

**AUTOMATED BUILDING FOOTPRINT EXTRACTION FROM
HIGH RESOLUTION LIDAR DEM IMAGERY**

A THESIS
PRESENTED TO
THE FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF MISSOURI-COLUMBIA

In Partial Fulfillment
Of the Requirements for the Degree

Master of Science

By
Mandar M. Gadre
Dr. Curt Davis, Thesis Supervisor

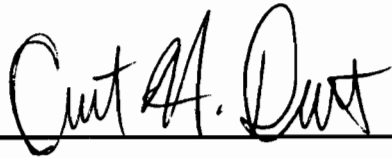
July 2005

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

**AUTOMATED BUILDING FOOTPRINT EXTRACTION FROM
HIGH RESOLUTION LIDAR DEM IMAGERY**

Presented by Mandar Gadre,

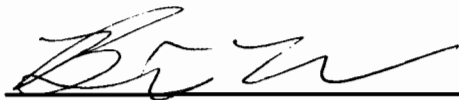
a candidate for the degree of Master of Science and hereby certify that in their opinion is worthy of acceptance.



Dr. Curt Davis, Thesis Advisor,
Professor,
Department of Electrical Engineering.



Dr. Justin Legarsky
Assistant Professor,
Department of Electrical Engineering.



Dr. Brian Adams
Assistant Professor,
Department of Agricultural Systems Management

ACKNOWLEDGEMENTS

First of all I would like to thank my advisor, Dr. Curt Davis. Without his enthusiasm, his inspiration, and his great efforts to explain things clearly and simply, this work would not be possible. Not only he suggested me the research topic but also provided encouragement, sound advice, good teaching and a lot of good ideas.

Without him I would have been lost. I wish to thank my thesis committee members Dr. Justin Legarsky and Dr. Brian Adams, for their valuable advice. I would also like to thank Dr. Aaron Shakelford, for his invaluable help and time.

I would like to thank my family back in India for all the support and the encouragement given by them. Last but not the least, I would also like to thank all my friends, especially Kapil, Arpit and Milind who have always been there for me during my graduate studies.

TABLE OF CONTENTS

Acknowledgements.....	ii
List of Illustrations.....	vi
List of Tables.....	viii
Abstract.....	ix
1. Introduction.....	1
1.1 Geographical Information products and Applications.....	1
1.2 High Resolution Satellite Imagery.....	2
1.2.1 Aerial Photography.....	4
1.2.2 Landsat.....	5
1.2.3 Commercial High Resolution Satellites.....	6
1.2.4 Image Pan Sharpening.....	7
1.3 Digital Elevation Models (DEMs).....	8
1.3.1 Low Resolution DEMs.....	10
1.3.2 High Resolution DEMs.....	10
1.4 Feature Extraction.....	12
1.5 Summary.....	13
1.6 Objective.....	13
2. Background.....	15
2.1 Introduction.....	15
2.2 Literature Review.....	15
2.2.1 Building Extraction using High Resolution Imagery.....	15
2.2.2 Building Extraction using DEM.....	17
2.2.3 Overview of initial work and results.....	18
2.2.3.1 DRAW Algorithm.....	18

2.2.3.2 RUBBERBAND Algorithm.....	19
2.3 Summary.....	20
2.4 Research Goals.....	21
3. Image processing Methodology.....	22
3.1 Introduction.....	22
3.2 Processing Overview	23
3.3 Image Preprocessing.....	25
3.3.1 Interpolation.....	26
3.3.2 Median Filtering	27
3.4 Differential Morphological Profiles.....	29
3.4.1 Morphological Opening and Closing.....	29
3.4.2 Opening and Closing by Reconstruction.....	30
3.4.3 Differential Morphological Profile.....	32
3.5 Results and Discussion.....	34
3.6 Summary	38
4. Feature Extraction & Fusion.....	40
4.1 Introduction	40
4.2 Fuzzy S and Π membership functions.....	40
4.2.1 Fuzzy S membership functions.....	41
4.2.2 Fuzzy Π membership functions.....	42
4.3 Minimum bounding rectangles.....	43
4.4 Fuzzy rectangular measure.....	45
4.5 Fuzzy Branchiness Measure.....	49
4.6 Multi Fusion Detection.....	51

4.7	Accuracy Assessment.....	53
4.7.1	Pixel Based Assessment	54
4.7.2	Object-Based Assessment.....	55
4.8	Summary.....	56
5.	Results and Discussions.....	59
5.1	Introduction	59
5.2	Results and Discussion.....	59
5.3	Summary and Conclusion.....	65
6.	Summary and future Work.....	66
6.1	Summary.....	66
6.2	Future Work.....	69
7.	Reference List.....	70

LIST OF ILLUSTRATIONS

FIGURE	PAGE
1.1. Panchromatic USGS DOQQ of St. Charles, MO.....	4
1.2. False color 7, 5, 3 band combination Landsat TM image....	5
1.3. color infrared IKONOS MS image of Columbia, MO.....	6
1.4. PAN IKONOS images of Columbia, MO.....	7
1.5. Airborne LIDAR system.....	11
1.6. LIDAR DEM image of Springfield, MO.....	11
3.1. Flowchart of automated building extraction approach.....	24
3.2. DEM subsection before and after interpolation.....	27
3.3. DEM subsection before and after median filtering.....	28
3.4. Example of Morphological Dilation.....	30
3.5. Example of Morphological Erosion.....	30
3.6. DEM subsection with DMPs.....	37
4.1. Fuzzy S-membership function.....	41
4.2. Fuzzy Π -membership function.....	42
4.3. Minimum Bounding Rectangles for some closed curves.....	43
4.4. DEM subsection and output of MBR measure.....	45
4.5. Image segments with their skeletons.....	46
4.6. Output of fuzzy rectangular shape measure on DEM.....	49

4.7	Output of fuzzy branchiness shape measure on DEM.....	51
4.8	Output after Multi Detector fusion.....	53
5.1	Original Image and final building output.....	60
5.2	Validation Image and final building output.....	61

LIST OF TABLES

TABLE		PAGE
1.1	Characteristics of Various Optical Remote Sensing Systems.....	3
1.2	Various Satellite Sensors and corresponding DEM Resolutions.....	10
3.1	Typical DEM data file header.....	27
5.1	Pixel Based and Building Based Accuracy Statistics for Development Image 1.....	59
5.2	Pixel Based and Building Based Accuracy Statistics for Validation Image 1.....	60
5.3	Pixel Based Accuracy Statistics for Development/Validation Images.....	61
5.4	Building Based Accuracy Statistics for Development/Validation Images.....	62

Automated Building Footprint Extraction from High Resolution LIDAR DEM Data

Mandar M. Gadre

Dr. Curt Davis, Thesis Supervisor

Abstract

Geographic Information Systems (GIS) are used in the fields of urban planning, environmental management, agriculture, transportation, utilities etc. because of their ability to provide geospatial information organized in multiple layers such as digital image basemap, land use zoning, political boundaries, parcel maps, land cover, road network, building footprints, utility networks (e.g. water, sewage and electricity), topography, and green space. Some urban features like roads and buildings change with the time and it is therefore necessary to update this information. The goal of this research is to provide a robust automated method to extract commercial buildings from the high resolution DEM data with high quality, accuracy, and detection rates. This processing strategy uses three different detectors which are fused to obtain a final output. Though multi-detector fusion has been used previously for satellite imagery, it is completely new for the DEM data. All three algorithms are developed using a fuzzy logic approach. The results of our algorithm show that we have obtained 82% correctness, 73% completeness and 65% quality pixel wise and 82% correctness, 97% completeness and 65% quality object wise for the tuning images and similar results for the test images. This approach can be expanded for the extraction of residential buildings which is left for future work.

Chapter 1

Introduction

1.1 Geographical Information Products and Applications

Since more than 50% of the world's population lives in urban and suburban areas, detailed and up-to-date building and road information is of great importance to residents, government agencies, and private companies (utilities, real estate, etc.).

For the past 20 years, Geographic Information Systems (GIS) have been broadly applied in fields such as urban planning, environmental management, agriculture, transportation, utilities, etc. GIS (for more details see Sheckelford, 2004) is basically an integrated database where each piece of information has a geographic location associated with it. Information is displayed in a variety of spatial formats and is organized in information layers such as image basemap, land use zoning, political boundaries, parcel maps, land cover, road network, building footprints, utility networks (e.g. water, sewage and electricity), topographic networks and green space.

Therefore it is not surprising that public government agencies and private companies spend millions of dollars each year to convert paper maps into digital format suitable for use in GIS. However, many of these maps and much of the resulting GIS data are obsolete because though some features do not change much with the time (e.g. elevation), some features change frequently due to construction of new roads and buildings. Due to its temporal capability and ability to create accurate maps, remotely

sensed imagery is one of the most efficient and widely used ways to acquire, extract new features, and update the old data.

1.2 High Resolution Satellite Imagery

Remote sensing from space began in April 1960 with the launch of the Television and Infrared Observational Satellite (TIROS-1) by the USA as an experimental weather satellite. In 1972 the first civilian satellite (Earth Resources Technology Satellite ERTS-1) designed specifically to collect information about the earth's surface and resources was launched. Since then, a number of satellite systems have been developed, such as LANDSAT TM (Thematic Mapper), French SPOT HRV (High Resolution Visible), Russian SPIN, Indian IRS, Japanese Earth Resources Satellite (JERS), etc.

Spatial and temporal characteristics of several other high resolution satellite systems are listed in Table 1. Alternatives to passive remote sensing systems that record naturally reflected electromagnetic radiation are active sensors such as radar, LIDAR, etc. They are used to collect earth information from a radar or light pulse of energy emitted from the sensing device.

Table 1.1: Characteristics of Various Optical Remote Sensing Systems (for more details see Song, 2001)

	Spatial Resolution (meters)	Temporal Revisit (days)	Spectral Band (μ m)
NOAA AVHRR	1100	1	0.58-12.5
LANDSAT MSS	79	18	0.50-1.1
LANDSAT TM	30	16	0.45-1.75
LANDSAT ETM+	15(pan) & 30(MSS)	16	0.45-12.5
SPOT HRV	10(pan) & 20 (MSS)	Variable	0.50-0.89
IRS-1A	72	22	0.45-0.86
IRS-1B	36.25	22	0.45-0.86
JERS-1	18*24	44	0.55-0.86
SPIN KVR-1000	2	9	0.49-0.59
SPIN TK-350	10	9	0.49-0.59
QuickBird	0.82 & 3.28	1-5	0.45-0.90
IKONOS(MSS)	1 (pan) & 4(MSS)	1-3	0.45-0.90
Orbview 3&4	1 (pan) & 4(MSS)	< 3	0.45-0.90
MODIS	250-1000	2	0.40-14.54
ERS-1	30	35	5.3 GHZ C-band
RADARSAT	25*28	4-6	5.3 GHz C-band

In the following section we present a brief overview of different imagery sources.

1.2.1. Aerial Photography

Digital Orthophoto Quarter Quadrangles (DOQQ) produced by the USGS are the most widely available high-resolution digital orthoimage dataset available to local governments in the U.S. These are digital images generated from an aerial photograph where the distortions due to terrain and sensor displacements have been removed giving it the positional accuracy and geometric qualities of map. DOQQs are available in either panchromatic (black/white), natural color, or color infrared with 1 m ground resolution and 8-bit per pixel dynamic range.



Figure 1.1: Panchromatic USGS DOQQ of St. Charles, Missouri (1-m resolution)

1.2.2. Landsat

Landsat 5 and Landsat 7 were the first widely available sources of satellite-based imagery, which collect data with 5 (Multi Spectral Scanner MSS and Thematic Mapper TM) and 7 sensors (Enhanced Thematic Mapper ETM+) respectively . Bands 1-5 and 7 have a spatial resolution of 30 m while band 6 (ETM+, TM thermal infrared band) has a spatial resolution of 60 m. The panchromatic band (PAN) present on the ETM+ sensor has a spatial resolution of 15 m.

Thus Landsat imagery has many applications in agriculture, geology, geography, water quality, etc but has very limited use in urban applications because of its inability to map man-made features like roads and buildings.



Figure 1.2: False color 7, 5, 3 band combination Landsat TM image of St. Charles (30 m resolution)

1.2.3. Commercial High Resolution Satellites

On September 24, 1999 the world's first privately owned commercial, high resolution satellite - IKONOS - was successfully launched by Space Imaging Inc. IKONOS has a 1 m resolution panchromatic band and 4 m resolution multi-spectral bands. Data is stored with 11-bits per pixel giving it more dynamic range and therefore allowing increased discrimination of subtle spectral differences between objects. IKONOS imagery is a cost effective alternative for the generation of basemaps for use by local governments.



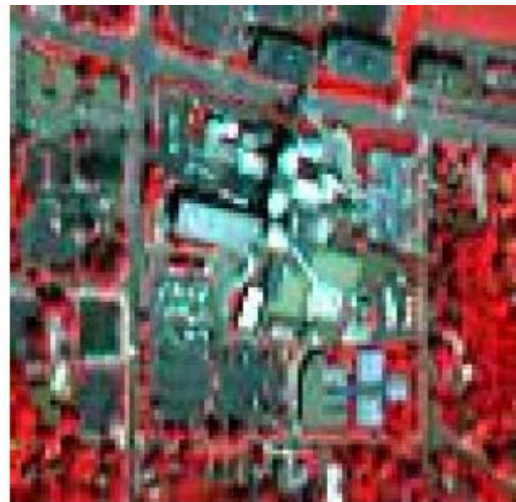
Figure 1.3 False color infrared IKONOS MS image of Columbia, Missouri (4 m resolution)

1.2.4. Image Pan Sharpening

Pan Sharpening or data fusion is the process of increasing the spatial resolution of lower resolution MS bands using higher resolution PAN band. This produces imagery with the resolution of the PAN band and the spectral characteristics of the MS bands. There are variety of PAN sharpening algorithms using wavelet-based techniques, principle component analysis, and arithmetic means.



(a)



(b)



(c)

Figure1.4: (a) PAN IKONOS image of Columbia (1 m resolution), (b) false color infrared IKONOS MS image of Columbia (4m resolution) (c) false colored infrared PS-MS IKONOS image of Columbia, Missouri

1.3 Digital Elevation Models

A digital elevation model (DEM) is a digital file consisting of terrain elevations for ground positions at regularly spaced horizontal intervals. The resolution or the distance between adjacent grid points is a critical parameter. The best horizontal resolution that is widely available is 30 m, with a vertical resolution of 1 m. The coverage of the entire globe, including the ocean floor can be obtained at lower resolutions.

DEMs can be created using several methods. Some of those are listed below:

Conversion of printed contour lines

- existing plates used for printing maps are scanned
- the resulting raster is vectorized and edited
- contours are "tagged" with elevations
- additional elevation data are created from the hydrography layer
 - i.e. shorelines provide additional contours
- a final algorithm is used to interpolate elevations at every grid point from the contour data

Photogrammetry

This can be done manually or automatically:

- i. manually, an operator looks at a pair of stereo photos through a stereo plotter and moves two dots together until they appear to be as one lying just at the surface of the ground

- ii. automatically, an instrument calculates the parallax displacement of a large number of points
- iii. e.g. for USGS 7.5 minute quadrangles, the Gestalt Photo Mapper II correlates 500,000 points

Extraction of elevation from photographs is confused by flat areas, especially lakes, and wherever the ground surface is obscured (buildings, trees). There are two techniques for choosing sample points when using manual photogrammetry:

Profiling

- iv. the photo is scanned in rows, alternately left to right and right to left, to create profiles
- v. a regular grid is formed by resampling the points created in this process
- vi. Because the process tends to underestimate elevations on uphill parts of each profile and overestimate on downhill parts, the resulting DEMs show a characteristic "herringbone" effect when contoured.

Contour following

- i. contour lines are extracted directly from stereo pairs during compilation of standard USGS maps
- ii. contour data are processed into profile lines and a regular grid is interpolated using the same algorithms used for manual profiling data

1.3.1 Low Resolution DEMs

Low resolution DEMs can be obtained from variety of sensors listed below:

Table 1.2: Various Satellite Sensors and their corresponding DEM Resolutions

<i>Satellite Sensor</i>	<i>DEM Resolution</i>
ASTER	15-25 meters
SPOT	5-10 meters
IKONOS	2-5 meters

1.3.2 High Resolution DEMs

High resolution DEMs with a 1 m horizontal resolution and a vertical accuracy of +/- 0.2 m can be generated using Light Detection and Ranging (LIDAR) technology.

An airborne sensor emits laser pulses that strike the ground and reflect to the sensor. When the precise altitude and position of the sensor is known, the elevation of the surface points can be determined from the time required for the pulse to return to the sensor. The following figure illustrates the airborne LIDAR concept.

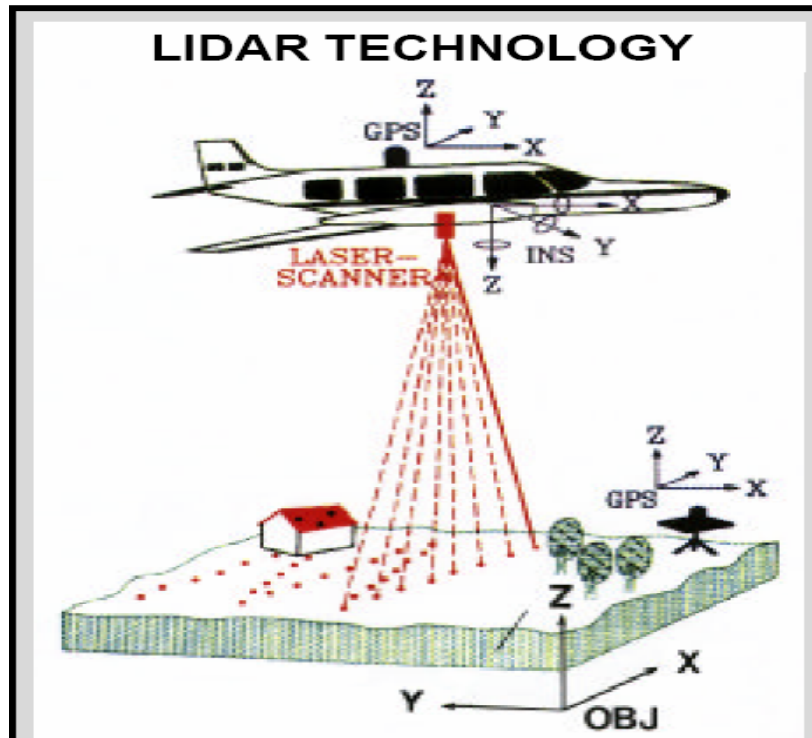


Figure 1.5: Airborne LIDAR system.

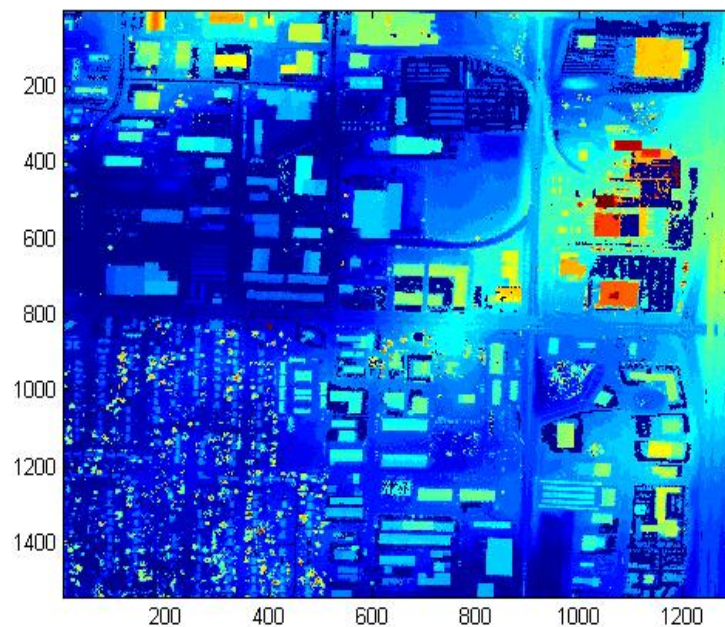


Figure 1.6: LIDAR DEM image of Springfield, Missouri with 1 m horizontal resolution.

1.4 Feature Extraction

Feature extraction is the detection, identification, classification, annotation, measurement, and delineation in two or three dimensions. The tasks of feature identification and delineation are often still done manually. These tasks are very time consuming, very labor intensive, and costly. Although man-made objects can be easily identified by human operator, it is very difficult for a computer to identify them. Semi-automatic and automatic extraction of man-made objects, which has been studied for more than 20 years in the photogrammetry and digital image processing fields, is still in its developmental stage.

One of the most promising applications of automated techniques is to assist image analysts in the feature extraction process. Here, the goal is to produce a highly accurate, but not necessarily complete extraction of the feature of interest. For example, if an automated building extraction technique is able to accurately identify 75% of the buildings present in the image being analyzed, then the image analyst only has to manually digitize 25% of the buildings, greatly reducing the time required for data production. Another application where automated extraction techniques are useful even if they are not completely accurate is information mining of remote sensing image database archives.

Because of the rapid pace of urban/suburban development, up-to-date cartographic information is required by public agencies and private companies. Among the most important features needed are buildings; hence the requirement for automated

procedures to rapidly identify and extract topologically correct building footprints from remote sensing data and/or imagery.

1.5 Summary

In this chapter, we presented a variety of background material on geospatial information products, high-resolution satellite imagery, and image processing techniques and methods. The utility of these for urban area applications was discussed. Geospatial information products developed from commercial high-resolution satellite imagery have numerous government applications at both the local and national level. For example, the USGS's National Map integrates a variety of geospatial information products (topography, roads, land cover, etc.). In addition to civilian government applications, there are many defense and intelligence applications for urban area geospatial information products. NIMA/NGA has integrated the previously disparate areas of mapping and imagery to produce a new intelligence tradecraft called geospatial intelligence or GeoINT.

Several earth image collection methods were also discussed; including both aerial photography and satellite based digital sensors. Then there was a brief overview of DEMs and their various generation methods. Also both high and low resolution DEMs were discussed followed by an overview of feature extraction.

1.6 Objective

In this research we propose an approach to automatic building extraction from high resolution DEM data using digital image processing techniques. A set of algorithms for automatically extracting buildings from the remotely sensed data is developed within

the MATLAB environment. This research focuses on extraction of commercial buildings, i.e. buildings with larger size compared to residential buildings with comparatively smaller size. Pixel-based evaluations for completeness, correctness, detection, quality, missing factor, branching ratio and building or object based evaluations for detection and quality percentages are conducted by comparing the automatically extracted buildings with manually acquired buildings.

The rest of this thesis is organized as follows. In Chapter 2, previous building extraction research using high resolution imagery and DEM data is reviewed. Image preprocessing and the differential morphological profile are discussed in Chapter 3. In Chapter 4, actual feature extraction using different shape parameters and their fusion is presented along with the accuracy measurement concepts. Chapter 5 gives the results and accuracy statistics of the feature extraction on different test images with a discussion. Finally, concluding remarks and future work for this research are discussed in Chapter 6.

Chapter 2

Background

2.1 Introduction

Research involving building extraction from satellite imagery using digital image processing techniques is a relatively new development. In the following sections we present a brief review of the most significant advances.

The review is divided into four sections. Initially, previous work done using high resolution satellite imagery like IKONOS is reviewed. Literature on building extraction using high resolution DEM data is presented. This is followed by the discussion of the DRAW and Rubber band algorithms implemented on DEM data used for this research. Chapter is concluded with the research goals set at the start of this project and provides a summary of the techniques involved.

2.2 Literature Review

2.2.1 Building Extraction using High Resolution Imagery

A technique for building extraction from supervised (i.e. semi-automated) classification of IKONOS imagery is presented by Lee *et al.* [2003]. First, a supervised classification of the MS data is performed, focusing on identification of different types of building roof materials. The classification is then used to define a window within the PAN imagery to search for buildings. Buildings are then extracted using unsupervised

clustering followed by a polygon squaring technique based on the Hough transform. Building detection rates of 64%-72% and quality values of 30%-51% are reported. Croitoru and Doytsher [2003] presented a model-based building extraction technique that relies on the detection of right-angle building corners. Buildings are detected using pose clustering, a voting technique, where right-angle corners are used as voting elements. The voting process is constrained by detection of shadowed regions. This approach is limited by the necessity of a model for every building shape that is to be detected. Building detection rates of 79%-81% and quality values of 32%-34% are reported. A technique for detecting small objects, such as house plots and vegetation patterns, in 4.5-m PAN satellite imagery based on combining supervised shape classification with unsupervised image segmentation was presented by Segal and Kaufmann (2001). This approach utilizes a thresholding segmentation technique that has been modified to identify specific shapes on which it has been trained. Extraction completeness values of 61%-97% and correctness values of 58%-94% are reported. Jaynes *et al.* (1994) present a task driven perceptual organization approach for extraction of building polygons. This approach has three steps: low-level feature extraction of line segments and orthogonal corners, perceptual grouping of those features into chains, and polygon hypothesis generation from the chains. The building extraction algorithm was tested on high-resolution aerial photography and resulted in the detection of 87% of the buildings; however the accuracy of the extracted buildings was not reported. A technique based on extracting buildings from high-resolution PAN aerial photography by first identifying their cast shadows was developed by Irvin and McKeown (1989). Building hypotheses are initially generated by grouping detected edges and corners into polygons. Shaded regions are then identified

using a variety of image processing techniques, followed by identification and simplification of the shadow edge. The shadow edges are then combined with the building hypotheses to identify buildings. Building extraction accuracy measures for this approach were not reported.

2.2.2 Building Extraction using High Resolution DEM data

LIDAR elevation data, which is acquired by airborne laser scanners, has been used as an attractive alternative to aerial photography due to high vertical accuracy and high point density. As a single source, LIDAR data has been used to reconstruct various types of building shapes; parametric models (Mass & Vosselman, 1999; Wang & Schenk, 2000), prismatic models with flat roofs (Weidner & Förstner, 1995), and polyhedral buildings with the restriction on the building orientation (Vosselman, 1999). Haithcoat, Song, and Hipple (2001) proposed an approach to extract buildings by generating a digital surface model (DSM) from LIDAR data and then using watershed segmentation analysis. Due to limitations of DEM data for extracting comparatively small structures like houses, DEM data can be combined with other source imagery like IKONOS to produce more efficient building extraction (Brunn and Weidner, 1997; Haala and Brenner, 1999).

2.2.3 Overview of Initial Work Done

Two basic automatic building extraction algorithms named “DRAW” and “RubberBand” were developed for Dr. Davis by an undergraduate research assistant. However, because of time limitations, these algorithms were not fully explored and tested. These algorithms are briefly described here.

2.2.3.1 DRAW Algorithm

First the gradient is computed for each point in the DEM grid and then all points above a certain threshold are marked. The gradient values, i.e. slopes, indicate the height of any vertical structure relative to the ground level. Since the DEM data contains artifacts and noise, this slope threshold is useful to eliminate some noise in the image. However, selection of the appropriate threshold is dependent on the typical slope produced by the building type one wishes to extract.

Once the points with a slope above the threshold are determined, the DRAW algorithm is applied to each point. This algorithm vectorizes linear features present in the DEM by starting at a DEM grid point whose slope is above the threshold. The algorithm then searches radially around the current DEM point for linear, contiguous strings of grid points above the threshold. If multiple strings are found, the algorithm selects the longest one and records the starting and ending points of the selected string. The algorithm then recursively calls itself with the end point of the selected segment as the new current DEM grid point. To avoid infinite recursion the algorithm keeps track of all points in a string selected as longest.

This algorithm has some major disadvantages. Since it's a searching algorithm it is very time consuming especially for large datasets. Also the selection of an appropriate threshold is a problem because some buildings can be inadvertently filtered out with the noise. If the threshold is slightly higher than some parts of a building it will be represented by broken line segments. Adaptive thresholding is one potential solution but this increases the complexity of the algorithm. To overcome some of these difficulties the "Rubber-Band" algorithm was proposed.

2.2.3.2 Rubber-Band Algorithm

This algorithm was inspired by its namesake, which, once stretched out, will contract and conform to the shape of whatever it encircles. This algorithm begins with the calculation of the centroid or geometric center, of a group of above threshold grid cells (i.e. grid points on the boundary of a building). The algorithm then examines grid points located at a distance, r , and angle θ , from the centroid. (Initially, r is set to a relatively large value r_0 . The initial value of θ doesn't matter). If this test point belongs to the building under current consideration, the location of the point is recorded. In this case θ is incremented by an angle $\Delta\theta$, r is again set to r_0 and then the point at the distance r and new angle θ is recorded. On the other hand, if the point does not belong to the building outline under current consideration, r is decremented, and the new test point is examined as before. This is done until a building point is found or until r is decreased to zero. In either case θ is incremented by $\Delta\theta$, r is again set to r_0 , and the process repeats until θ reaches 2π radians.

The main advantage of the rubber-band algorithm is it generates simple, closed polygons. Also, it is free of the broken lines problem of the Draw algorithm. Unfortunately two complications limit the effectiveness of this algorithm. First, if the centroid of the building lies outside the building (e.g. “L” shaped buildings) this algorithm produces a polygon with few or no vertices along the portion of the building outline. Sometimes the same building number is assigned to two different buildings closely lying near each other. Thus, due to real world noise and vegetation (tall trees) present in the DEM, the rubber-band algorithm often merges two or more buildings into one single polygon.

Unfortunately both of these algorithms were time consuming and there were no accuracy statistics available. Also neither of these two algorithms proved robust enough for further modifications. Hence, the research presented in the next few chapters is based on a completely new approach.

2.3 Summary

Here we briefly reviewed recent literature regarding semi-automated and automated building extraction. Various studies using high resolution imagery like IKONOS were discussed along with their accuracy statistics. Since DEMs have their own advantages like high point density, good vertical accuracy, and a 3-dimensional building perspective, we reviewed some of the recent building extraction research using DEM data. Finally, we presented an overview of the DRAW and Rubber Band algorithms developed as initial attempts for building extraction from LIDAR DEMs.

2.4 Research Goals

The goal set at the start of this research was to design a robust algorithm to extract comparatively large man-made structures with a high accuracy, quality, and detection rates. As said before, the approach developed here is completely new and not used before for DEM data. The final building extraction algorithm has many steps but it is simple to implement and produces high detection and quality percentages. The following two chapters will discuss the algorithm in detail.

Chapter 3

Image Processing Methodology

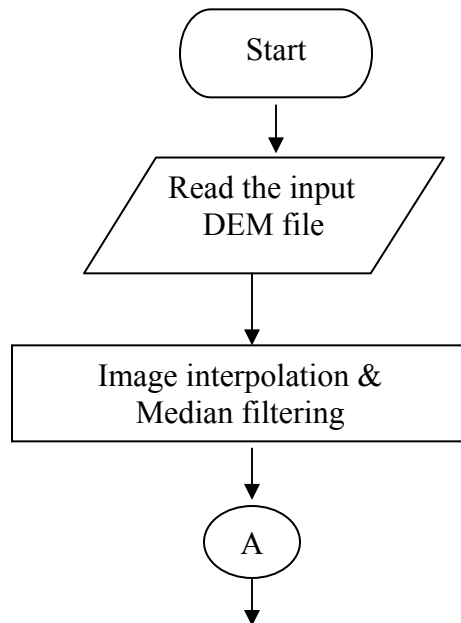
3.1 Introduction

Chapters 3 and 4 explain the image processing methodologies developed in this research for automated building extraction from high-resolution lidar DEMs. Chapter starts with the description of the overall processing chain. Then the image preprocessing strategy that is needed to remove noise and/or replace erroneous data is presented. These problems are tackled with Image interpolation and median filtering techniques. The image interpolation is used to correct for image artifacts, while median filtering is used to remove or lessen noise present in the DEM data.

Then the most important step in the processing chain, i.e. the Differential Morphological Profile (DMP) is discussed. Results for this step are presented after detailed explanation of the theory. The basic morphological operations like opening and closing are presented initially and then are extended to describe opening profiles and differential opening profiles. The chapter concludes with the results subsection where actual DMP output is shown for structural elements with different radii.

3.2 Processing Overview

Figure 3.1 presents a flowchart summarizing all major steps in the automated building extraction processing chain. After the image is read into the input matrix, the first step is interpolation which is needed to recover missing data points or “No Data Value” pixels. Next, the interpolated image is median filtered to remove random noise present in the image. After image preprocessing, eleven levels comprising the “Differential Morphological Profile” (DMPs) are calculated using disk structuring elements with variable radii. After this, three different shape measures, fuzzy rectangularness, fuzzy minimum bounding rectangle, and fuzzy branchiness are applied to the DMP levels. Three different images are formed - one each for each shape measure - by calculating the maximum value at each pixel for all the DMP levels. Finally, the three images are fused together to form a final image with building polygons represented as different objects. The following sections describe each block in more detail.



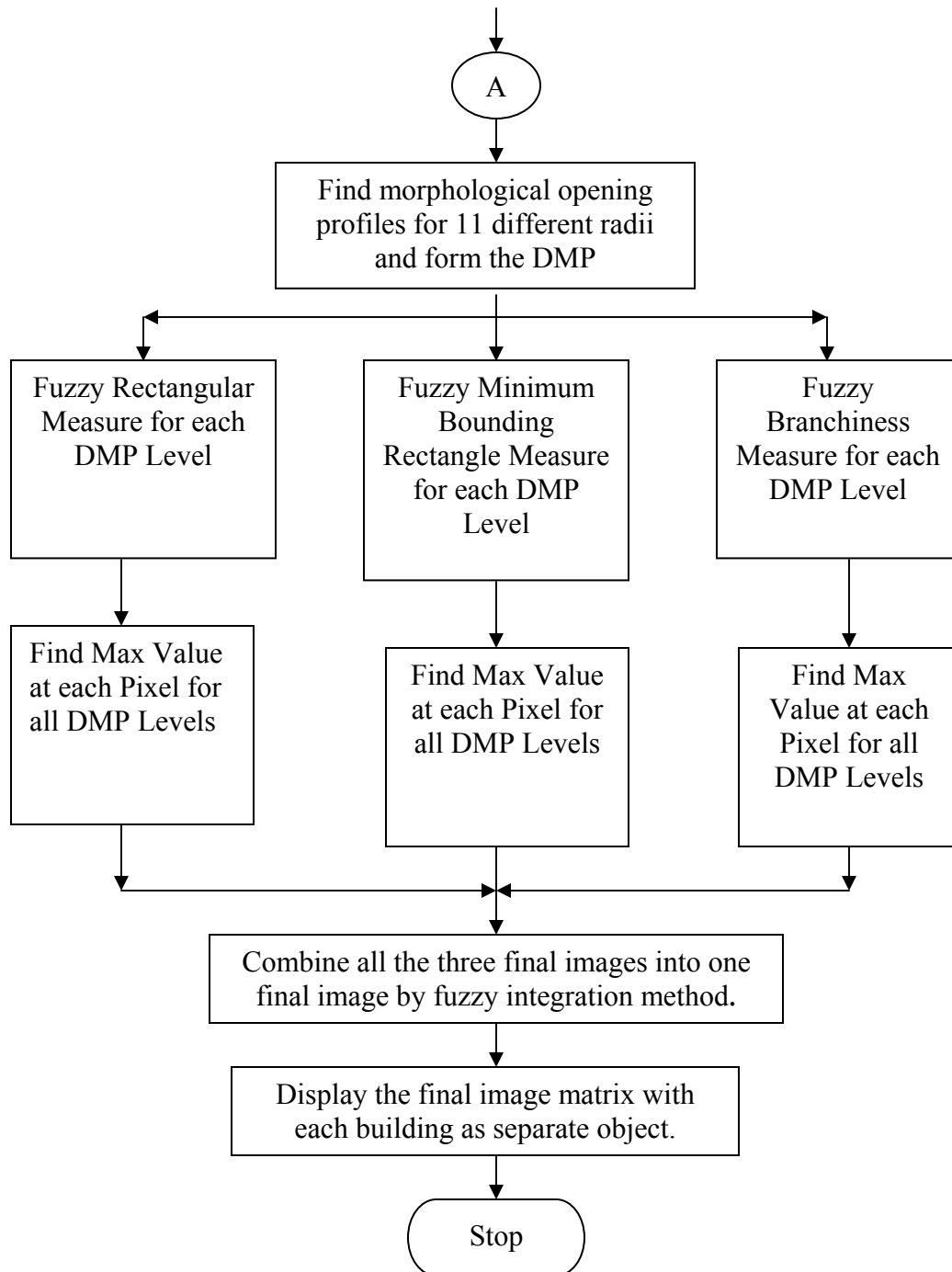


Figure 3.1: Flowchart illustrating processing chain for automated building extraction approach.

3.3 Image preprocessing

The raw input DEM file is a text file that was generated by exporting an ARCInfo binary format to an ASCII format. The ASCII file output by ARCInfo begins with a six line header containing information about the DEM, followed by space-separated DEM data points. An example header is as follows:

ncols	340
nrows	688
xllcorner	421970.5
yllcorner	149330.5
cellsize	1
NODATA_value	-9999

Table 3.1: Typical DEM data file Header

The number of rows and columns in the input DEM file can be obtained from the header lines. In the MATLAB programming environment, the ASCII file can be directly read into a matrix format if the header lines are removed. Thus, the header lines should be eliminated leaving only data values in the ASCII file. The DEM data values represent the height of each data location above sea level. Thus, when the input DEM is represented by a black and white raster (2D) image, the pixel brightness indicates the relative height at each DEM location, i.e. brighter pixels represent larger elevation values.

3.3.1 Interpolation

The DEMs used in this research were generated from LIDAR technology. While this method of data collection is cost effective and accurate, the raw data file contains many missing data points, i.e. “NODATA_values” (data points with value -9999). These missing points are caused by surfaces that reflect most of the incident energy from the laser pulse away from the sensor. Valid pixels have values greater than zero indicating that these points have elevations above sea level. Missing data points will limit the effectiveness of any building extraction algorithm. Thus, this problem is addressed by using an image interpolation strategy.

Initially all pixels with values -9999 are identified. For each “no data” pixel, a 3*3 neighborhood is searched for valid data points. If no valid data point is found then the mask size is increased to 4*4 and again the neighborhood values are searched. This process is repeated until a data value is successfully found in the neighborhood. After this, the average of all valid data values is calculated and the original pixel and all the searched pixels with NODATA_values are replaced by that value. Here instead of the average value, the pixel can be replaced by a binary interpolated value. The process of interpolation is explained with the help of following figures.

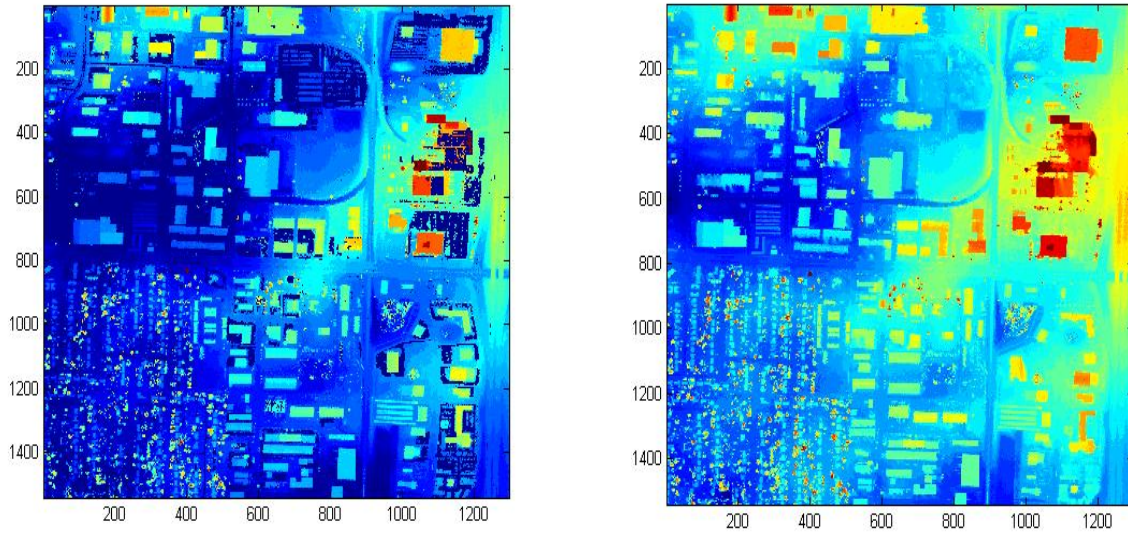


Figure 3.2: DEM sub-section (a) before interpolation and (b) after interpolation.

The effect of interpolation on the image can be easily observed especially in the case of parking lots which tend to have “No Data Values”.

3.3.2 Median Filtering

Another problem that can limit the effectiveness of the building extraction algorithm is random noise present in the image. This problem is mitigated using a traditional median filtering technique.

The median filter is one of the most widely used order statistic filters. Order statistic filters are based on ordering the pixels (i.e. either ascending or descending) contained in the filter’s window and then replacing the value of the center pixel by the ranking result. In the case of median filtering, the center pixel is replaced by the median of the pixels in its immediate neighborhood. For certain types of random noise like salt and pepper noise, median filtering provides excellent noise reduction capabilities with considerably less blurring.

Here, the center pixel and its 3*3 neighborhood are sorted and then the center pixel is replaced by the median value, i.e. the 5th value in the rank order. The principle function of the median filter is to force the points with distinct grey levels to be more like their neighbors. Isolated pixel clusters that are light or dark with respect to their neighbors and whose area is less than $n^2/2$ (one half of the filter area), are eliminated by an $n*n$ median filter. Large clusters are affected considerably less.

The figure below illustrates the effect of median filtering on a sub-section of the lidar DEM.

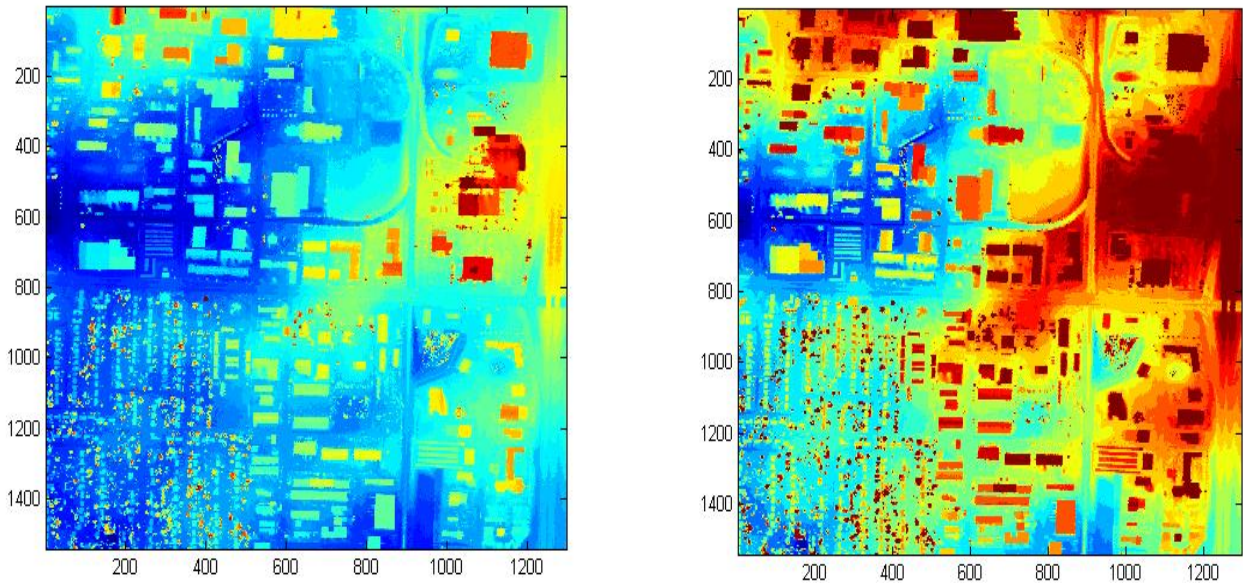


Figure 3.3: DEM subsection (a) before median filtering and (b) after median filtering

3.4 Differential Morphological Profile

A differential morphological profile (DMP) is the most important component in the automated building extraction processing chain because it extracts bright and dark structures in the image depending on their shape and size. Mathematical morphology employs a set of image operators to extract and analyze image components based on the shape and size of quasi-homogeneous regions in the image. Mathematical morphology has been applied to a wide variety of practical problems such as image pre-processing, noise filtering, shape detection and decomposition, and pattern association. The DMP is a relatively new technique. It was first introduced by Pesaresi and Benediktsson (2001) who used it to segment high-resolution satellite imagery. The DMP has also been used for urban land cover classification by Benediktsson *et al.* (2003). The DMP is a multi-scale image analysis technique where a morphological profile of the image is constructed through the use of morphological opening and closing by reconstruction operations while varying the size of the structuring element (SE). The following sections introduce the concepts of morphological opening and closing, then opening by reconstruction, and finally the DMP itself.

3.4.1 Morphological Opening and Closing

Morphological dilation of A by B , denoted $A \oplus B$, is defined as

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (1)$$

The dilation of A by B then is the set of all displacements, z , such that B and A overlap by at least one element. Set B is commonly referred to as the **structuring element** in all morphological operations.

The process of dilation can be visually illustrated as

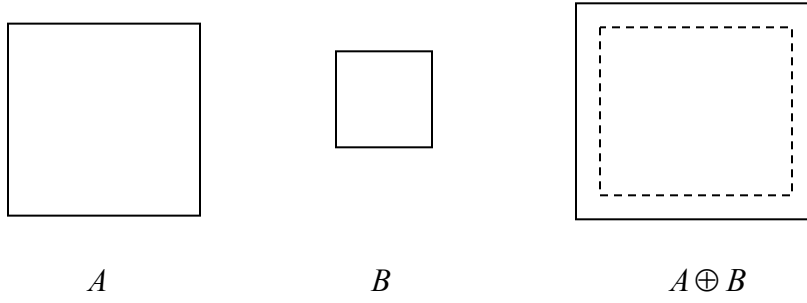


Figure 3.4: Example of Morphological Dilation

One of the simplest applications of dilation is bridging the gap between two points.

Similarly morphological erosion of A by B , denoted $A \ominus B$, can be defined as

$$A \ominus B = \{z \mid (\hat{B})_z \subseteq A\} \quad (2)$$

Thus erosion of A by B is the set of all points z such that B , translated by z , is contained in A . Visually, erosion can be illustrated as

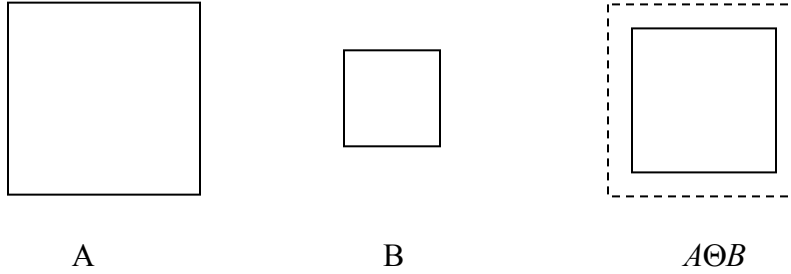


Figure 3.5: Example of Morphological Erosion

One of the simplest applications of erosion is eliminating irrelevant detail in an image. Two important and most commonly used morphological operators are opening and closing. Each of these operators uses sequential combinations of dilation and erosion.

Opening A by B is the erosion of A by B followed by a dilation of the result by B . Mathematically this is defined as

$$A \circ B = (A \ominus B) \oplus B \quad (3)$$

Similarly closing A by B is the dilation of A by B followed by erosion of the result by A

$$A \bullet B = (A \oplus B) \ominus B \quad (4)$$

Opening generally smoothes the contour of an object, breaks narrow isthmuses and eliminates thin protrusions. Closing also tends to smooth contours but, in contrast to opening, it generally fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour. See (Gonzalez and Woods, 2002) for more details.

3.4.2 Opening and Closing by Reconstruction

The opening by reconstruction γ^* of an image I with structuring element $SE=B$ is defined as the erosion of I with $SE=B$ followed by grayscale geodesic dilation with $SE=BI$ iterated until stability is reached. The opening by reconstruction can be calculated as follows:

$$\begin{aligned} J(0) &= I \ominus B \\ n &= 0 \\ \textbf{Repeat} \\ J(n+1) &= (J(n) \oplus BI) \wedge I \\ n &= n+1 \\ \textbf{Until } J(n+1) &= J(n) \\ \gamma^* &= J(n+1) \end{aligned}$$

Where $(J(n) \oplus BI) \wedge I$ is dilation of $J(n)$ by BI i.e. structuring element with radius 1 and \wedge is the point wise minimum. The effect of opening by reconstruction is to remove all structures in an image that are both smaller than SE and brighter than the surroundings.

The closing by reconstruction ϕ^* of an image I with structuring element $SE=B$ is defined as the dilation of I with $SE=B$ followed by grayscale geodesic erosion with $SE=BI$ iterated until stability is reached. The closing by reconstruction can be calculated as follows:

$$J(0) = I \oplus B$$

$$n = 0$$

Repeat

$$J(n+1) = (J(n) \ominus B) \vee I$$

$$n = n + 1$$

Until $J(n+1) = J(n)$

$$\Phi^* = J(n+1)$$

Where $(J(n) \ominus B) \vee I$ is erosion of $J(n)$ by B i.e. structuring element with radius 1 and \vee is the point wise maximum. The effect of closing by reconstruction is to remove all structures in an image that are both smaller than the SE and darker than the surroundings. For more detailed description see Pesaresi and Benediktsson [2001] and Vincent (1993).

In our application, we just need the bright structures (i.e. large buildings have higher elevation than surroundings and are bright in a raster black and white image) so we only utilize opening by reconstruction. Closing by reconstruction can be used in the case of dark structures, i.e. structures with relatively less height above ground level like roads, etc.

3.4.3 Differential Morphological Profile

Let γ^* be a morphological opening by reconstruction operator with $SE = \lambda$. The opening profile at pixel x of image I is defined as a vector

$$\Pi_\gamma(x) = \{\Pi_{\gamma_\lambda} : \Pi_{\gamma_\lambda} = \gamma^*_\lambda(x), \forall \lambda \in [o, n]\}. \quad (5)$$

Let Φ^*_λ be a morphological closing by reconstruction with operator $SE = \lambda$. Then, the closing profile at pixel x of the image I is defined as a vector

$$\Pi_{\varphi}(x) = \{\Pi_{\varphi_{\lambda}} : \Pi_{\varphi_{\lambda}} = \varphi_{\lambda}^*(x), \forall \lambda \in [0, n]\}. \quad (6)$$

In the above two equations, $\Pi_{\gamma}(0) = \Pi_{\varphi}(0) = I(0)$ for $\lambda = 0$ by the definition of opening and closing by reconstruction. The derivative of the morphological profile is defined as a vector storing the slope of the opening/closing profile for each step of increasing SE size. The derivative of the opening profile is defined as the vector

$$\Delta\gamma(x) = \{\Delta\gamma_{\lambda} : \Delta\gamma_{\lambda} = |\Pi\gamma_{\lambda} - \Pi\gamma_{\lambda-1}|, \forall \lambda \in [1, n]\}. \quad (7)$$

By duality, the derivative of the closing profile is defined as the vector

$$\Delta\varphi(x) = \{\Delta\varphi_{\lambda} : \Delta\varphi_{\lambda} = |\Pi\varphi_{\lambda} - \Pi\varphi_{\lambda-1}|, \forall \lambda \in [1, n]\}. \quad (8)$$

In general, the derivative of the morphological profile (DMP) can be written as the vector

$$\Delta(x) = \left\{ \Delta_c : \begin{cases} \Delta_c = \Delta_{\varphi_{\lambda=n-c+1}}, \forall c \in [1, n] \\ \Delta_c = \Delta_{\varphi_{\lambda=c-n}}, \forall c \in [n+1, 2n] \end{cases} \right\}. \quad (9)$$

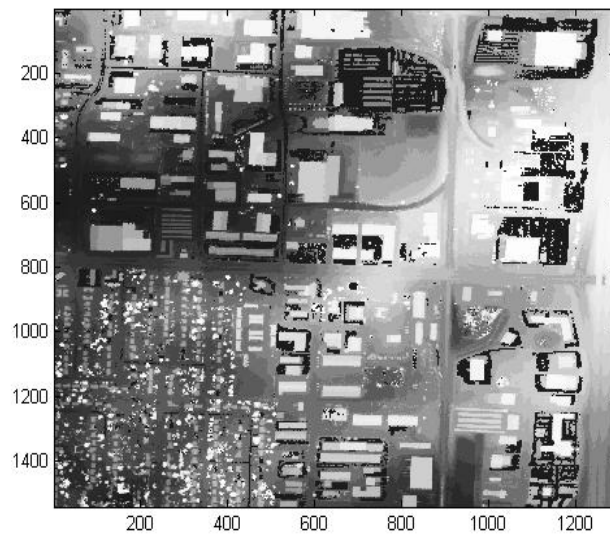
Where n is the total number of iterations (SE sizes), and $c = 1 \dots 2n$. The responses of the derivative calculated using small SEs are near the central position of the DMP vector, while the responses for the largest SEs in the closing and opening profile are recorded at the beginning ($c=1$) and the end ($c=2n$), respectively.

The signal recorded in the DMP provides information about the size and contrast of multi-scale structures in the image. Small structures will have strong response for small SE sizes, yielding a large response near the center of the DMP, while large structures will have strong response for large SE sizes, yielding a large response near the beginning or end of the DMP. Also, brighter structures will have high response in the opening portion of the DMP ($c = n+1 \dots 2n$). Since we are concerned with only commercial buildings and other large structures, structuring elements with radii starting

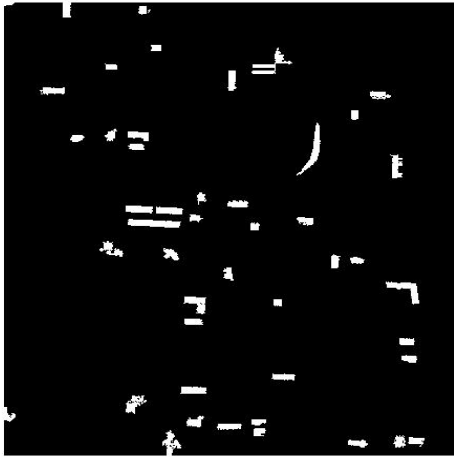
from 10 m through 65 m are used with a variable step size, i.e. the step size increases as the SE size moves towards the larger values (i.e. 65 m).

3.5 Results and Discussion

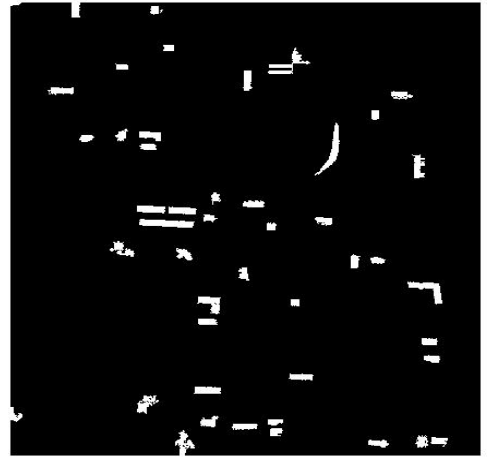
As discussed in the theory section, each level in the morphological profile corresponds to the objects of different shapes and sizes. As the radius of the structuring element increases, the object size captured in each level of the DMP also increases. Small structures have strong response for small SE sizes, yielding a large response near the center of the DMP, while large structures have strong response for large SE sizes, yielding a large response near the beginning or end of the DMP. We can also observe some non-building structures in the DMP levels because of their high contrast in the original image. So the next step of the algorithm is essentially to eliminate these non-building structures while keeping the building structures intact. Figure 3.4 demonstrates the various DMP levels for one of the lidar DEMs.



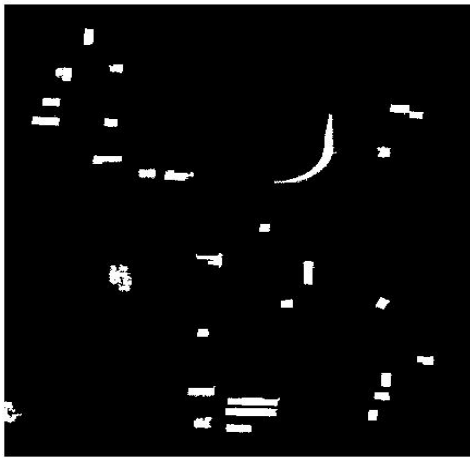
(a)



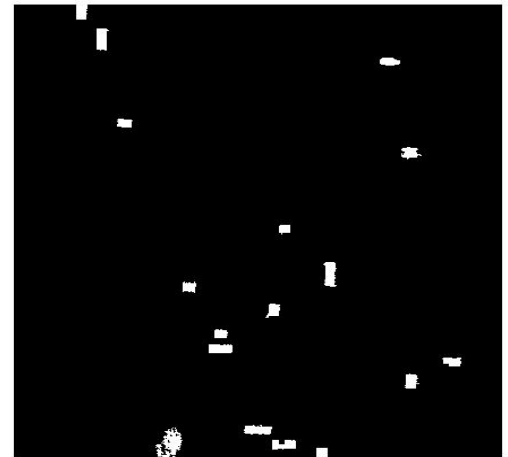
(b)



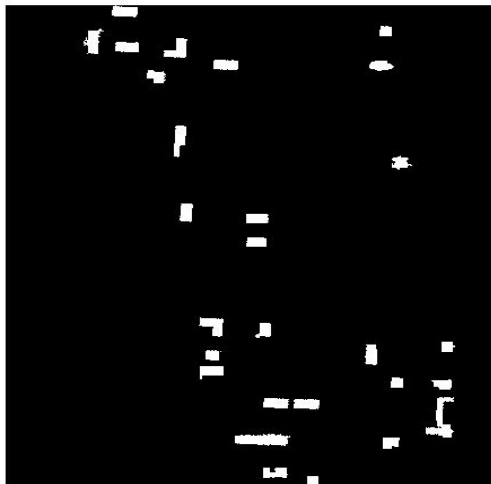
(c)



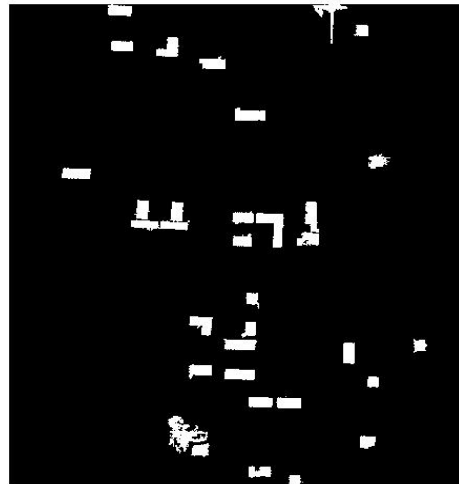
(d)



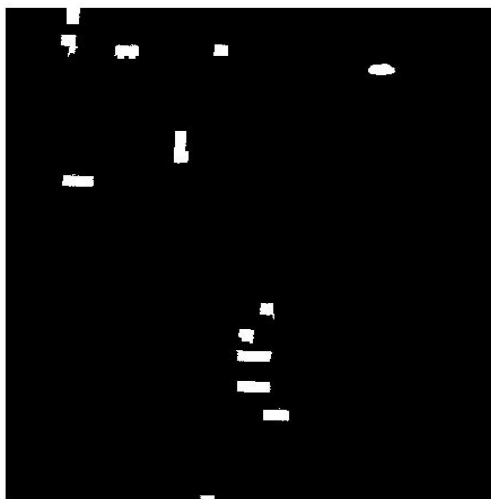
(e)



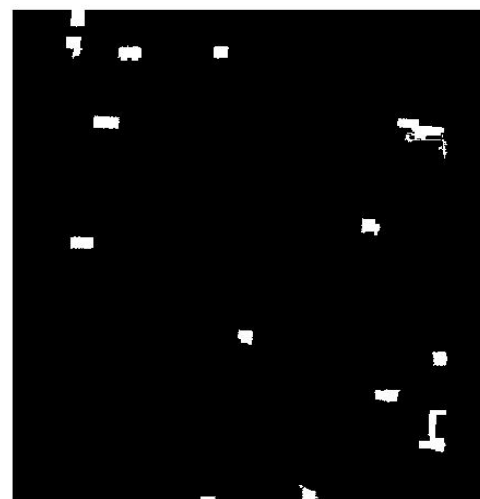
(f)



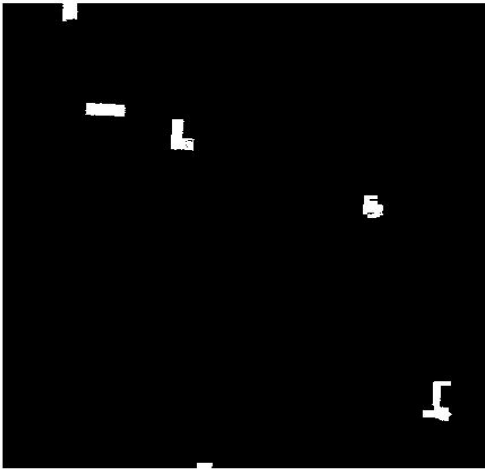
(g)



(h)



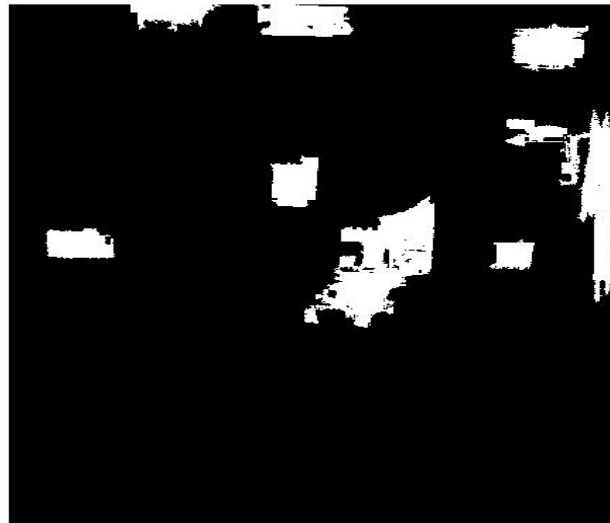
(i)



(j)



(k)



(l)

Figure 3.6: Original Image (a) & Differential Morphological Profile Levels (b)-(l) corresponding to disks of radii 12 m, 13 m, 14 m, 15 m, 16 m, 17 m, 18 m, 19 m, 20 m, 36 m, 65 m respectively.

3.6 Summary

This chapter's primary focus was to present an overview of the processing chain and then the Differential Morphological Profile, which is the backbone of this research. The chapter starts with a overview of the processing chain used in this research along with a brief discussion. Then, the next section discusses the image interpolation strategy needed to remove the "No Data Value" pixels and the median filtering step which is essential for removing random noise present in the image.

After this, a brief introduction of morphological erosion, dilation, opening, and closing is presented. This is followed by a detailed discussion of opening and closing by reconstruction and the Differential Morphological Profile (DMP). The DMP is a multi-scale image analysis technique where a morphological profile of the image is constructed through the use of morphological opening and closing by reconstruction operations while varying the size of the structuring element (SE).

The signal recorded in the DMP provides information about the size and contrast of multi-scale structures in the image. Small structures will have strong response for small SE sizes, yielding a large response near the center of the DMP, while large structures will have strong response for large SE sizes, yielding a large response near the beginning or end of the DMP. Also, brighter structures will have high response in the opening portion of the DMP ($c = n+1 \dots 2n$). Since we are concerned with only commercial buildings and other large structures, structuring elements with radii starting from 10 m through 65 m are used with a variable step size, i.e. the step size increases as the SE size moves towards the larger values (i.e. 65 m).

These DMPs have non-building structures along with the building structures. So the next step is to eliminate these non-building structures using different shape measures and combine their outputs to form a final output. This is explained in detail in the next chapter. The chapter concludes with the actual images showing the DMP levels of an image for different disk sizes.

Chapter 4

Feature Extraction and Fusion

4.1 Introduction

Once the different DMP levels are obtained, a variety of shape measures are applied to extract different feature parameters like rectangularness, elongation, branchiness, compactness, etc. Since each DMP level is composed of objects at one specific scale, the above mentioned shape features vary between DMP levels (different scales). So each DMP level is evaluated for the three shape features, i.e. rectangularness, elongation, and minimum area rectangle. Then, three final feature images are obtained from the 11 different DMP levels, one for each shape measure. Then from these three feature images, a final fused image is obtained that represents the final building extraction output. Various statistical measures are then used to quantify the accuracy of the building extraction results.

4.2 Fuzzy S and II Membership Functions

Classification and feature extraction techniques normally use crisp classification, i.e. every pixel or object is assigned to only one class. However, remote sensing images, especially for urban environments, often contain mixed pixels i.e., smaller sub-pixels with different signatures that together form a larger mixed pixel. Fuzzy classification techniques allow pixels to have more than one class and therefore better represent the imprecise nature of the data Wang [1990].

Fuzzy classification techniques are based upon the fuzzy set theory. For the basic concept of the fuzzy set theory please see Klir and Yuan (1995). This section describes the fuzzy S - and Π -membership functions which are the primary functions used in this research.

4.2.1 Fuzzy S – Membership Function

According to the definition of fuzzy set theory, membership values must fall within the range of 0 to 1. As the name suggests, the S -membership function maps the original values along an S – curved function. This function has three control points a , b and c as shown in Fig. 4.1.

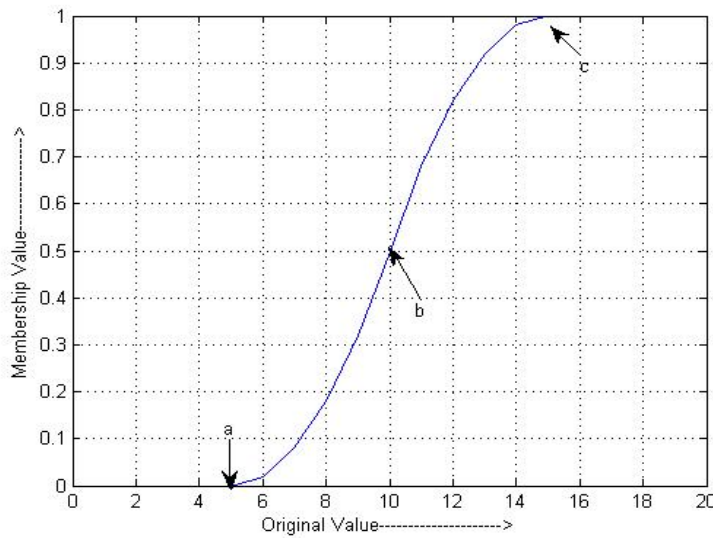


Figure 4.1: A fuzzy S -membership function

All points below a have a zero membership value. Points above a and below b have membership values between 0 and 0.5. Similarly, points above b and below c have membership values between 0.5 and 1.0. Finally, points above c all have a membership value of 1. The equations controlling these values are as follows:

$$S(x ;, a, b, c) = 0 \quad \text{for } x \leq a \quad (1)$$

$$S(x ;, a, b, c) = 0.5 \times \left(\frac{x - a}{b - a} \right)^2 \quad \text{for } a < x \leq b \quad (2)$$

$$S(x ;, a, b, c) = 1 - 0.5 \times \left(\frac{x - c}{c - b} \right)^2 \quad \text{for } b < x \leq c \quad (3)$$

$$S(x ;, a, b, c) = 1 \quad \text{Otherwise} \quad (4)$$

4.2.2 Fuzzy II–Membership Function

A Π – membership function consists of 2 S -functions as shown in Fig. 4.2.

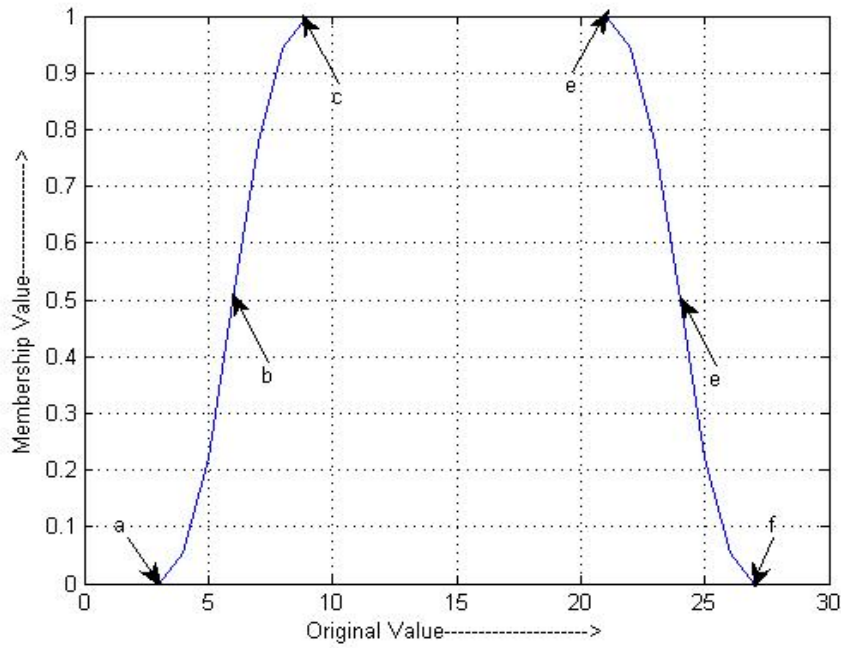


Figure 4.2: A fuzzy Π – membership function

The equations governing this function are:

$$S(x ;, a, b, c, d, e, f) = 0 \quad \text{for } x \leq a \text{ or } x \geq f \quad (5)$$

$$S(x ;, a, b, c, d, e, f) = S(x; a, b, c) \quad \text{for } a < x < c \quad (6)$$

$$S(x ;, a, b, c, d, e, f) = 1 \quad \text{for } c \leq x \leq d \quad (7)$$

$$S(x ;, a, b, c, d, e, f) = 1 - S(x; d, e, f) \quad \text{for } d < x < f \quad (8)$$

Thus, each point is assigned a value depending upon its proximity to the threshold points (i.e. a, b, c, d, e , or f).

4.3 Minimum Bounding Rectangles

Minimum Bounding Rectangles (MBRs) are also called minimum enclosing rectangles or minimum area encasing rectangles. These are defined as the rectangular structure that completely encloses a given closed curve ensuring a minimum area ratio (i.e. ratio of area of the closed curve to the area of the enclosing rectangle).

From the DMP objects there are sometimes closed curved objects along with the building structures caused by vegetation areas, noise, or other non-building structures. These structures need to be eliminated and one of the most efficient ways is to use MBRs. Fig. 4.3 illustrates the MBR concept.

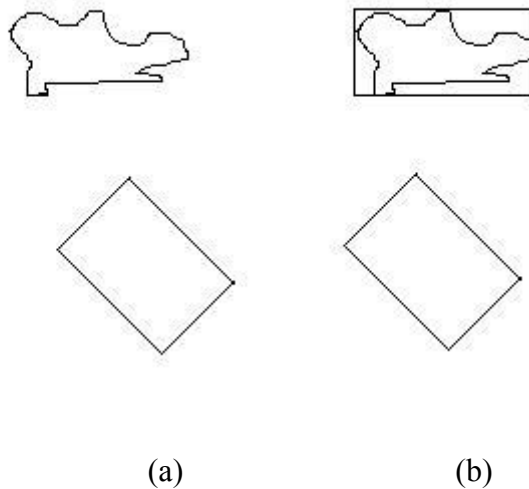


Figure 4.3: Minimum bounding rectangles for some closed curves. Original curves are shown in (a) and their corresponding MBRs are shown in (b)

In the above figure, the first curve has an area ratio (closed curve area divided by MBR area) less than that for the second one because the second curve itself is a rectangle so the MBR is the same as the curve (area ratio of 1). From the shape of the two curves

we can see that second object is more likely to be a building structure than the first one. Therefore, we can eliminate the first object (irregular object) by setting an appropriate MBR threshold to reject it. When calculating the minimum bounding rectangle we must consider the following three theorems:

- Given a rectangle with four points arbitrarily chosen such that no edge contains more than one point, there exists another rectangle such that each edge contains one and only one of these points and the area of the rectangle is less than the given rectangle.
- The rectangle of minimum area enclosing a convex polygon has a side collinear with one of the edges of the polygon.
- The minimum area rectangle encasing the convex hull of a simple, closed, chain coded curve is one and the same as the minimum area rectangle enclosing the curve.

For more details regarding the theory and the algorithm please refer to Freeman and Shapira (1975) and Gonzalez and Woods (2002). Thus for every object in each DMP level, a minimum enclosing rectangle is found. Depending upon the mean value of the area ratios of all objects in each DMP level, a fuzzy value is assigned to each object in that particular DMP level using an *S*-membership function. Since each pixel has a different membership value in each DMP, a maximum value from all the DMP levels is found for each pixel to form a final image. A final MBR output is shown in Fig. 4.4.

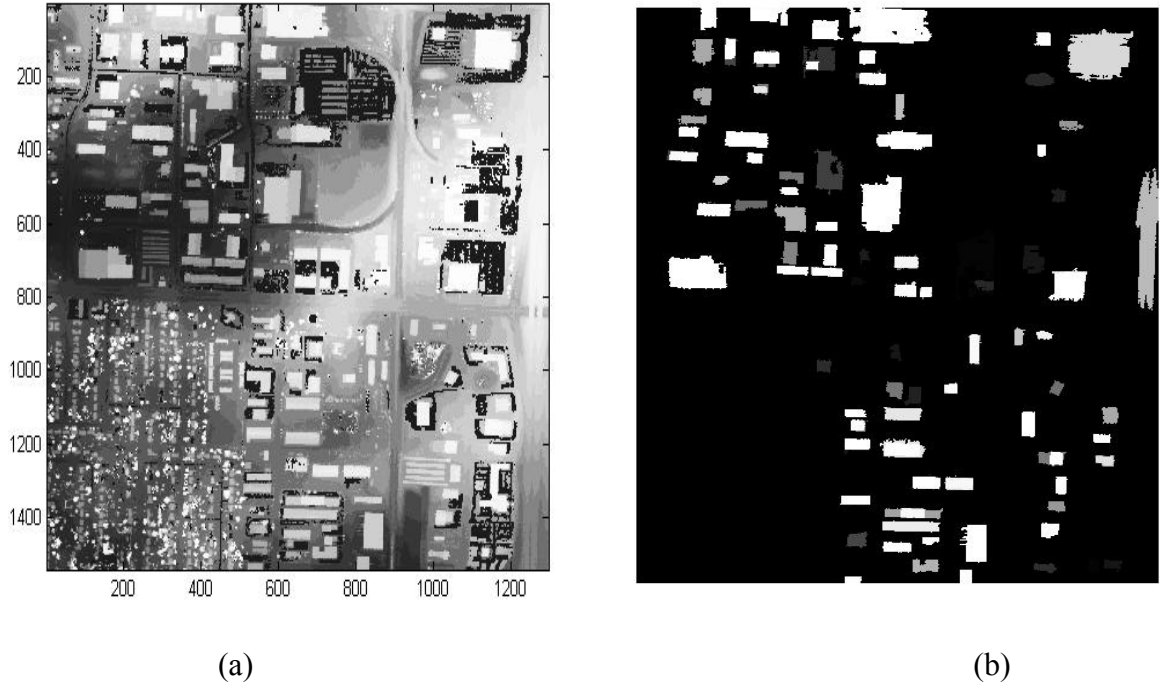


Figure 4.4: (a) original DEM image (b) output of fuzzy minimum bounding rectangle measure.

4.4 Fuzzy Rectangular Measure

Shape information for each DMP object can also be measured using the morphological skeleton of the object. Morphological skeletonization is a process of reducing a binary shape into a graph that largely preserves the extent and connectivity of the shape while discarding the foreground pixels. There are a variety of methods to produce a skeleton. The method used here is part of the image processing toolbox in the Matlab environment¹ and it always produces a connected and one pixel thick skeleton.

Since the skeleton is very sensitive to small-scale variations in the object's original shape, the original shape should be initially smoothed by performing a

¹ Bwmorph (I, 'skel' , inf): I is the input image matrix

morphological closing operation with suitable structuring element. The morphological filtering operation ensures that the smoothening will preserve the shape.

Once the skeleton is found for each object in a DMP level, the endpoints of the skeleton are identified using a binary hit and miss transform. For more details see Gonzalez and Woods (2002). Line segments are then found to connect the endpoints and then the angle between the two line segments connected to each endpoint is calculated. This process is illustrated in the Figure 4.5.

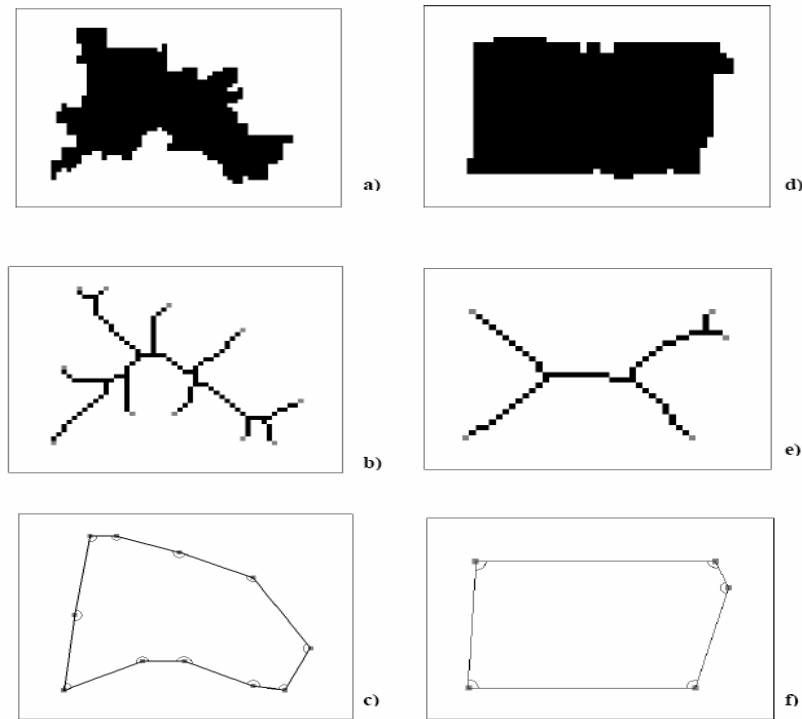


Figure 4.5²: Morphological skeleton processing for image segments shown in (a) and (d). Result of skeletonizing algorithm shown in (b) and (e) for image segments (a) and (d), respectively. The endpoints of the skeletons have been highlighted in gray. In (c) and (f), the skeleton endpoints have been connected by line segments and the endpoint angles identified.

² From Shackelford and Davis (2002).

The angles and the line segment lengths are used as shape information to help classify the building as “approximately rectangular”. A shape with about four endpoints with angles close to 90^0 and large separation, about two or less endpoints with angles much larger than 90^0 and about two or less endpoints with angles much smaller than 90^0 is classified as an “approximately rectangular”. Since these attributes are imprecise and the shapes of buildings vary, fuzzy membership functions are used to measure how closely the endpoint angles and the line segment lengths match the different criteria.

First, three fuzzy membership values are calculated for each skeleton endpoint angle θ_i : about 90^0 , $\mu_{\sim 90}(\theta_i)$, larger than 90^0 , $\mu_{>90}(\theta_i)$, and smaller than 90^0 , $\mu_{<90}(\theta_i)$. First, $\mu_{\sim 90}(\theta_i)$ is calculated using a π -function as

$$\mu_{\sim 90}(\theta_i) = \Pi(\theta_i; 70, 80, 85, 95, 100, 110) \quad (9)$$

Where $\Pi(\theta_i; 70, 80, 85, 95, 100, 110)$ is a Π -membership function as explained earlier.

$\mu_{>90}(\theta_i)$ and $\mu_{<90}(\theta_i)$ can be calculated using S -functions as

$$\mu_{>90}(\theta_i) = S(\theta_i; 95, 100, 110) \quad (10)$$

$$\mu_{<90}(\theta_i) = 1 - S(\theta_i; 70, 80, 85) \quad (11)$$

The reference angle values are chosen so that angles matching the description of the membership functions would have membership values close to one, while angles deviating from the descriptions would have membership values approaching zero. Similarly a membership value to quantify the large separation between the skeleton endpoints, $\mu_{LS}(e_i)$ is calculated for each line segment connecting the segment skeleton endpoints as

$$\mu_{LS}(e_i) = S(\min[d_i, d_{i+1}]; 3, 5, 7) \quad (12)$$

where e_i is the i^{th} endpoint and d_i is the distance between e_i and e_{i-1} . “Large separation” is defined as being sufficiently larger than 5 m. To identify endpoints that have both angles near 90° and large separation ($\mu_{\sim 90, \text{LS}}(e_i)$), the fuzzy intersection (i.e. the minimum value) of (1) and (4) is calculated.

Fuzzy quantifiers (Klir, 1996) are used to calculate membership values for fuzzy sets describing ‘about four’ or ‘about two or less’. The membership value for a segment having about four endpoints with angles close to 90° and large separation, μ_4 , is calculated as

$$\mu_4 = \Pi \left(\sum_{i=1}^n \mu_{\sim 90, \text{LS}}(e_i); 1.0, 1.5, 2.5, 5.5, 6.5, 7.0 \right) \quad (13)$$

The membership values for a segment having about two or less angles much larger than 90° , $\mu_{2>}$ and having about two or less angles much smaller than 90° , $\mu_{2<}$ are calculated as:

$$\mu_{2>} = 1 - S \left(\sum_{i=1}^n \mu_{>90}(\theta_i); 2.0, 3.0, 4.5 \right) \quad (14)$$

$$\mu_{2<} = 1 - S \left(\sum_{i=1}^n \mu_{<90}(\theta_i); 2.0, 3.0, 4.5 \right) \quad (15)$$

Finally, the membership value for a shape being “approximately rectangular” μ_R is calculated using fuzzy intersection of (13), (14) and (15) with the minimum operator.

As with the fuzzy MBR, this procedure is repeated for each DMP level and then a final image is obtained by taking a pixel wise maximum from all DMP levels. Fig. 4.6 demonstrates the output of the fuzzy minimum bounding rectangle measure on one of the images.

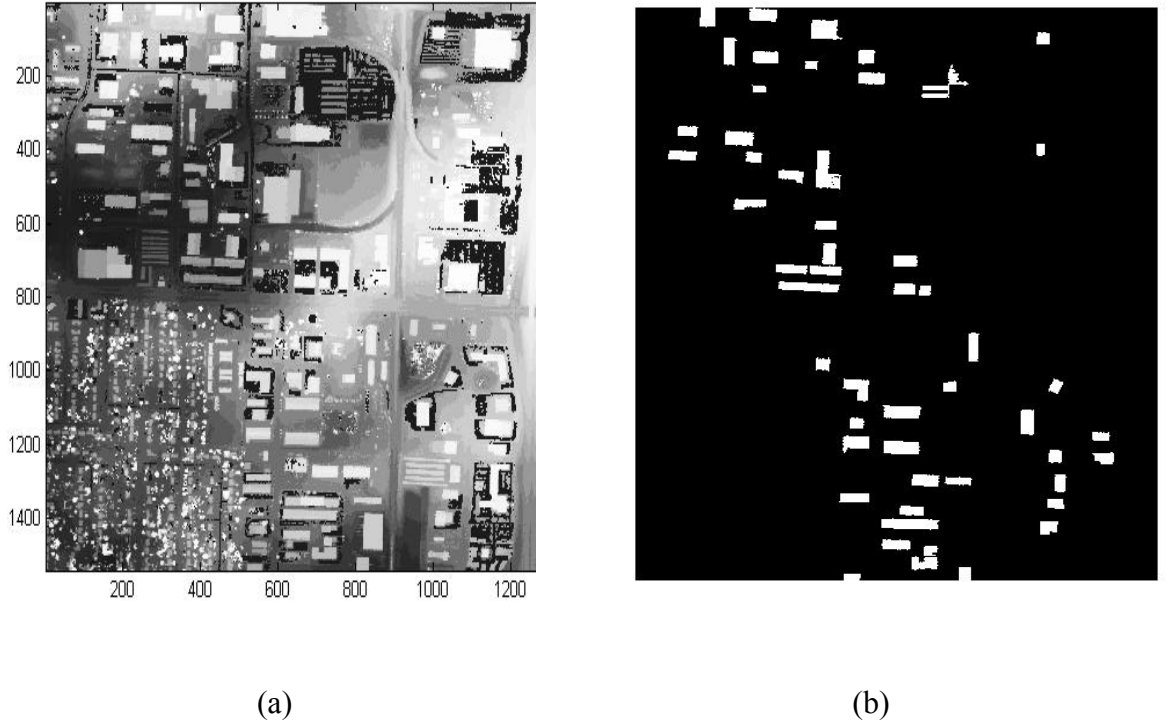


Figure 4.6: (a) original DEM (b) output of fuzzy rectangular shape measure.

From this figure we see that almost all the shapes delineated are rectangular or L-shaped. Hence when we fuse all three detector outputs into one final image, a confidence value of 1 is assigned to this measure. This is explained in more detail in the coming sections.

4.5 Fuzzy Branchiness Measure

One efficient way to eliminate non-building structures is to measure the “area porosity” or the “branchiness” (Soh, Tsatsoulis, Gineris and Bertoia, 2002) of the structure. This is given as

$$area_porosity = \frac{\max_length \times \max_width}{area}, \quad (16)$$

where *area* is area of the object under consideration, *max_length* and *max_width* are the lengths of the object's major axis and minor axis respectively. These can be calculated as follows:

- First we need to determine the orientation of the object using

$$\text{Orientation} = \frac{1}{2} \times \tan^{-1} \left[\frac{2\bar{\mu}_{1,1}}{\bar{\mu}_{2,0} - \bar{\mu}_{0,2}} \right] \quad (17)$$

Where,

$$\mu_{0,2} = \sum_{ally} (y - \mu_y)^2, \mu_{2,0} = \sum_{allx} (x - \mu_x)^2, \mu_{0,2} = \sum_{allx} \sum_{ally} (x - \mu_x)(y - \mu_y) \quad (18)$$

x and y are the lists of all the x -coordinates and y -coordinates of the all the points in that object and μ_x and μ_y are their means.

- Then for each pixel $\langle x, y \rangle$ find

$$\alpha = x * \cos(\text{orientation}) + y * \sin(\text{orientation}) \quad (19)$$

$$\beta = -x * \sin(\text{orientation}) + y * \cos(\text{orientation}) \quad (20)$$

- Identify maximum and minimum values of both α and β and call them α_{\min} , α_{\max} and β_{\min} and β_{\max} .
- Then the principle axes can be given as

$$d_\alpha = \alpha_{\max} - \alpha_{\min} \text{ and } d_\beta = \beta_{\max} - \beta_{\min} \quad (21)$$

$$\text{max_length} = \max(d_\alpha, d_\beta) \text{ and } \text{max_width} = \min(d_\alpha, d_\beta)$$

It is observed that usually non-building structures will have area porosity values that are higher than for building structures. Normally buildings tend to be more rectangular and smoother than their non-building counterparts. If the structure has a branchy nature then the numerator of the area porosity ratio becomes large as compared to the area of the object.

As with the first two measures, every object in each DMP level is assigned a fuzzy value depending on the mean value of the branchy measures of all the objects. This procedure is then repeated for all the DMP levels and a final image is obtained by taking the pixel-wise maximum. At the end of this step we have three detector outputs; one from the fuzzy minimum bounding rectangle measure, one from the fuzzy rectangular shape measure, and one from the fuzzy branchiness measure.

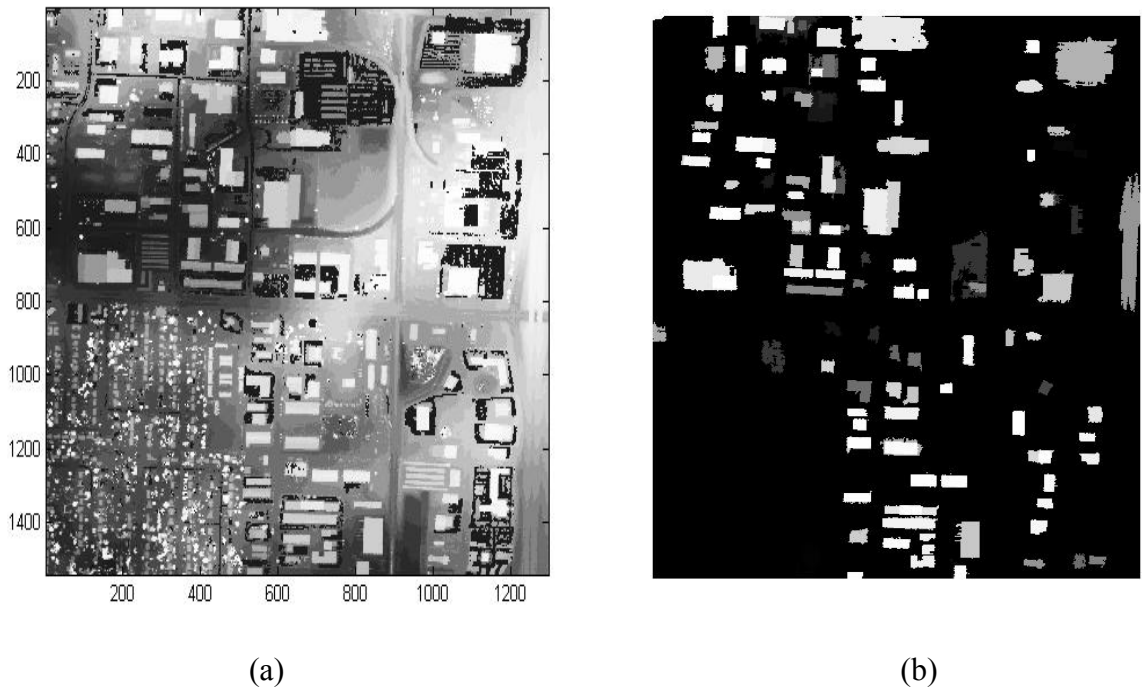


Figure 4.7: (a) Original DEM (b) Output of fuzzy branchiness measure.

4.6 Multi-Detector Fusion

We have three different detector outputs that must be combined into one final output image. However, we must consider the fact that each detector output has some incorrectly identified objects, e.g. non-building structures. Conversely, each detector output may have some building structures that are not identified by the other two

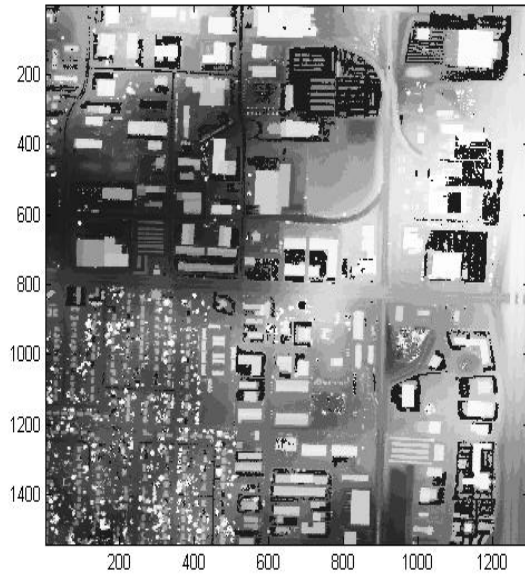
detectors. Thus, our goal is to fuse these three detector outputs in such a way that minimizes the number of non-building structures while maximizing the correct identified building structures.

One approach to this problem is to assign confidence values (on the scale of 0 to 1) to each of the detector outputs and set a fuzzy rule base for their fusion. Here we assign a maximum confidence (i.e. 1) to the output from the fuzzy rectangular shape measure because it actually considers the angles and the lengths of the sides of each object, and, accordingly, most of the structures identified by this measure are perfectly rectangular or L-shaped structures. The rules used for the fusion are as follows:

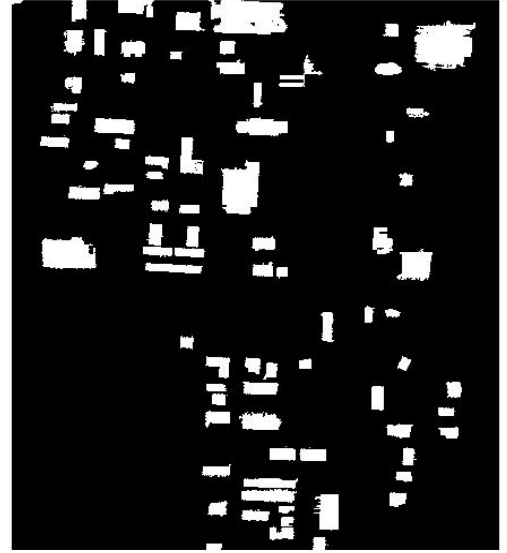
Given pixel is an output pixel if

- *It is in the fuzzy rectangular measured image **OR***
- *If the confidence value of that pixel in the fuzzy minimum bounding rectangle measure is above 0.75 **OR** in fuzzy branchiness measure is 0.7.*
- ***OR IF** the confidence value of that pixel in the fuzzy minimum bounding rectangle measure is above 0.6 **AND** in fuzzy branchiness measure is 0.6; **ELSE** pixel is not an output pixel.*

The threshold values used above were determined from empirical evaluations on two tuning images and were shown to produce good results for these tuning images.



(a)



(b)

Figure 4.8: Final output after combining detector outputs from 4.3, 4.5, and 4.6 using the fusion method discussed in the text.

4.7 Accuracy Assessment

Accuracy assessment of the final algorithm output can be done in two ways. The first way is to compare the final output image and a reference image on a pixel by pixel basis. The second approach is to compare these on an object or building basis. The first method represents the best case while the second method is a worst- case assessment. The following subsections explain both methods.

4.7.1 Pixel-Based Assessment

In this method every pixel in the output image is either marked as True Positive, True Negative, False Positive, or False Negative (Jin and Davis,2004) using the following definitions.

- **True Positive (TP):** Both manual and automated methods label the pixel belonging to the same building.
- **True Negative (TN):** Both manual and automated methods label the pixel belonging to the background.
- **False Positive (FP):** The automated method incorrectly labels the pixel as belonging to the building
- **False Negative (FN):** The automated method incorrectly labels the pixel as belonging to the background.

Once we determine the number of pixels in each category we can evaluate the algorithm performance using the following statistical measures:

- **Branching factor** = $\frac{FP}{TP}$ (22)

- **Miss factor** = $\frac{FN}{TP}$ (23)

- **Completeness percentage** = $100 * \frac{TP}{TP_R}$ (24)

- **Correctness percentage** = $100 * \frac{TP}{TP + FP}$ (25)

- **Quality percentage** = $100 * \frac{TP}{TP + FP + FN}$, (26)

where TP_R represents the number of building pixels (TP+FN) in a reference image.

The **detection percentage** denotes the percentage of building pixels correctly labeled by the automated process. The **branching factor** is a measure of the commission error of the system that incorrectly labels the background pixels as building. The **miss factor** measures the omission error of the system that incorrectly labels valid building pixels as the background. The **miss factor** can be derived from the **detection percentage**. Among these statistics, the **quality percentage** measures the absolute quality of the extraction and is the most stringent measure. To obtain 100% quality, the extraction algorithm must correctly label every object pixel ($FN = 0$). The accuracy of any algorithm will be lower when assessed at the pixel level. Low completeness values for the pixel-level accuracy assessment are caused by discrepancies between the shape of the buildings extracted by the algorithm and the reference building footprints.

4.7.2 Object-Based Assessment

The object-based assessment is very similar to the pixel-based assessment. It provides a more generalized idea about the accuracy of the algorithm compared to the pixel-based counterpart since it compares the number of objects in the output image rather than the pixels. Here again the buildings in the reference image are manually labeled (TP_R) and, similar to the previous method, objects are marked as TP , FP or FN . Here usually **completeness**, **correctness**, **detection percentage**, and **quality** measures are calculated using the same formulae provided by the pixel-based accuracy assessment section.

4.8 Summary

In Chapter 3 we discussed the Digital Morphological Profile in detail. This chapter described the application of a variety of shape measures followed by a multi-detector fusion technique to generate the final output image with 2D building footprints (polygons).

Classification and feature extraction techniques normally use crisp classification, i.e. every pixel or object is assigned to only one class. Here we adopted fuzzy classification techniques that allow the image pixels to have multiple shape measures to better represent the imprecise nature of the image objects. Thus, this chapter first provided a brief introduction for fuzzy S and Π membership functions which are primarily used in this research.

Next, Minimum Bounding Rectangles (MBRs) are explained in detail. These are defined as the rectangular structure that completely encloses a given closed curve ensuring a minimum area ratio (i.e. ratio of area of the closed curve to the area of the enclosing rectangle). Since buildings are usually rectangular and L-shaped, a high area ratio indicates a higher probability that a given object is a building. Depending upon the mean value of the area ratios of all objects in each DMP level, a fuzzy value is assigned to each object in that particular DMP using an S -membership function. Since each pixel has a different membership value in each DMP, a maximum value from all the DMP levels is found for each pixel to form a final image.

After this, a shape measure is used in this research to quantify the “rectangularness” of a building. First the skeleton of each object is determined. Morphological skeletonization is

a process of reducing a binary shape into a graph that largely preserves the extent and connectivity of the shape while discarding the foreground pixels. Once the skeleton is found for each object in a DMP level, the endpoints of the skeleton are identified. Line segments are then found to connect the endpoints and then the angle between the two line segments connected to each endpoint is calculated. A shape with about four endpoints with angles close to 90^0 and large separation, about two or less endpoints with angles much larger than 90^0 and about two or less endpoints with angles much smaller than 90^0 is classified as an “approximately rectangular”. Since these attributes are imprecise and the shapes of buildings vary, fuzzy membership functions are used to measure how closely the endpoint angles and the line segment lengths match the different criteria. As with the fuzzy MBR, this procedure is repeated for each DMP level and then a final image is obtained by taking a pixel wise maximum of each DMP.

The third and final measure used here is area porosity or “branchiness”. It is observed that usually non-building structures will have area porosity values that are higher than building structures because normally buildings tend to be more rectangular and smoother than their non-building counterparts. As with the first two measures, every object in each DMP level is assigned a fuzzy value depending on the mean value of the branchy measures of all the objects. This procedure is then repeated for all the DMP levels and a final image is obtained by taking the pixel-wise maximum.

At the end of these steps we have three detector outputs; one from the fuzzy minimum bounding rectangle measure, one from the fuzzy rectangular shape measure, and one from the fuzzy branchiness measure. Confidence values (on the scale of 0 to 1)

are then assigned to each of the detector outputs and a fuzzy rule base is used for their fusion.

Finally the output image is assessed on a pixel-level and object-level basis. Initially each pixel is marked either as true positive or true negative or false positive or false negative depending on whether it belongs to building or background. Once we determine the number of pixels in each category we can evaluate the algorithm performance using some statistical measures like detection, completeness, correctness and quality percentages. The object-based assessment is very similar to the pixel-based assessment. It provides a more generalized idea about the accuracy of the algorithm compared to the pixel-based counterpart since it compares the number of objects in the output image rather than the pixels. Here again the buildings in the reference image are manually labeled (TP_R) and, similar to the previous method, objects are marked as TP , FP or FN . The statistical measures remain the same as for the pixel-based assessment.

In the next chapter, these statistical measures are presented for five different images along with their detailed discussion. Two of the images are the tuning images or the images used to design the algorithm while the rest three are validation images on which the algorithm is run without making any changes.

Chapter 5

Results and Discussion

5.1 Introduction

In Chapters 3 and 4 the building extraction processing strategy developed in this research was presented. This chapter presents the actual output of the processing strategy for five high resolution LIDAR DEM test images along with the accuracy statistics and a brief discussion.

5.2 Results and Discussion

The algorithms designed in this research were developed and tested on two images to obtain the best possible pixel- and object-based accuracy statistics. Once we finalized the technique, we then tested it on three other images without any further modification to demonstrate the repeatability of its performance.

In this section we will show actual polygonal building footprint outputs for two images. The first image was used to develop and test the algorithms, and the second image was used to validate the results on a independent image. The statistics for the other three images are presented in the later section. All the test images used for this research were obtained from Springfield, Missouri and have a resolution of 1 m.

After this, the accuracy statistics are reported and this is followed by a brief discussion about the algorithm's performance. The chapter is then concluded with a brief summary and discussion.

Development/Test Image 1 (size 1540*1295):

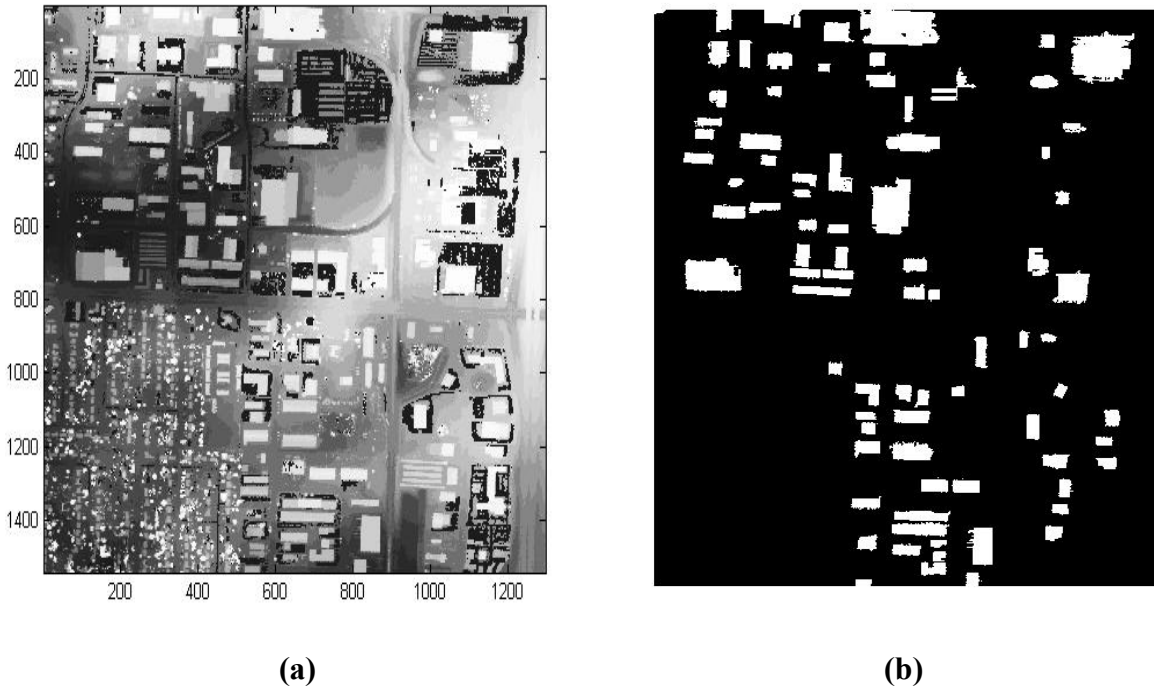


Figure 5.1: (a) Original DEM image and (b) Final building output.

**Table 5.1: Pixel Based and Building Based Accuracy Statistics for Development
Image 1**

a. Pixel-Based Statistics:

<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>Completeness%</i>	<i>Correctness%</i>	<i>Quality%</i>
171451	1724083	62157	36609	82.4	73.39	63.45

b. Object-Based Statistics:

<i>TP</i>	<i>FP</i>	<i>FN</i>	<i>Total buildings</i>	<i>Completeness%</i>	<i>Correctness%</i>	<i>Quality%</i>
97	3	22	119	81.51	97	79.5

Discussion:

Since this image is one of the two images used to develop and test the algorithms, the accuracy statistics are better than those for the validation images. The correctness percentage is the probability that the algorithm will extract a building from all the extracted structures. A high correctness value of 97% for the object-level extraction indicates that most of the extracted building objects are at least partially overlapping with the actual buildings present in the image. However, a lower completeness value of 81.5% indicates that about one fifth of the buildings present in the image are not being identified at all. As expected, the pixel statistics are considerably less than their object counterpart. Low completeness values for the pixel-level accuracy assessment are caused by discrepancies between the shape of the buildings extracted by the algorithm and the reference building footprints. It can be observed that false negatives (*FNs*) play an important role in the accuracy assessment. False negative pixels are pixels that are incorrectly identified as background pixels. The process of multi-detector fusion eliminates some valid building structures (resulting in false negative pixels) which are detected by individual detectors.

Validation Image 1(size 1000*1000):

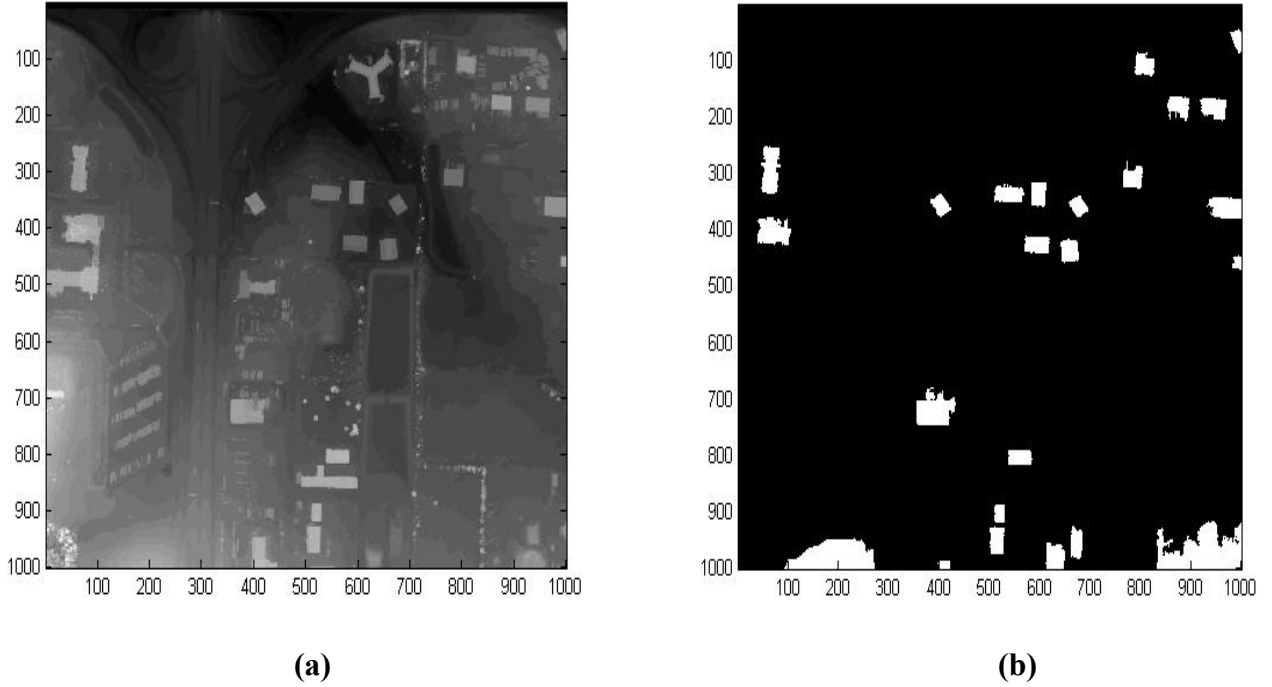


Figure 5.2: (a) Validation image and (b) Final building output.

Table 5.2: Pixel Based and Building Based Accuracy Statistics for Validation Image

1

a. Pixel Based Statistics:

TP	TN	FP	FN	Completeness%	Correctness%	Quality%
26371	938067	14882	20680	56.05	63.92	42.58

b. Object Based Statistics:

TP	FP	FN	Total buildings	Completeness%	Correctness%	Quality%
24	2	5	29	82.75	92.3	77.42

Discussion:

As mentioned earlier, this image is a validation image since the algorithm was NOT tuned for this image. Accordingly, the statistical results are not as good as they were for the first image. As we can observe in Figure 5.2 and its statistical results, almost all the buildings (24 out of 29) are correctly extracted, but two buildings (bottom left and right corners) are falsely detected which leads to a large number of false positives deteriorating the pixel statistics. Also we can observe that the algorithm has failed to detect a triangular-shaped building. This was expected because the shape measures used in this algorithm only consider rectangular and L-shaped buildings.

The following tables show accuracy statistics for three more images. Here, the first one is for a development/test image and the remaining two are validation images.

Table 5.3: Pixel Based Accuracy statistics for Development/Test Images (All Images Of size 1000*1000 pixels)

<i>Images</i>	<i>TP</i>	<i>TN</i>	<i>FP</i>	<i>FN</i>	<i>Completeness%</i>	<i>Correctness%</i>	<i>Quality%</i>
Validation Image 1	169143	1172375	49345	39137	77.41	81.2	65.65
Validation Image 2	45345	817077	134652	2926	25.19	93.94	24.78
Validation Image 3	27822	942015	17731	12432	61	69.11	47.98

Table 5.4: Building Based Accuracy statistics for Development/Test Images (All Images of size 1000*1000 pixels)

<i>Images</i>	<i>TP</i>	<i>FP</i>	<i>FN</i>	<i>Total buildings</i>	<i>Completeness%</i>	<i>Correctness%</i>	<i>Quality%</i>
Validation Image 1	62	6	22	84	73.8	91.17	68.89
Validation Image 2	16	0	9	25	64	100	64
Validation Image 3	24	5	6	30	80	82.76	68.57

Discussion:

As expected for the tuning image 2 both the statistics are convincing. Though Test image 2 has fewer false negative pixels, it is the presence of false positives in a large number which lowers the completeness and quality significantly. It can be seen from the object statistics of test image 2 that it is possible to get 100% correctness (i.e. zero FP). To reduce the false negatives and hence improve the completeness and quality more effective preprocessing can be used such as morphological filtering than just using median filtering.

5.3 Summary and Conclusion

In this chapter we presented actual algorithm output and provided both pixel-based and object-based statistics. The first image (Fig 5.1) was used along with one additional image to develop and test the processing algorithms. Accuracy statistics were calculated for each image as explained in the Chapter 4.

The correctness percentage is the probability that the algorithm will extract a building out of all the extracted structures. A high correctness value for the object-level extraction indicates that most of the extracted building objects are at least partially overlapping with the actual buildings present in the image. It is possible to obtain 100% correctness in the object level extraction which can be observed from the statistics of validation image 2. The completeness percentage indicates the number of buildings algorithm failed to extract. Since this algorithm focuses only on rectangular and L-shaped buildings, completeness percentage is adversely affected by buildings with irregular shapes (e.g. triangular). Also the process of multi-detector fusion eliminates some building structures that are detected by individual detectors. This fusion algorithm can be further improved by implementing fuzzy integral techniques resulting in better completeness and quality percentages.

Chapter 6

Summary and Future Work

6.1 Summary

The goal of this research was to develop a robust automated building footprint extraction technique from high-resolution DEM raster data generated by airborne LIDAR. The development of automated and semi-automated techniques for generation of urban geospatial information products is important not only because of the many applications for which they can be used, but also because the large volume of the data collected by the remote sensing systems exceeds the capacity of trained geospatial professionals to analyze. Many applications, especially military and intelligence related activities require near real time exploitation of geospatial data. Both the quantity of data and the speed with which that data must be analyzed illustrate the need for automated geospatial information extraction.

In Chapter 1 we introduced Geographical Information Systems (GIS) and discussed their role in different applications like urban planning, environmental management, agriculture, transportation, utilities (water, electricity, sewage etc.). Then a brief history of remote sensing imagery was presented and spatial and temporal characteristics of different satellite imaging systems were listed with a variety of background material on different high resolution geospatial image and data including aerial photography, Landsat, IKONOS and Digital Elevation Models (DEMs)etc.

In Chapter 2, a literature review for building extraction from high-resolution satellite images like IKONOS and DEM data was presented along with accuracy and quality statistics. This was followed by a brief description of two algorithms (DRAW and Rubberband) implemented before on the DEM data used in this research. Since the accuracy statistics of these two algorithms were not satisfactory and because they were very computationally complex, a completely new approach was needed.

Chapter 3 presented the overall processing strategy explaining the approach used in this research. The image preprocessing (Image Interpolation and Median filtering) and the Differential Morphological Profile (DMP), which is the backbone of this algorithm, were presented. The generation of the different DMP levels from the interpolated median filtered data was described for a predefined spatial scale range. The DMP is a multi-scale image analysis technique that provides information about both the size and contrast of multi-scale structures in the image. Once the DMP levels are generated, three different shape extraction algorithms namely Fuzzy Minimum bounding rectangles, Fuzzy Rectangularness measure and Fuzzy Branchiness measure were applied to these DMPs to get the three different detector outputs. The final building footprint output is obtained from the three detector outputs using a fusion technique based on some simple fuzzy rules. Finally, a variety of accuracy statistics were calculated to evaluate the performance of our algorithm.

Chapter 5 presented the actual accuracy statistics (pixel-based and object-based) like completeness, correctness, detection rate, miss factor, branching ratio, and quality. Completeness is the percentage of the buildings extracted compared to the total number of buildings in the reference image. Correctness represents the percentage of the correctly

extracted buildings from the total number of buildings in the data. Detection rate is the probability of correctly extracting a building. The quality percentage measures the absolute quality of the extraction and is the most stringent measure.

The results of our algorithm show that we have obtained 82% correctness, 73% completeness and 65% quality pixel wise and 82% correctness, 97% completeness and 65% quality object wise for the tuning images and similar results for the test images. The accuracy of the algorithm is significantly lower when assessed at the pixel level. The low completeness value in the pixel-level accuracy assessment is caused by discrepancies between the shape of the buildings extracted by the algorithm and the reference building footprints. The high correctness value of the object-level extraction indicates that most of the extracted building objects at least partially overlap the actual buildings present in the image. However, the lower completeness value indicates that almost one third of the buildings present in the image are not being identified at all.

6.2 Future Work

There are a number of areas that were not explored in this research due to time constraints. For example, all three detectors utilize fuzzy membership functions to model the uncertainty and ambiguity associated with the shape of the structures. In each case, the parameters governing the shape of these membership functions have been manually set. While these parameters are intuitively chosen based on characteristics of the raster data and the type of features they are to represent, it would be beneficial to develop techniques for learning these directly from the data. One possible approach would be to utilize neural network type techniques.

There are only two shape parameters - minimum bounding rectangles and area porosity or branchiness - that were considered here. However there are number of other shape or object-level parameters that could be explored including contrast, perimeter porosity, and texture measures like average roughness, elongation, roundness, etc. See Soh, Tsatsoulis, Gineris and Bertoia [10] for more details. In the image preprocessing stage, morphological opening and closing filters can be used to further image smoothening. The fuzzy integral approach used here to combine the multiple detector outputs is in its basic stage.

References

1. Aaron Sheckelford, December 2004 Doctoral Dissertation: "Development of Urban Geospatial Information Products from High Resolution Satellite Imagery Using Advanced Image Analysis Techniques", University of Missouri Columbia.
2. Aaron Sheckelford and Curt Davis, Oct. 2003, "A combined Fuzzy Pixel-Based and object -Based approach for classification of high resolution multispectral data over urban areas.", IEEE transactions on Geoscience and Remote sensing Vol. 41, No. 10.
3. H. Freeman and R. Shapira, July 1975, "Determining the minimum area encasing rectangle for an arbitrary closed curve.", Communications of the ACM, vol. 18, No. 7.
4. Martino Pearesi and Jon Atli Benediktsson, Feb. 2001, "A new approach for the morphological segmentation of high resolution satellite imagery.", IEEE transactions on Geoscience and Remote sensing, Vol. 39, No. 2.
5. Xiaoying Jin and Curt Davis, 2004, "Automated building extraction from high resolution imagery in urban areas using structural, contextual, and spectral information."
6. Haithcoat T., Song and Hipple J., 2001, "Building footprint extraction and 3-d reconstruction from LIDAR data.", IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas.
7. Luc Vincent, April 1993, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms.", IEEE Transactions on Image Processing, Vol. 2, No. 2.
8. Gonzalez and Woods, 2002, "Digital Image Processing.", 2nd edition, Prentice Hall: New Jersey.
9. Dr. Wenbo Song, May 2001, Master of Arts Thesis: "An Integrated Approach of Automatic Road Extraction and Evaluation from Remotely sensed Imagery", University of Missouri Columbia.
10. Leen-Kiat Soh, Costas Tsatsoulis, Denise Gineris, and Cheryl Bertoia, 2002, "ARKTOS: An Intelligent System for SAR Sea Ice Image Classification"
11. Lee, D. S., J. Shan and J. S. Bethel, 2003, "Class-guided building extraction from IKONOS imagery," Photogrammetric Engineering & Remote Sensing, Vol. 69, No. 2, pp. 143-150.

12. Croitoru, A. and Y. Doytscher, 2003, "Monocular right-angle building hypothesis generation in regularized urban areas by pose clustering," *Photogrammetric Engineering & Remote Sensing*, Vol. 69, No. 2, pp. 151-169.
13. Segl, K. and H. Kaufmann, 2001, "Detection of small objects from high-resolution panchromatic satellite imagery based on supervised image segmentation," *IEEE Trans. Geosci. Remote Sensing*, Vol. 39, No. 9, pp. 2080-2083.
14. Jaynes, C. O., F. Stolle and R. T. Collins, 1994, "Task driven perceptual organization for extraction of rooftop polygons," *IEEE Workshop on Applications of Computer Vision*, pp. 152-159, December 1994.
15. Irvin, R. B., and D. M. McKeown, 1989, "Methods for exploiting the relationship between buildings and their shadows in aerial imagery," *IEEE Trans. Sysm. Man Cyb.*, Vol. 19, No. 6 pp. 1564-1575.
16. Mass, H.G., and G. Vosselman, 1999, Two algorithms for extracting building Models from raw laser altimetry data, *ISPRS Journal of Photogrammetry & Remote Sensing*, 54(2-3): 153-163.
17. Weidner, U., and W. Forstner, 1995, Towards automatic building reconstruction from high resolution digital elevation models, *ISPRS Journal of Photogrammetry & Remote Sensing*, 50(4), 38-49.
18. Haala, N., and C. Brenner, 1999, Extraction of building and trees in urban environments, *ISPRS Journal of Photogrammetry & Remote Sensing*, 54(2-3): 130-137.
19. Brunn, A. and Weidner, U., 1997. Extracting buildings from digital surface models. In: *International Archives of Photogrammetry and Remote Sensing*, 32(3-4W2):27-34.
20. Wang, Z. and Schenk, T., 2000. Building extraction and reconstruction from lidar data. In: *International Archives of Photogrammetry and Remote Sensing*, 33(B3):958-964.
21. Benediktsson, J. A., M. Pesaresi, and K. Arnason, 2003, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sensing*, Vol. 41, No. 9, pp. 1940-1949.
22. Klir, G. J., and B. Yuan, 1995, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall: New Jersey.