\liora\temp\sub_j3.img d:\liora\temp\par2_1_patch.img
Input file size is 8193, 7585
Computed -srcwin 1303 2081 2186 956 from projected window.
0...10...20...30...40...50...60...70...80...90...100 - done.
after translate
found parcelFieldId for PARCELID
in runTranslate
*******************************
 command line:
 gdal_translate -projwin 235531.151806 3878317.840756 236186.956189 3878031.1490
32 -scale 0 255 0 0 d:\liora\temp\par2_1_patch.img d:\liora\temp\par2_1_mask.img

Input file size is 2186, 956
Computed -srcwin 0 0 2186 956 from projected window.
ting\i3_111_res\TIFS\i3_111_res_SP.shp  MY_PAR_ID
```

Figure 132– Running the Image Sub-Setting Procedure. The inputs: image "sub_j3.img, shape file "par2.shp, attribute name "PARCELID".

APPENDIX B . Graham Algorithm for calculating the Convex Hull

Graham published his algorithm for calculating the convex hull in 1972. The algorithm is commonly referred to as Graham's scan. The algorithm assumes a finite number of points and begins by locating the most extreme point on the y axis (the lowest or the highest y coordinate). This point would always be part of the hull and is the center of the following analysis. Step 2 includes calculating the angle based on the extreme point in an increasing manner. Step 3 includes the creation of the hull by calculating the turn between the points. A left turn represents a point on the hull edge. A right turn requires the scan to go a step back. The complexity of the process is evaluated as O(n log n).

Example of the scan is provided below:
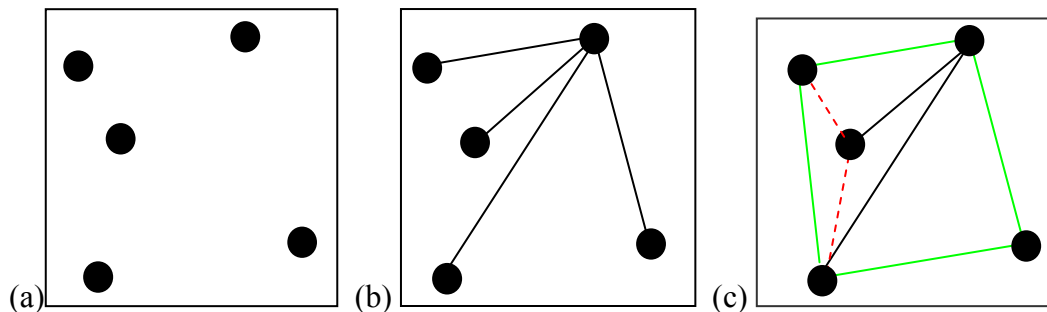
Figure 133 – Graham scan. (a) A finite set of points (b) sorting the points by angle (c)creating the hull (green). Wrong turns (red)

APPENDIX C . Calculating confidence for a segment

For each segment during the geometry analysis (section 4.1.5), a set of geometric characteristics is calculated. The characteristics include rectangularity, convexity, ellipticity and compactness. This section provides the VB script used to calculate the confidence based on the different indices. The probability was calculated for each range based on the number of observations for the building and non-building on the graph (normalized by the total number of features in the class–buildings or non-buildings). For example, the total number of building features is 250/ non-buildings 334. For a point in the graph that has 150 buildings and 50 non-buildings, a normalized number is calculated as 150/250=0.6 and 50/334=0.15 respectively. Hence we can conclude an 80% (0.6/0.75) probability for the range for representing a building feature.

***Lines that begin with an asterisk indicate my comments.

```
Dim conf As Double
Dim rect_m As Double
Dim rect_d As Double
Dim rect As Double
Dim convex As Double
Dim compact As Double
Dim ellip As Double
Dim convex_area As Double
Dim convexty As Double


***rectangularity based on moments
Select Case [rect_momen]
   Case 0.8 To 1
      rect_m=100
   Case 0.5 To 0.8
      rect_m=50
   Case 0.3 To 0.5
      rect_m=20
   Case Else
      rect_m=10
```

```
End Select

***rectangularity based on MBR
Select Case [rect_disc_]
    Case 0.85 To 1
        rect_d=100
    Case 0.7 To 0.85
        rect_d=80
    Case 0.6 To 0.7
        rect_d=40
    Case Else
        rect_d=20
End Select

If rect_m > rect_d Then
    rect= rect_m
Else
    rect= rect_d
End If


***compactness
Select Case [compactnes]
    Case 0.9 To 1
        compact=90
    Case 0.6 To 0.9
        compact=80
    Case 0.5 To 0.6
        compact=70
    Case Else
        compact=30
End Select

*** ellipticity
Select Case [elps_index]
    Case 0.85 To 1
        ellip=80
    Case 0.7 To 0.85
        ellip=30
    Case Else
        ellip=10
End Select

***convexity based on generalized polygon
Select Case [convexity]
    Case 0.95 To 1
```

```
        convex=100
    Case 0.85 To 0.95
        convex=40
    Case 0.7 To 0.85
        convex=20
    Case Else
        convex=10
End Select

***convexity based on area ratio
Select Case [convex_are]
    Case 0.5 To 1.35
        convex_area=80
    Case 1.35 To 1.5
        convex_area=30
    Case Else
        convex_area=0
End Select

If convex > convex_are Then
    convexty= convex
Else
    convexty = convex_are
End If

***final average of the probabilities into a confidence
    conf=(rect+convexty+ellip+compact)/4
```

APPENDIX D . Shape Recognition Using Moments

The human brain can fairly easily identify and classify the shape of a building by simply observing the feature. An automatic process, that accurately identifies and classifies the shape of buildings, may result in valuable time saving as well as accurate, consistent classification. Automatic classification measures intend to imitate human pattern recognition and decision making mechanism, which in many cases become "fuzzy". Even when testing "clean", lab created features, the difference between a rectangle and an "O" shape may not be clear. How big should be the island inside the "O" in order to be clearly classified as an "O" and not as a rectangle with a small island? Building footprints, as extracted from an image, may vary greatly in their shape depending on the method of extraction. Manual digitization and for greater extent, automatic processes, introduce noise to the edge of the extracted feature. As a result, the "fuzziness" of the extraction becomes a greater obstacle for the automatic classification to overcome.

Based on shape identification techniques that use and expand the moment theorem (Hu, 1962), this chapter evaluates the ability to identify not only simple feature such as triangles and rectangles (Rosin, 1999; Rosin, 2003), but two other common building shapes. The chapter outlines the steps taken to develop the index for "O" and "I" shapes. Based on the index formulas, different "clean" and "extracted" features are tested to evaluate the validity of each index and identify the "fuzzy" classification area, where the index may not clearly identify the shape.

Moment Invariants

Moment invariants have been commonly used for the purpose of shape recognition (Hu 1962; Dudani et al, 1977; Blumenkrans, 1991; Bookstein, 1991; Jiang and Bunke, 1991; Li, 1992; Safaee-Rad et al., 1992; Trier et al., 1996; Loncaric, 1998; Realpe and Velazquez, 2006) . Different moments have been introduced in the literature (such as the Hu, Legendre, Zernike, pseudo-Zernike polynomials or Chebychev moments) and allow unique description of a shape. By normalizing central moments, Hu (1962) introduced his set of invariant moments that include seven moments invariant to translation, rotation and scale. Following sets of moments tried to overcome some of the shortcomings of Hu's moments. Since Hu's moments were found to not be orthogonal, they sustain a high degree of redundancy. Modifications to Hu's moments allow minimizing the degree of redundancy (Rothe et al. 1996; Zhang et al. 2003). Moment invariants methods are frequently used and are theoretically sound, although they bear certain disadvantages such as sensitivity to noise within high order moments.

Moment definition

For a continuous function f(x,y), we can define raw and central moments (translation invariants) as follows:

Raw moments are defined as:

$$M_{p,q} = \iint\limits_{x\,y} X^p Y^q f(x, y) dxdy$$

Central Moments are defined as:

$$\mu_{p,q} = \int_x \int_y (X - \bar{X})^p (Y - \bar{Y})^q f(x, y)\,dxdy$$

$(\bar{X}, \bar{Y})$ is the centroid of the shape.

The moments can be modified to the special case of an image as the following scalar values. For images, Raw moments are defined as:

$$M_{p,q} = \sum_x \sum_y X^p Y^q I(x, y)$$

I represents the intensity of the image in the location (x,y)

Central Moments are defined as:

$$\mu_{p,q} = \sum_x \sum_y (X - \bar{X})^p (Y - \bar{Y})^q I(x, y)$$

The centroid (center of mass) of the shape is defined as: $\bar{X} = M_{10}/M_{00} \quad \bar{Y} = M_{01}/M_{00}$

M_{00} is the area of the shape.

Based on the moments, we can calculate the orientation of the shape using:

$$\theta = \frac{1}{2}\tan^{-1}\frac{2\mu_{11}}{2\mu_{20} - 2\mu_{02}}$$

Shape Indices

Rosin (1999) defines a rectangularity index based on image moments as follows:

$R = 144 \dfrac{\mu_{22}}{m_{00}^3}$. R should peak at 1 for rectangles and range between 0-1 for non

rectangles. Value greater then 1 should be calculated as $\dfrac{1}{R}$.

The rectangularity index is based on a rectangle (a x b), centered at the origin (figure 134).
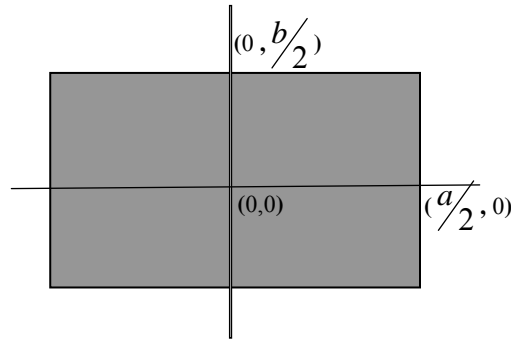
Figure 134 – Rectangle centered at the origin used for Rectangularity index definition

According to Rosin (1999), the moments of that rectangle are:

$m_{00} = ab$ and $m_{22} = \dfrac{a^3 b^3}{144}$ (the proof for m22 is provided in appendix E).

The rectangularity index is a direct result of those two moments.

Based on the rectangularity index, as defined by Rosin (1999) and the mathematical

proof, as provided in E, an O shape index was defined. The O shape similarly is centered

at the origin and is defined as indicated in figure 135. Each side of the O shape is half the

side of the rectangle, and is defined as a/2 x b/2.

The figure shows a shaded rectangle with a white rectangular hole (O shape) centered at the origin, with labeled points $(0, b/2)$, $(0, b/4)$, $(a/4, 0)$, and $(a/2, 0)$.
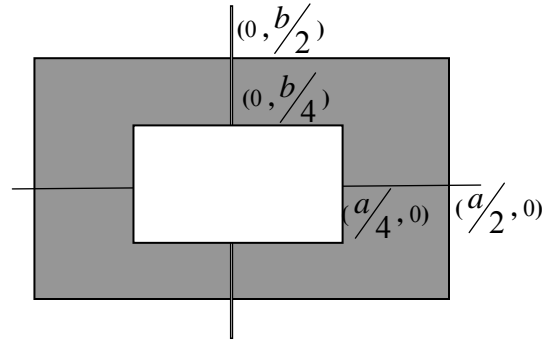
Figure 135 – O shape centered at the origin used for O shape index definition

The O index is defined as: $O = \dfrac{432}{7} \dfrac{\mu_{22}}{m_{00}^3}$. The proof for the O index is provided in

appendix F.

An I shape is defined centered at the origin and defined as indicated in figure 136. The leg width of the shape is defined as 1 forth of the entire width of the shape. The

index was calculated as $\dfrac{16*1125}{449} \dfrac{m_{22}}{m_{00}^3}$. An I shape with leg width of one half of the

entire width was also calculated as $\dfrac{1296}{19} \dfrac{m_{22}}{m_{00}^3}$.

241
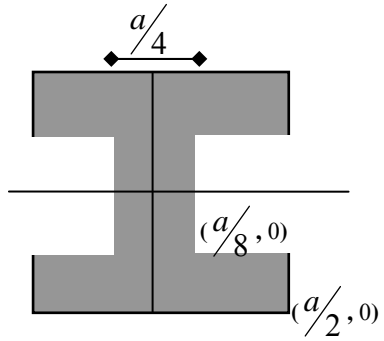
Figure 136 - I shape centered at the origin used for O shape index definition

The second I index (half width leg) was added due to confusion between the O index and the "I" index. To avoid the confusion, a second definition of the "I" shape was added as a half width leg rather then a quarter width leg. Enriching the shape library with a second definition allows to evaluate whether multiple definitions manages to discern better between different shapes.

Shape recognition testing

The table below contains the shape indices as calculated for multiple shapes. The shapes are either "clean", manually created shapes or automatically extracted shapes. Based on the equations provided above, the rectangularity, "O" shape and "I" shape indices are calculated. For the "I" shape, two values are calculated, one for the one quarter leg width and another for the one half leg width.

Table 10 - Shape indices testing results

	Shape	Rectangularity	O Index	I Index	Kurtosis
1		0.452647	**0.935853**	0.613674 **0.958547**	0.989674
2		0.622854	**0.680113**	0.445976 **0.758159**	0.986052
3		**0.816677**	0.518701	0.340132 0.578224	0.995884
4		**0.999979**	0.423602	0.277772 0.472212	1.000001
5		0.369648	**0.872611**	0.751466 0.782783	0.985981
6		0.416654	**0.983578**	0.666686 0.882327	0.979240
7		**0.985832**	0.429699	0.281770 0.479009	1.031014

Table 11 continued
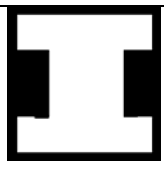
8		0.613165	0.690860	0.453023 0.770139	1.033631
9		0.420427	**0.992483**	0.660704 0.890316	0.950839
10		**0.999619**	0.423450	0.277672 0.472189	0.999923
11		**0.746289**	0.567623	0.372212 0.632761	0.996371
12		0.284244	0.671001	**0.977253** 0.601928	0.621817
13		0.527854	0.802515	0.526240 **0.894607**	0.722123

Table 12 continued

14		**0.681899**	0.621223	**0.407359** **0.692511**	0.832732
15		0.247866	0.585127	**0.892318** 0.524893	0.600655
16		0.619456	0.683844	0.448422 **0.762318**	0.786115
17		0.322194	0.760589	**0.862144** 0.682293	0.618506
18		0.077464	0.182867	0.278872 0.164042	0.978075
19		0.712442	0.594590	0.389895 0.662822	1.120515

Table 9 illustrates the calculated indices for rectangular shapes, "I" shapes, "O" shapes and random blobs. The initial testing included the indices for each shape including the rectangularity, "O" index and "I" index columns. The test shapes show progressive transition from an ideal shape to rectangular shape. For example, rows 1-4 show the transition from an "O" shape with a large island to a complete rectangular shape (4). The indices calculated for the shape indicate the difference between the shapes. Since the index was designed to peak at 1 for the "perfect shape" (as defined in the previous section), the shape 1 and shape 9 show a high "O" index value (0.935853/ 0.992483). Shape number 2 present two separate phenomena as the rectangularity index value (0.622854) is close to the "O" index (0.680113) and one of the "I" index values (0.758159) is actually higher then the "O" index. As the island within the "O" shape get smaller, the shape resembles more a rectangle. This transition is clear for shapes 3 and 4 that show the highest index value for "rectangularity". The confusion between the "O" shape and "I" shape is discussed below. Since the result of image segmentation is rarely a clean shape, each shape was tested for a "noisy shape" scenario. Shapes 5-8 present two lab created shapes (5,6) and two segmentation results (7,8). Clearly, noise added to the shape for shapes 5 and 6, do not significantly affect the index result. Both, shape 5 and shape 6, show the highest values for the "O" index. Shape 7 indicate a high rectangularity value, similar to shape 2 and shape 8 shows confusion between the "I" (0.770139) and "O" (0.690860) indices.

Similar to the transition from the "O" shape to rectangle, transition shapes were defined for the "I" shape. Shapes 12-14 illustrates the change in indices as shape number 12, "the ideal shape" shows a very high "I" index (0.9772563); shape 13 shows a high "I"

246

index (0.894607) but also a high "O" index; shape 14 which is very close to a rectangle shows a high "I" index (0.692511) as well as high rectangularity index (0.681899). Shape 14 may be considered close enough to a rectangle by manual classification which alleviates the confusion. Shapes 15-17 represent noisy "I" with indices values that do not deviate considerably from the clean shapes. As a result, the noise does not reflect change in the classification result (0.892318 vs 0.977253; 0.762318 vs 0.692511; 0.862144 vs 0.977253). Shapes 18 and 19 provide examples of automatically extracted blob shapes and their calculated indices. Shape 18 receives very low scores for all indices and shape 19 receives a higher value for the rectangularity index (0.712442).

It is obvious that moment based indices can be used to identify shapes and distinguish between shapes. The index provides a classification tool that compares a shape to a pre-defined shape and calculates a probability. Further investigation is required to examine the confusion between the "O" and "I" shapes as indicated in shapes 1 and 2 for the "O" shape and shapes 12 and 13 for the "I" shape.

The indices, as designed and defined above, are based on second order moments. Second order moments provide a fundamental description of the shape by defining the orientation and elongation of an ellipse that best fit the shape. As a result, obvious confusion occurs between the shapes. In order to distinguish better between the shapes, higher order moments were examined. The third order moments define the symmetry of the shape for each axis. Since both shapes are highly symmetrical relative to the shape center axes, the third order moment does not appear to provide a solution. The forth order moments define the kurtosis of the shape. Kurtosis is a common statistical term that defines the peakedness of a curve. As relates to images, kurtosis defines the intensity

distribution along an axis (Mukundan and.Ramakrishnan, 1998) and is defined as

$K_x = \dfrac{\mu_{40}}{\mu_{20}^2}$ for the x axis and $K_y = \dfrac{\mu_{04}}{\mu_{02}^2}$ for the y axis. Due to the nature of the shapes,

there is a clear difference between the distributions along the axes between the "I" and

"O" shape. The "O" shape has a similar distribution along both axes, while the "I" shape

show different trends along each axis. As a result, a kurtosis index was defined as $\dfrac{K_x}{K_y}$.

The index is designed to peak at 1 for shapes with similar x and y axis distribution such

as "O" and a rectangle. As seen in table 9, the kurtosis values for all the "O" shapes is

very close to 1 and varies between 0.950839 (shape 9) and 1.033631 (shape 8). The "I"

shape kurtosis calculation values are much lower and range between 0.621817 (shape 12)

and 0.722123 (shape 13). It is noted that shape 14 that resemble greatly a rectangle was

not used for the comparison between the "I" and "O" shapes. As a result, the kurtosis

index provides a method to distinguish between the "I" and "O" shapes and to avoid the

confusion as described above.

The above results present a methodology to describe shapes and differentiate between

shapes using image moments. The indices can be used at the end of a feature extraction

process to identify a shape with a calculated probability or to better discern between

buildings and non-buildings blob segments.

APPENDIX E Mathematical proof for the second order moment used for Rectangularity Index.

We would like to prove that $m_{22} = \dfrac{a^3 b^3}{144}$ is true for a rectangle a x b, centered at the origin (figure 137).

Based on $M_{p,q} = \sum_x \sum_y X^p Y^q I(x,y)$:

$$m_{22} = \int_0^{a/2} \int_0^{b/2} X^2 Y^2 = 0(0 + 1^2 + ... + (\tfrac{a}{2})^2) + 1(0 + 1^2 + ... + (\tfrac{a}{2})^2) + ... + (\tfrac{b}{2})^2 (0 + 1^2 + ... + (\tfrac{a}{2})^2) =$$

$$(0 + 1^2 + ... + (\tfrac{a}{2})^2)(0 + 1^2 + ... + (\tfrac{b}{2})^2)$$

Since $\int_a^b X^n dx = \dfrac{b^{n+1} - a^{n+1}}{n+1}$ for n≠-1

$$(0 + 1^2 + ... + (\tfrac{a}{2})^2) = \int_0^{a/2} X^2 = \dfrac{a^3}{24}$$

$$(0 + 1^2 + ... + (\tfrac{b}{2})^2) = \int_0^{b/2} Y^2 = \dfrac{b^3}{24}$$

$$m_{22} = \dfrac{a^3 b^3}{24 * 24} * 4 = \dfrac{a^3 b^3}{144}$$

He result is multiplied by 4 to indicate that the initial integral is calculated for one quarter of the rectangle!

APPENDIX F . Mathematical proof for the O index

We would like to prove that O shape index = $\dfrac{432}{7}\dfrac{m_{22}}{m_{00}^3}$ for an O shape, centered at

the origin (figure 135). As shown in appendix E, $m_{22} = \dfrac{a^3 b^3}{24 * 24}$ for a quarter of a

rectangle (axb).

For each quarter of the internal rectangle (the island of the O shape), we need to

calculate the second order moment. Since the dimension of the island are $(\dfrac{a}{4}, \dfrac{b}{4})$,

similar to the proof in appendix E, $\displaystyle\int_0^{a/4} X^2 = \dfrac{a^3}{4^3 * 3}$ and $\displaystyle\int_0^{b/4} Y^2 = \dfrac{b^3}{4^3 * 3}$ are the result of

the integral on one quarter of the island. As a result, for each quarter of the O shape, we

have to subtract the moment of the island from the moment of the island:

$$m_{22} = \dfrac{a^3 b^3}{24 * 24} - \dfrac{a^3 b^3}{9 * 64 * 64} = \dfrac{7 a^3 b^3}{4096}$$

For the entire rectangle, we multiply by 4, hence:

$$m_{22} = \dfrac{7 a^3 b^3}{1024}. \text{ We know that } m_{00} = \dfrac{3}{4} ab.$$

From m_{22} and m_{00}, for the index to peak at 1 for the O shape, the index is: $\dfrac{432}{7}\dfrac{m_{22}}{m_{00}^3}$

APPENDIX G . Mathematical proof for the I index

We would like to prove that I shape index $= \dfrac{16*1125}{449}\dfrac{m_{22}}{m_{00}^3}$ for an I shape, centered at

the origin (figure 136). As shown in appendix E, $m_{22} = \dfrac{a^3 b^3}{24*24}$ for a quarter of a

rectangle (axb). For the I shape, the two side intrusions have to be subtracted from the

entire rectangle. Hence, for each quarter, a second order moment for the side intrusion

rectangle has to be calculated. Since the dimension of the rectangle are $({}^{3a}\!/_8, {}^{b}\!/_4)$, it

spans from ${}^{a}\!/_8$ to ${}^{a}\!/_2$ on the x axis and from 0 to ${}^{b}\!/_4$ on the y axis. similar to the proof

in appendix E, $\displaystyle\int_{{}^{a}\!/_8}^{a/2} X^2 = \dfrac{63a^3}{512*3}$ and $\displaystyle\int_{0}^{b/4} Y^2 = \dfrac{b^3}{4^3*3}$ are the result of the integral on one

quarter of the island. As a result, for each quarter of the O shape, we have to subtract the

moment of the island from the moment of the island:

$$m_{22} = \frac{a^3 b^3}{24*24} - \frac{63a^3 b^3}{9*64*512} = \frac{449a^3 b^3}{9*64*512}$$

For the entire rectangle, we multiply by 4, hence:

$$m_{22} = \frac{449a^3 b^3}{4608*16}.$$ We know that $m_{00} = {}^{5}\!/_8\, ab$.

From m_{22} and m_{00}, for the index to peak at 1 for the O shape, the index is:

$$\frac{16*1125}{449} \frac{m_{22}}{m_{00}^3}.$$ For an I shape with leg width of one half of the width rather then one

forth, the index is: $\dfrac{1296}{19} \dfrac{m_{22}}{m_{00}^3}$

APPENDIX H . Database attribute scheme for the parcels and building inventory

Table 13 - Attribute Scheme for the parcel dataset

Field Name	Description
PARCELID	Parcel Identifier (Unique; No Duplicates allowed)
COMDAT_REC	TRUE for Commercial Parcels
DWLDAT_REC	TRUE for Residential Parcels
LAT	Latitude of Parcel Centroid in NAD 1983
LONG	Longitude of Parcel Centroid in NAD 1984
PARID	Alternate Parcel Identifier
APRLAND	Appraised Value of Land
APRBLDG	Appraised Value of all buildings in the parcel
RTOTAPR	Total Appraised Value (of land and buildings)
ASMT_CLS	Assessment Class (roughly corresponds to General Occupancy)
ASMT_LUD	Assessed Land Use of parcel (roughly corresponds to Specific Occupancy)
OCC_CLASS	Occupancy Class of parcel (roughly corresponds to HAZUS Occupancy Categories)

Table 14 - Attribute scheme for the building inventory

Field Name	Description
PAR_ID	Parcel Identifier (Duplicates allowed for multiple buildings in the same parcel)
PARID_CARD	Building Identifier (Unique, No Duplicates allowed for this field)
LAT	Latitude of Parcel Centroid in NAD 1983
LONG	Longitude of Parcel Centroid in NAD 1984
STR_TYPE	General Structure Type (used for summarized tabulations in the accompanying workbook)
YEAR_BLT	Year of Construction
OCC_TYPE	HAZUS Occupancy Category
OCC_DETAIL	Specific Occupancy Category, describing the detailed use of the building
TOT_APPR	Total Value of the Building and Contents
BLDG_VAL	Total Value of the Building (obtained by distributing the appraised value of the parcel across all the buildings contained within the parcel, weighted by each building's square footage)
CONT_VAL	Total Value of the Contents within the Buildings (derived from multiplier functions of appraised building value; multipliers vary by occupancy category of building)
NO_DU	Number of Dwelling Units in the building (usually for residential categories; represent multiple use if found in other occupancy categories)
SQ_FEET	Square Footage of the building
EF	Essential Facility Designation for the building
NO_STORIES	Number of Stories for the building
tract	2000 Census Tract Identifier in which the building is located
CT_LAT	Longitude of Census Tract Centroid in NAD 1983
CT_LON	Latitude of Census Tract Centroid in NAD 1984
STR_TYP2	HAZUS Structure Type
pidc	Alternate Building Identifier (not unique, Duplicates allowed)
OCC_BROAD	General Occupancy Categories (used for summarized tabulations in the accompanying workbook)

APPENDIX I . Preliminary Results - Edge detection Approach

In much of the existing literature, edge detection and line segmentation techniques have been suggested heavily for feature extraction. Using the primary data available in this project, I have tried to extract buildings from high resolution imagery via edge detection techniques. This approach includes large scale ortho-rectified aerial images of 1 ft resolution. Locating structures in an entire scene that spans over several miles and contains hundreds of buildings at different size, shape and spectral signature is not an easy task. In order to simplify the problem, the first step in this methodology was to cut the images into smaller areas of analysis. To accomplish this, subsets of the digital image are created using parcel boundaries, in order to reduce the amount of data to process, and to limit the noise in data. Parcel boundary layers are easily available, cover most built areas within the USA and restrict the search for the buildings to a very small area that will be referred to here forth as the "patch". Moreover, each parcel has a usage attribute attached, that classifies the use of the building on that parcel (industrial, residential, vacant etc). This information was first used to eliminate all residential parcels from the analysis. The process is semiautomatic and consists of several major stages as depicted in Figure 137.
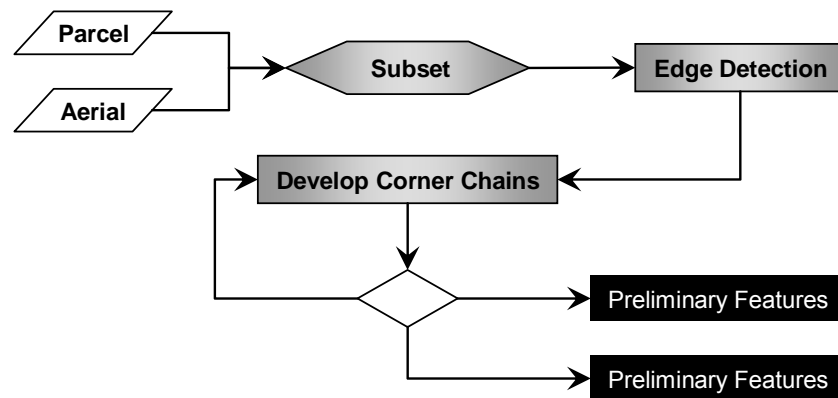
Figure 137 – Building extraction process

The analysis begins with edge-detection – a technique that locates an edge by examining the image for abrupt changes in pixel values (change in contrast). The direction of the edge is perpendicular to the direction of the change (Haralick and Shapiro, 1992). The common edge detection operators can be divided into "directional" and "non-directional". "Non directional" operators locate edges everywhere in the image while "directional" operators require pre-defined edge orientation. One prominent directional operator is Sobel (1970) that defines two masks, one horizontal and one vertical. The edge detection procedure is accomplished via applying a filter-mask or several masks to the image. This process, in the image space, is called "convolution" and it assigns a value to the pixel in the middle of the filter. The filter determines the value based on neighboring pixels. Filters can be low-pass filters that smooth the image and high pass filters that emphasize certain elements in the image, such as edges. The LoG (Laplacian of Gaussian) operator revolutionized the edge detection procedure. During the convolution, smoothing operator is applied to image, followed by a Gaussian operator that emphasized the edges. One drawback of that operator is the possible shift of the

255

edge location when the edge is not a straight line, as in corners. Canny (1986) operator is one prominent edge detector based on smoothing before the edge detection. The Optimal Zero Crossing Operator (OZCO, Sarkar and Boyer, 1991), for which public domain code is available, was used for the edge detection and segmentation.

The building outline usually appears as an edge. Edge points are then grouped into straight lines using a "line segmentation" routine. These steps are performed on the first principal component image created from the original 3 bands. Corners are created by enforcing rules on the line segments derived in the segmentation process. Rules are both proximity- and orthogonality-based with built-in and adjustable tolerances, leading to the creation of L- and T-shaped corners (see Figure 138). Pairs of corners satisfying conditions of distance/orientation and displacement then form short chains. Short chains are grouped into longer chains based on "*shared*" corners. This grouping occurs *iteratively*, successively building longer chains, until the chain is *closed*, or until a *maximum specified number of corners in a building have been reached*. The different chains are grouped according to their corner locations in order to extract multiple buildings in one image. Each corner chain is then **graded** based on several conditions, including distance between corner points covered by corner legs; displacement offsets between lines; orthogonality of the corners; number of corners extracted; type of corners. The chain with the highest score is selected, and the coordinate locations of its corners form the output – these corners are then "built" into a preliminary polygon for downstream analysis. If no chain is found with a high enough score, the process is repeated after tolerance thresholds are incrementally relaxed.

90±σ

L shaped corner

d
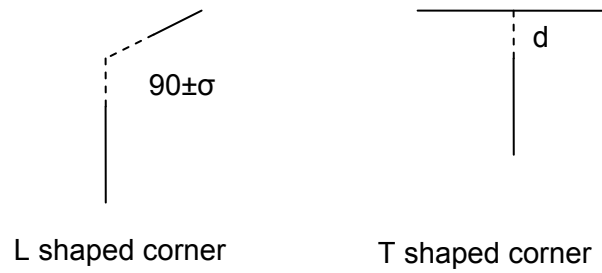
T shaped corner

Figure 138 – Types of corners

Figure 139 below indicates successful implementations of the process for rectangular and L-shaped buildings. Each figure consists of 4 panels from left to right, beginning with the raw image subsetted to the parcel, to edge detection, then to line segmentation and formation of chains and through to the final vectorized building footprint polygon.

(a)

(b)
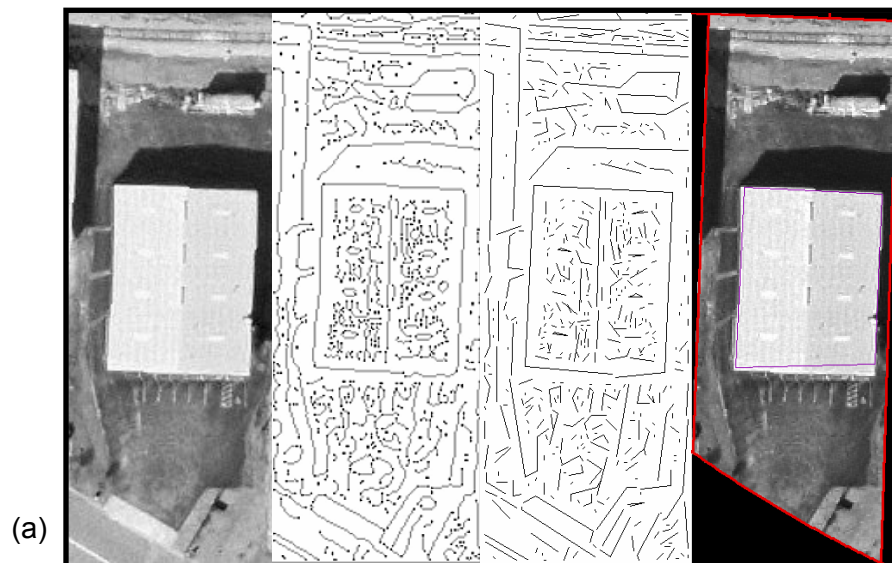
Figure 139 – Successful implementation for a rectangular building (a) and L building (b)

(a)   (b)   (c)
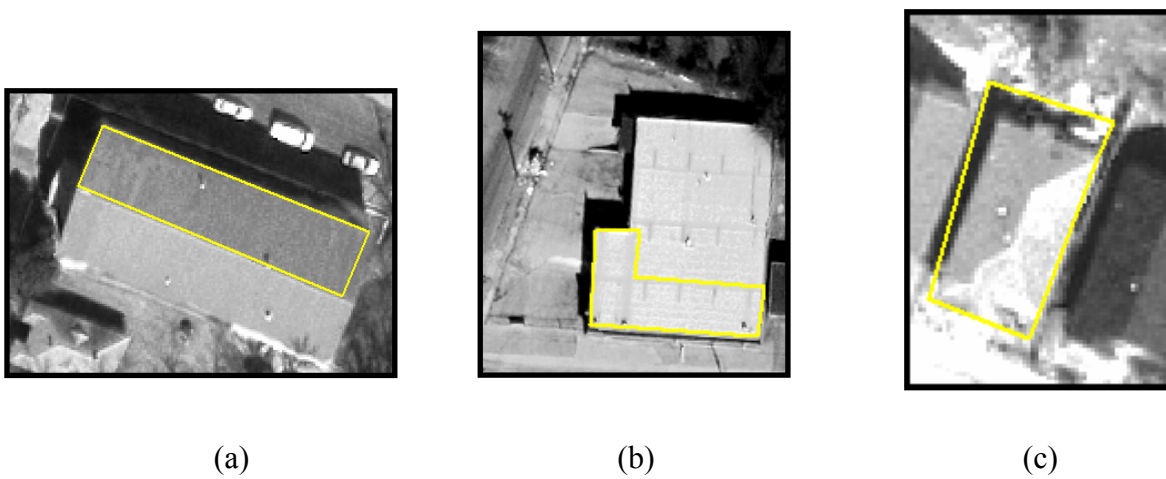
Figure 140 –Examples of result problems with the extraction process

The examples in Figure 140 demonstrate several of the limitations of the implementation that relies heavily on edge detection techniques and geometric characteristics of a building. Even within a parcel sized image, objects around the building, on the roof and shadows may interfere with the extraction process. The different features have strong, straight edges, are within close proximity to the building and become part of the candidate building polygon. Figure 140(a) demonstrates the problem of a non-flat roof. Each side of the roof creates strong enough edges due to sun illumination angle, and be considered as a valid rectangular roof. In figure 140(b) lines on the roof interfere with the external outline edges and create an L shape roof. Figure 140(c) illustrates extraneous objects and shadows around the building that obstruct with the detection of the true building outline.

The initial testing did not produce satisfying results as a stand alone approach. One of the conclusions, consistent with previous research, is the need for integration of approaches. Only geometry or only shadow extraction for example, will not provide a sufficient solution. As a result, we propose a different, more robust methodology as will be illustrated below.

APPENDIX J . Code of ordinance of Memphis, TN – Zoning section

Taken from the "code of ordinances of Memphis, TN"

Available in: http://municipalcodes.lexisnexis.com/codes/memphis/

Footnotes for are found at the end of the chart.

Table 15- Memphis Zoning Ordinanace

District and Use	Minimum Lot Requirements Area (square feet)	Minimum Lot Requirements Width (feet)	Minimum Yard Requirements [7, 20, 23] Front (feet)	Minimum Yard Requirements [7, 20, 23] Side (feet)	Minimum Yard Requirements [7, 20, 23] Rear (feet)	Maximum Height [16] (feet)	Maximum Gross Density Per Acre (in units)	Land Use Intensity Ratios Maximum FAR	Land Use Intensity Ratios Minimum LSR	Land Use Intensity Ratios Minimum OSR
AG district:										
1. Single-family detached dwellings	43,560 (1 acre)	60	60	5	30	35	1.0	None	None	None
2. Institutional uses	1 acre	60	60	FN [1]	FN [1]	35	None	None	None	None
3. Country clubs	1 acre	60	60	FN [1]	FN [1]	35	None	None	None	None
4. Other uses	1 acre	60	60	FN [1]	FN [1]	35	None	None	None	None
R-E district: Dwellings and other uses	22,000	120	50 [22]	15	40	35	0.51	None	None	None

Table 13 continued

R-S15 district: Dwellings and other uses	15,000	100	30 [2]	5 [3]	25	35	2.9	None	None	None
R-S10 district: Dwellings and other uses	10,000	65	20 [2]	5	20 [2]	35	4.4	None	None	None
R-S8 district: Dwellings and other uses	8,000	60	20 [2]	5	20 [2]	35	5.4	None	None	None
R-S6 district: Dwellings and other uses	6,000	50	20 [2]	5	20 [2]	36	7.3	None	None	None
R-S6I district: Dwellings and other uses	See §§ 16-76-6 (E) and (F).									
R-D district:										
1. Single-family detached dwellings and other uses except	6,000	50	20 [2]	5	20 [2]	35	7.3	None	None	None

Table 13 continued

2. Two-family dwellings	6,000	50	30 [2]	5	25	35	14.6	None	None	None
3. Single-family attached dwellings 12	3,000	25	30 [2]	5 [4]	25	35	14.6	None	None	None
R-TH district:										
1. Single-family detached dwellings	6,000	50	20 [2]	5	20 [2]	35	7.3	None	None	None
2. Townhouse dwellings	2,500	16	30 [2]	5 [4]	25	35	12.0	.28	.48	.74
3. Two-family dwellings	6,000	50	30 [2]	5	25	35	14.6	None	None	None
4. Single-family attached dwellings 12	3,000	25	30 [2]	5 [4]	25	35	14.6	None	None	None
5. Other uses	10,000	100	30 [2]	5	20	35	None	None	None	None
R-ML district:										

Table 13 continued

1. Single-family detached dwellings	6,000	50	20 [2]	5	20 2	35	7.3	None	None	None
2. Townhouse dwellings	2,500	16	30 [2]	5 [4]	25	35	15.0	.28	.48	.78
3. Two-family dwellings	5,500	50	30 [2]	5	25	35	15.8	None	None	None
4. Single-family attached dwellings 12	2,750	25	30 [2]	5 [4]	25	35	15.8	None	None	None
5. Multiple-family dwellings	10,000	100	30 [2]	FN [5]	FN [6]	35	15.0	.35	.45	.73
6. Other uses	10,000	100	30 [2]	5	20	35	None	None	None	None
R-MM district:										
1. Single-family detached dwellings	5,000	50	20 [2]	5	20 [2]	35	8.7	None	None	None
2. Townhouse dwellings	2,500	16	30 [2]	5 [4]	25	35	15.0	.28	.48	.74

Table 13 continued

3. Two-family dwellings	5,500	50	30 [2]	5	25	35	15.8	None	None	None
4. Single-family attached dwellings 12	2,750	25	30 [2]	5 [5]	25	35	15.8	None	None	None
5. Multiple-family dwellings	10,000	100	30 [2]	FN [5]	FN [6]	125	30.0	.75	.40	.69
6. Other uses	10,000	100	30 [2]	5	20	125	None	None	None	None
R-MH district:										
1. Single-family detached dwellings	5,000	50	20 [2]	5	20 [2]	35	8.7	None	None	None
2. Townhouse dwellings	2,400	16	30 [2]	5 [4]	25	35	18.2	.28	.48	.74
3. Two-family dwellings	4,800	50	30 [2]	5	25	35	18.2	None	None	None

Table 13 continued

4. Single-family attached dwellings 12	2,400	25	30 ^{2}	5 ^{4}	25	35	18.2	None	None	None
5. Multiple-family dwellings	10,000	100	30 ^{2}	FN 5	FN 6	None	75.0	1.75	.45	.68
6. Other uses	10,000	100	30 ^{2}	5	20	250	None	None	None	None
R-MO district:										
1. Mobile homes	4,000	40	25 ^{2}	5	15	12	8.0	None	None	None
2. Recreation buildings	4,000		25 ^{2}	10	25	35	None	None	None	None
O-L district:										
1. Office and other uses except	5,000	50	30	5 ^{10}	20 ^{10}	35	None	None	None	None
2. Single-family detached dwellings	5,000	50	30 ^{2}	5	20	35	8.7	None	None	None

Table 13 continued

3. Two-family dwellings	5,500	50	30 [2]	5	25	35	15.8	None	None	None
4. Single-family attached dwellings 12	2,750	25	30 [2]	5 [4]	25	35	15.8	None	None	None
O-G district:										
1. Offices and other uses except	5,000	50	30 [2]	5 [10]	20 [10]	125	None	3.0	None	None
2. Single-family detached dwellings	5,000	50	30 [2]	5	20	35	8.7	None	None	None
3. Single-family attached dwellings 12	2,500	25	30 [2]	5 [4]	25	35	34.8	None	None	None
C-P and C-L districts:										
1. All uses except offices	None	None	30 [14, 15]	10 [8]	15 [24]	35	None	None	None	None
2. Offices	None	None	30 [14, 15]	10 [8]	15 [2]	35	None	1.5	None	None

Table 13 continued

C-H district:										
1. All uses except offices	None	None	30 [14, 15]	10 [8]	15 [2]	50 [13]	None	None	None	None
2. Offices	None	None	30 [14, 15]	10 [8]	15 [2]	50 [13]	None	3.0	None	None
CBD district:										
1. Single-family detached dwellings	5,000	50	None	5	15	35	8.7	None	None	None
2. Townhouse dwellings	1,200	16	None	5 [4]	15	35	36.3	.28	.48	.74
3. Two-family dwellings	2,500	36	None	5	15	35	34.8	None	None	None
4. Single-family attached dwellings 12	1,250	16	None	5 [4]	15	35	34.8	None	None	None
5. Multiple-family dwellings	None	None	None	None	None	Unlimited	None	1.75	.45	.68

Table 13 continued

6. Offices	None	None	None	None	None	Unlimited	None	7.00	None	None
7. Commercial uses	None	None	None	None	None	Unlimited	None	7.00	None	None
8. Other uses	None	None	None	None	None	Unlimited	None	7.00	None	None
I-L district: All uses	None	60	30 [14, 15]	10 [8]	15	FN [9]	None	None	None	None
I-H district: All uses	None	60	30 [14, 15]	10 [8]	15	FN [9]	None	None	None	None
H district:										
1. Single-family detached dwellings	5,000	50	30 [2]	5	20	35	8.7	None	None	None
2. Townhouse dwellings	2,500	16	30 [2]	5 [4]	20	35	17.4	.28	.48	.74
3. Two-family dwellings	2,500	25	30 [2]	5	20	35	34.8	None	None	None

Table 13 continued

4. Single-family attached dwellings 12	1,250	16	None	5^4	15	35	34.8	None	None	None
5. Multiple-family dwellings	10,000	100	30^2	FN^5	FN^6	125	30.0	.75	.40	.69
6. Offices	5,000	50	$30^{14,\,15}$	5^{10}	20^{10}	125	None	2.00	None	None
7. Hospitals	None	50	None	10^{10}	20^{10}	125	None	2.00	None	None
8. Other uses	5,000	50	$30^{14,\,15}$	5^{10}	20^{10}	125	None	None	None	None
CU district:										
1. Single-family detached dwellings	5,000	50	30^2	5	20	35	8.7	None	None	None
2. Townhouse dwellings	2,500	16	30^2	5^4	20	35	17.4	.28	.48	.74
3. Two-family dwellings	2,500	25	30^2	5	20	35	34.8	None	None	None

Table 13 continued

4. Single-family attached dwellings 12	1,250	16	30 [2]	5 [4]	20	35	34.8	None	None	None
5. Multiple-family dwellings	10,000	100	30 [2]	FN [5]	FN [6]	125	30.0	.75	.40	.69
6. Offices	5,000	50	30 [14, 15]	5 [10]	20 [10]	125	None	2.00	None	None
7. Educational institutions	None	50	None	10 [10]	20 [10]	125	None	2.00	None	None
8. Other uses	5,000	50	30 [14, 15]	5 [10]	20 [10]	125	None	None	None	None
P district:										
1. Parking lots	5,000	50	FN [11]	5 [8]	5 [8]	None	None	None	None	None
2. Parking structure	5,000	50	30	10 [10]	15 [10]	50	None	None	None	None
FW district	None	None	50	None	None	None	None	None	None	None
C-N district	None	None	None	None	None	35 [21]	None	None	None	None

270

FOOTNOTES

1 Buildings shall be set back from side and rear lot lines, two feet for each foot of building height.

2 The minimum front yard setback set forth on Chart 2 shall apply except as follows.

(a) If the property has a front yard along a major road or parkway identified in the Major Road Plan a minimum forty-(40) foot front yard setback is required.

(b) If the property has a rear yard in a reverse frontage orientation along a major road or parkway identified in the Major Road Plan a minimum twenty-five (25) foot rear yard is required.

(c) If the property has a front yard along a major or minor collector street as defined by the subdivision regulations a minimum thirty (30) foot front yard setback is required.

(d) All subdivisions recorded prior to the date of approval of this zoning text amendment shall be governed by the required minimum thirty-(30) foot from yard and twenty-five- (25) foot rear yard setbacks unless a modification is approved by the land use control board to allow the reduced setbacks for the entire subdivision. Consideration of such modification shall require signature of all property owners. Appeal of land use control board action is to the appropriate local legislative body(ies).

(e) Approval of the reduced setbacks on individual lots created prior to the date of approval of this zoning text amendment shall be the jurisdiction of the board of adjustment.

(f) Lots along cul-de-sac streets serving twenty-five (25) or fewer dwelling units and where sidewalks are not required shall be allowed a minimum front yard setback of fifteen (15) feet.

3 The minimum side yard allowed shall be five feet; provided, however, that the total number of feet combined of the side yards on each side of a dwelling or other use shall be fifteen (15) feet.

4 The side yard requirements shall apply to only one side yard of the first and last attached houses in each set of attached houses. Each attached single-family dwelling shall have only one five-foot wide side yard.

5 The size of the required side yard shall be the greater of ten (10) feet or the number of feet derived under the following formula for buildings on the perimeter of a multiple dwelling development. The size of the required side yard for buildings located on the interior of a multiple-family dwelling development shall be determined as follows:

(a) If a required side yard abuts a building wall which contains any living room windows, the size of the required yard shall be computed as follows: ten (10) feet plus two feet for every ten (10) feet of wall height and fraction thereof plus one foot for every fifteen (15) feet of wall length or fraction thereof.

(b) If a required side yard abuts a building wall which contains any window other than living room window, the size of the required yard shall be computed as follows: five feet plus one foot for every ten (10) feet of wall height and fraction thereof plus one foot for every fifteen (15) feet of wall length or fraction thereof.

(c) If a required side yard abuts a windowless building wall, the size of the required side yard shall be computed as follows: Five feet plus one foot for every ten (10) feet of a wall height and fraction thereof.

6 The size of the required rear yard shall be the greater of twenty (20) feet or the number of feet derived under the following formula for buildings in the perimeter of the multiple-family dwelling development. The size of required rear yards for buildings located on the interior of a multiple-family dwelling shall be determined as follows:

(a) If a required rear yard abuts a building wall which contains any living room windows, the size of the required side yard shall be computed as follows: ten (10) feet plus two feet for every ten (10) feet of wall height and fraction thereof plus one foot for every fifteen (15) feet of wall length or fraction thereof.

(b) If a required rear yard abuts a building wall which contains any windows, other than living room windows, the size of the required yard shall be computed as follows: five feet plus one foot for every ten (10) feet of wall height and fraction thereof plus one foot for every fifteen (15) feet of wall length or fraction thereof.

(c) If a required rear yard abuts a windowless building wall, the size of the required rear yard shall be computed as follows: five feet plus one foot for every ten (10) feet of wall height and fraction thereof.

7 A corner lot used for single-family purposes shall have two front yards and two side yards; i.e., no rear yard. A corner lot used for nonsingle-family purposes shall have two front yards, one side yard and one rear yard.

8 The minimum side yard requirement shall apply only if the property abuts or is adjacent to property zoned or used for residential purposes or the residential portion of an approved planned development. Otherwise, no side yard is required.

9 The maximum height for buildings located in the I-L and I-H district shall be one hundred (100) feet except where property in the I-L and I-H district abuts or is adjacent to property zoned or used for residential purposes or the residential portion of an approved planned development, in which case the maximum height shall be forty-five (45) feet.

10 The required side and rear yards adjacent to property zoned or used for residential purposes and the residential portion of an approved planned development shall be increased one foot for every foot of building height and fraction thereof above two stories.

11 A minimum front yard of thirty (30) feet shall be required if the front yard of the property abuts or is adjacent to property zoned or used for residential purposes or the residential portion of a planned unit development; otherwise, a minimum front yard of five feet shall be required.

12 Only two attached single-family dwellings are permitted in a building.

13 Provided that the height limitations of property in the C-H zone which was formerly in C-3 commercial zone shall be no more restrictive than were the height limitations of the former C-3 commercial zone under the preexisting ordinance.

14 Where an existing principal structure is set back less than forty-two (42) feet from the rights-of-way line, a new principal structure, or an extension thereof may be constructed at less than the required forty-two (42) foot setback provided:

(1) The new construction is within one hundred (100) feet of an existing principal structure and is located on the same side of the same street and within the same block;

(2) The new construction is not closer to the right-of-way than the existing principal building(s) as specified above;

(3) The minimum setback permitted is fifteen(15) feet from the proposed right-of-way line as indicated on the Memphis urban area transportation plan or forty-five (45) feet from the center line of a street not designated on the MUATS plan; and

(4) Landscaping shall be provided in accordance with Section 16-120-4(D) if parking is provided within the required front yard.

15 Where an existing principal building is set back less than thirty (30) feet from the right-of-way line, a new principal building(s) or an extension thereof may be constructed at less than the required thirty- (30) foot setback provided:

(1) The new construction is within one hundred (100) feet of the existing principal building(s);

(2) The new construction is not closer to the right-of-way than the existing principal building(s) as specified above;

(3) The minimum setback permitted is fifteen (15) feet; and

(4) Landscaping shall be provided in accordance with Chapter 16-12.

16 Church steeples, spires and belfries are permitted to exceed the maximum permitted height limitations.

[17--19 Reserved.]

20 Modification of required building setbacks may be approved by the director of planning and development (or designated representative) up to a maximum of ten (10) percent of the required building setback subject to the following criteria:

(1) A modified building setback shall not be approved unless a specific plan for placement of a building on the site is presented which justifies that the changed setback is needed to accommodate the development.

(2) The modified building setback shall not conflict with streets, sidewalks, easements or landscape requirements.

(3) The modified building setback shall not injure or damage the use, value or enjoyment of surrounding property or hinder or prevent the development of surrounding property.

21 Maximum height - thirty-five (35) feet, except where abutting residential uses are less than thirty-five (35) feet, then height standard needs to be compatible with site plan review standards of Ordinance No. 5026AM.

22 In the R-E district, attached garages facing the street shall not be permitted.

23 Within the city a corner lot that is fifty (50) feet or less in width may be allowed one front yard setback of fifteen (15) feet subject to the following criteria:

(1) The second front yard shall conform with the requirements of the zoning ordinance.

(2) Along a major road, the fifteen (15) foot setback shall be subject to the approval of the land use control board.

(3) The building shall not encroach into utility, drainage or other easements.

(4) The building shall not block the vision triangle.

24 When a commercial use abuts the rear yard of an existing or proposed residential use, the minimum rear yard setback for the commercial structure shall be twenty (20) feet for the first ten thousand (10,000) square feet of building, plus an additional five feet of setback for each additional ten thousand (10,000) square feet or fraction thereof of commercial building.

(Ord. 5026AM § 2 (part), 2004; Ord. 5026 § 2 (part), 2003; Ord. 5025 § 2 (part), 2003; Code 1985 Appx. A, Chart 2)

APPENDIX K . Building Extraction GUI

Figure 141 illustrates the user interface for the building extraction process. The interface was developed using ERDAS-IMAGINE interface tools.
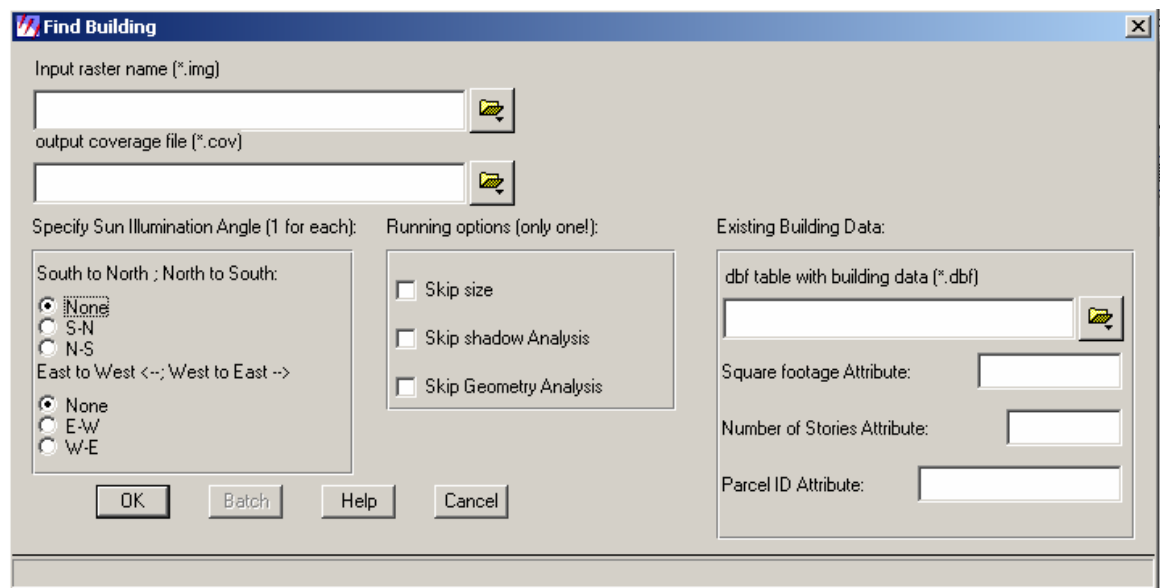
Figure 141 – Building Extraction GUI

The user input includes the input image and output vector file at the top of the window. The middle section of the window includes 3 parts. In the left part the user defines the sun illumination direction as appears in the image. In the middle section, the user defines which elimination analysis to skip for a specific run. The default is a full run that includes the entire set of analysis processes. The right section includes the GIS input. The input includes a dbf table with the building data, including the structure size and the specific fields names (size column, number of stories and the parcel id attribute).

278

The bottom of the window provides the user with the option to run the process in a batch mode. The batch mode runs the building extraction procedure for multiple parcel-sized images.

REFERENCES

Alt, F. L. (1962). "Digital Pattern Recognition by Moments." *J. ACM*, 9(2).

Aviad, Z., Carnine, P. D., Jr., and masters, f. "Road finding for road-network extraction." *Computer Vision and Pattern Recognition, 1988. Proceedings CVPR '88., Computer Society Conference on*, 814-819.

Avrahami , Y., Raizman,Y and Doytsher, Y. . (2008). "A Polygonal Approach for Automation in Extraction of Serial Modular Roofs." *PE&RS*, 74(11).

Baartz et al. (2001). "eCognition User Guide." Munich: Definiens Imaging GmbH.

Ball, G. H., and Hall, D. J. (1967). "ISODATA. an iterative method of multivariate analysis and pattern recognition." *Behavior Science*, 153.

Baltsavias, E. P. (2004). "Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems." *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), 129-151.

Benz, U. (2001). "Definiens Imaging GmbH: Object-Oriented Classification and feature detection." *IEEE Geoscience and Remote Sensing Society Newsletter*, September, 16-20.

Berretti, S., Del Bimbo, A., and Pala, P. (2000). "Retrieval by shape similarity with perceptual distance and effective indexing." *Multimedia, IEEE Transactions on*, 2(4), 225-239.

Bin, F., Wynne, H., and Mong Li, L. (2002). "Tumor cell identification using features rules." Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, Edmonton, Alberta, Canada.

Brenner, C. (2005). "Building reconstruction from images and laser scanning." *International Journal of Applied Earth Observation and Geoinformation*, 6(3-4), 187-198.

Canny, J. F., and masters, f. (1986). "A computational approach to edge detection." *IEEE Transaction on Pattern Analysis and machine Intelligence*, 8(6), 679-698.

Carson, C., Belongie, S., Greenspan, H., and Malik, J. (2002). "Blobworld: image segmentation using expectation-maximization and its application to image querying." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(8), 1026-1038.

Chen, Y., and Wang, J. Z. (2004). "Image Categorization by Learning and Reasoning with Regions." *Journal of Machine Learning Research 5 (2004)* 913-939.

Cover, T., and Hart, P. (1967). "Nearest neighbor pattern classification." *Information Theory, IEEE Transactions on*, 13(1), 21-27.

Davis, C. H., and Xiangyun, W. "Urban land cover classification from high resolution multi-spectral IKONOS imagery." *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International*, 1204-1206 vol.2.

Douglas, D., and Peucker, T. (1973). "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature." *The Canadian Cartographer*, 10(2), 112-122.

Duan, J., Prinet, V. and Lu. H. "Building extraction in urban areas from satellite images using GIS data as prior information." *Proceedings from the 2004 IEEE International, Geoscience and Remote Sensing Symposium, IGARSS* 4762- 4764

Elad, M., Milanfar, P., and Golub, G. H. (2004). "Shape from moments - an estimation theory perspective." *Signal Processing, IEEE Transactions on*, 52(7), 1814-1829.

Foody, G. M. (1996). "Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data." *International Journal of Remote Sensing*, 17(7), 1317 - 1340.

Fraser, C. S., Baltsavias, E., and Gruen, A. (2002). "Processing of Ikonos imagery for submetre 3D positioning and building extraction." *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(3), 177-194.

Fromm, F. R., and Northouse, R. A. (1976). "Class: A nonparametric clustering algorithm." *Pattern Recognition*, 8(3), 107-114.

Gamba, P. a. B. H. (2000). "Digital surface models and building extraction: A comparison of IFSAR and LIDAR Data." *IEEE Transactions on Geoscience and Remote Sensing*, 38 (4 Part 2), 1959-1968.

Graham, R. L. (1972). "An efficient algorith for determining the convex hull of a finite planar set." *Information Processing Letters*, 1(4), 132-133.

Gruen, A., and Li, H. (1995). "Road extraction from aerial and satellite images by dynamic programming." *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), 11-20.

Guienko G., D. Y. (2003). "GIS Data for Supporting Feature Extraction from High Resolution Aerial and Satellite Images." *Journal of Surveying Engineering - ASCE*, 129(4), 158-164.

Gunay, A., Arefi, H., and Hahn, M. "Semi-automatic true orthophoto production by using LIDAR data." *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007.* .

Habib, A. F., Kim, E.-M., and Kim, C.-J. (2007). "25 New Methodologies for True Orthophoto Generation." *PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING* 73.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). "Textural Features for Image Classification." *Systems, Man and Cybernetics, IEEE Transactions on*, 3(6), 610-621.

Haralick, R. M., and Shapiro, L. G. (1992). *Computer and Robot Vision*, Addison-Welsey: Reading, MA.

Hartigan, J. A., and Wong, M. A. (1979). "Algorithm AS 136: A K-Means Clustering Algorithm." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1), 100.

Hodgson, M. E., Jensen, J. R., Tullis, J. A., Riordan, K. D., and Archer, C. M. (2003). "Synergistic Use of LIDAR and Color Aerial Photography for Mapping Urban Parcel Imperviousness." *Photogrammetric Engineering and Remote Sensing*, 69(9), 973-980.

Hord, R. M. (1982). Digital Image Processing of Remotely Sensed Data, New York: Academic Press.

Hu, M. K. (1962). "Visual Pattern Recognition by Moment Invariants." *IRE Trans. Info. Theory*, IT-8, 179-187.

Huertas, A., and R. Nevatia. (1988). "Detecting buildings in aerial images." *Computer vision, Graphics and Image processing*, 41, 131-152.

Hui, Z., Rouhollah, R., Sharath, R. C., and Sally, A. G. (2006). "Local image representations using pruned salient points with applications to CBIR." Proceedings of the 14th annual ACM international conference on Multimedia, ACM, Santa Barbara, CA, USA.

Irvin, B. R., and D. M. Mckeown Jr. (1989). "Methods for exploiting the relationship between buildings abd their shadows in aerial imagery." *IEEE Transactions on systems, Man and Cybernetics*, 19(6), 1564-1575.

Jahne, B. (1991). *Digital Image Processing*, New York: Springer-Varlag.

Jarvis, C. H., and Stuart, N. (1996). "The sensitivity of a neural network for classifying remotely sensed imagery." *Computers & Geosciences*, 22(9), 959-967.

Jensen, J. R. (2000). *Remote Sensing of the Environment*, Prentice Hall.

Jensen, J. R. (2005). Introductory digital Image Processing- A remote sensing perspective, Pearson.

Jensen, J. R., Qiu, F., and Ji, M. (1999). "Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data." *International Journal of Remote Sensing*, 20(14), 2805-2822.

Ji, C. Y. (2000). "Land-Use Classification of Remotely Sensed data Using Kohonen Self-Organizing Feature Map Neural Networks." *Photogrammetric Engineering & Remote Sensing*, 66(12), 1451-1460.

Jia, L., and Wang, J. Z. (2003). "Automatic Linguistic Indexing of Pictures by a statistical modeling approach." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9), 1075-1088.

Jiang, N., Zhang, J. X., Li, H. T., and Lin, X. G. "Semi-automatic building extraction from high resolution imagery based on segmentation." *Earth Observation and Remote Sensing Applications, 2008. EORSA 2008. International Workshop on*, 1-5.

Jin, X. a. D., C. H. . (2005). "Automated Building Extraction fromHigh-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information." *EURASIP Journal on Applied Signal Processing* 14, 2196-2206.

Jing Yi, T., Yong Haur, T., and Phooi Yee, L. "One-dimensional Grey-level Co-occurrence Matrices for texture classification." *Information Technology, 2008. ITSim 2008. International Symposium on*, 1-6.

Khoshelham, K. (2004). "Building extraction from multiple data sources: A data fusion framework for reconstruction of generic models." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B3), 980-986.

Kim et al. (2004). "Semi-Automated Map Object Extraction from 1m Resolution Space Images." The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34.

Koc, D. a. M. T. "Automatic building detection from high resolution satellite images " *Proceedings of 2nd International Conference on Recent Advances in Space Technologies*, 617 - 622

Kolhe, T. H., Plumer, L. and A. B. Cremers. (2000). "Identifying buildings in aerial images using constraint relaxation and variable elimination." *IEEE, Intelligent Systems and Their Applications*, 15(1), 33 - 39.

KONECNY, G. (1985). "The International Society for Photogrammetry and Remote Sensing-75 Years Old, or 75 Years Young" *Photogrammetric Engineering & Remote Sensing*, 51(7), 911-933.

Konecny, G. (2003). Geoinformation: Remote Sensing, Photogrammetry and Geographic Information Systems, CRC Press.

Kuplich, T. M., Curran, P. J., and Atkinson, P. M. "Relating SAR image texture and backscatter to tropical forest biomass." *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, 2872-2874 vol.4.

Laha, A., Pal, N. R., and Das, J. (2006). "Land cover classification using fuzzy rules and aggregation of contextual information through evidence theory." *Geoscience and Remote Sensing, IEEE Transactions on*, 44(6), 1633-1641.

Lee, D. S., and Shan, J. "Generalization of Building Polygons Extracted from IKONOS Imagery." *Symposium on Geospatial Theory Processing and Applications. ISPRS Commission IV*, Ottawa, Canada.

Lee et al. "Class-Guided Building extraction from IKONOS Imagery." Photogrammetric Engineering & Remote Sensingat the 2002 ASPRS Annual Conference, 143-150.

Li, H. Y., Wang, H. Q., and Ding, C. B. "A New Solution of Automatic Building Extraction in Remote Sensing Images." *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, 3790-3793.

Liang, T., Weixin, X., and Jianjun, H. "Automatic high-rise building extraction from aerial images." *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on*, 3109-3113 Vol.4.

Lo, C. P., and Yeung, A. K. (2002). *Concepts and Techniques of Geographic Information Systems*, Prentice-Hall, Upper Saddle River, NJ.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwire, K. C., and Tinney, L. R. (1991). "Remote Sensing and Geographic Information system Data Integration: Error Source and Research Issues." *Photogrammetric Engineering & Remote Sensing*, 57(6), 677-687.

M, A. P., and Lewis, P. (2000). "Geostatistical Classification for Remote Sensing: An Introduction." *Computers & Geosciences* 26, 361-371.

M¨uller, S. a. Z., D. W "Robust building detection from aerial images." *IAPRS*, Vienna, Austria.

Madani, M. "Importance of Digital Photogrammetry for a complete GIS." *5th Global Spatial data Infrastructure Conference*, Cartagena, Columbia

Makhfi, P. (2007). "Introduction to Knowledge Modeling."

Marr, D., and Hildreth, E. (1980). "Theory of edge detection." *In Proceeding, Royal Society of London* B(207), 186-217.

McCulloch, W. S., and Pitts, W. (1943). "A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biology*, 5, 115-133.

McIver, D. K., and Friedl, M. A. (2002). "Using prior probabilities in decision-tree classification of remotely sensed data." *Remote Sensing of Environment*, 81(2-3), 253-261.

Ming, D. e. a. "Features based parcel unit extraction from high resolution image." Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International, pp: 1875 - 1878

Mukundan, R., and Ramakrishnan, K. R. (1998). *Moment Functions in Image Analysis - Theory and Applications* World Scientific Publishing Company, Incorporated.

Nardinocchi, C., Scaioni, M. and G. Forlani. "Building extraction from LIDAR data." *Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Workshop* 79 - 84

Ohlhof, T. e. a. (2004). "Semi-Automatic Extraction of Line and Area Features from Aerial and Satellite Images." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *34*(*XXX*).

Ozesmi, S. L., and Bauer, M. E. (2002). "Satellite remote sensing of wetlands." *Wetlands Ecology & Management*, 10(5), 381-402.

Prokop, R. J. a. R., A. P. (1992). "A survey of moment-based techniques for unoccluded object representation and recognition." *CVGIP: Graph. Models Image Process*, 54(5), 438-460.

Qiu, F., and Jensen, J. R. (2004). "Opening the black box of neural networks for remote sensing image classification." *International Journal of Remote Sensing*, 25(9), 1749-1768.

Rosin, P. L. (1999). "Measuring rectangularity." *Mach. Vision Appl*, 11(4), 191-196.

Rosin, P. L. (2000). "Shape Partitioning by Convexity." *IEEE Trans. Systems, Man, and Cybernetics,*, vol. 30 (no. 2), 202-210.

Rosin, P. L. (2003). "Measuring shape: ellipticity, rectangularity, and triangularity." *Mach. Vision Appl*, 14(3), 172-184.

Rottensteiner, F., Trinder, J. and S. Clode. (2005). "Data acquisition for 3D city models from LIDAR extracting buildings and roads." *Proceedings from Geoscience and Remote Sensing Symposium, IEEE International, 2005 (IGARSS '05),* , 1, 521-524.

Sahar, L. a. K., A (1999). "Semiautomatic extraction of building outlines from large-scale aerial images.". *Photogrammetric Engineering and Remote Sensing*, 65(4), 459-465.

Samaniego, L., Bardossy, A., and Schulz, K. (2008). "Supervised Classification of Remotely Sensed Imagery Using a Modified <formula formulatype="inline"><tex>k</tex></formula>-NN Technique." *Geoscience and Remote Sensing, IEEE Transactions on*, 46(7), 2112-2125.

Sarkar, S. a. K., L. Boyer (1991). "An Optimal infinite impulse response edge detection filters." *IEEE Transaction on Pattern Analysis and machine Intelligence*, 13(11), 1154-1171.

Schowengerdt, R. A. (1997). Remote Sensing: Models and Methods for Image Processing, 2nd ed, Academic Press, San Diego, CA.

Schweitzer, H. a., Straach, J (1998). "Utilizing Moment Invariants and Grobner Bases to Reason about Shapes." *Computational Intelligence*, 14, 461-474

Scientific, C. (2007). "BrainMaker Neural Network Application Examples." California Scientific.

Shi, J., and Malik, J. (2000). "Normalized cuts and image segmentation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8), 888-905.

Shi, Z. a. R. S. (1996). "Towards automated house detection from digital stereo imagery for GID database revision." *International Archives of Photogrammetery and Remote Sensing*, 3 (B4), 780-785.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). "Content-based image retrieval at the end of the early years." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12), 1349-1380.

Sobel, I. E., and masters, f. (1970). "Camera Models and machine perception," Stanford, Stanford, CA.

Soh, L. K., and Tsatsoulis, C. (1999). "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices." *Geoscience and Remote Sensing, IEEE Transactions on*, 37(2), 780-795.

Sohn, G. a. D., I. "Building Extraction using LiDAR DEMs and IKONOS Images." Proceedings of the ISPRS working group III/3 workshop, `3-D reconstruction from airborne laserscanner and InSAR data', Dresden, Germany.

Sohn, H. G., Park, C. H., Kim, H. S. and J. Heo. "3-D building extraction using IKONOS multispectral images." *Proceedings from Geoscience and Remote Sensing Symposium, IEEE International, 2005 (IGARSS '05), , 1432 - 1434.*

StatSoft, I. (2003). *Neural Networks*.

Sung Chun, L., Price, K., Nevatia, R., Heinze, T., and Irvine, J. M. "Semi-automated 3-D Building Extraction from Stereo Imagery." *Applied Imagery and Pattern Recognition Workshop, 2006. AIPR 2006. 35th IEEE*, 34-34.

Tao, C., Chapman, M. A., and masters, f. "Semiautomatic object pointing using multi-image matching for mobile mapping." *ACSM/ASPRS Annual convention*, 265-274.

Theiler, J., and Gisler, G. (1997). "A contiguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation." *Proc SPIE* 3159, 108-118.

Tremeau, A., and Borel, N. (1997). "A region growing and merging algorithm to color segmentation." *Pattern Recognition*, 30(7), 1191-1203.

Vosselman, G. (1998). "Cartographic Feature Extraction <http://www.itc.nl/personal/vosselman/papers/vosselman1998.bonn.pdf>." Institut für Photogrammetrie and Landesvermessungsamt NRW, Bonn, Germany, 1-10. Accessed 10/9/2006.

Wang, J. Z., Jia, L., Gray, R. M., and Wiederhold, G. (2001). "Unsupervised multiresolution segmentation for images with low depth of field." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(1), 85-90.

Wei, Y., Zhao, Z. and J. Song. "Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection." *Proceedings from Geoscience and Remote Sensing Symposium, IEEE International, 2004 (IGARSS '04)*, 2008 - 2010

Weidner, U. a. F. (1995). "Towards automatic building extraction from high resolution digital elevation model." *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), 38-44.

Wijnant, J. a. S., T. "Per-Parcel Classification of Urban IKONOS Imagery." *7th AGILE Conference on Geographic Information Science, Heraklion, Greece* 447-455.

Xiao, Y., Lim, S.K., Tan, T.S. and S. C. Tay. "Feature extraction using very high resolution satellite imagery." *Proceedings from Geoscience and Remote Sensing Symposium, IEEE International*, 2004 - 2007.

Yixin, C., and James, Z. W. (2004). "Image Categorization by Learning and Reasoning with Regions." *J. Mach. Learn. Res.*, 5, 913-939.

Yixin, C., and Wang, J. Z. (2002). "A region-based fuzzy feature matching approach to content-based image retrieval." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(9), 1252-1267.

Zebedin, L., Klaus, A., Gruber-Geymayer, B., and Karner, K. (2006). "Towards 3D map generation from digital aerial images." *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(6), 413-427.

Zhengjun, L., Shiyong, C., and Qin, Y. "Building extraction from high resolution satellite imagery based on multi-scale image segmentation and model matching." *Earth Observation and Remote Sensing Applications, 2008. EORSA 2008. International Workshop on*, 1-7.

Zhu, S. C., and Yuille, A. (1996). "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(9), 884-900.

Ziyu, W., Wenxia, W., Shuhe, Z., and Xiuwan, C. "Object-oriented classification and application in land use classification using SPOT-5 PAN imagery." *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, 3158-3160 vol.5.