

**UCGE Reports
Number 20345**

Department of Geomatics Engineering

**AUTOMATED URBAN FEATURES
CLASSIFICATION AND RECOGNITION FROM
COMBINED RGB/LIDAR DATA**

(URL: <http://www.geomatics.ucalgary.ca/graduatetheses>)

by

HASSAN ELSAID ELHIFNAWY EID

July 2011



UNIVERSITY OF CALGARY

AUTOMATED URBAN FEATURES CLASSIFICATION AND RECOGNITION
FROM COMBINED RGB/LIDAR DATA

by

HASSAN ELSAID ELHIFNAWY EID

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

JULY, 2011

© HASSAN ELSAID ELHIFNAWY EID 2011

Abstract

Although a Red, Green and Blue (RGB) image provides rich semantic information for different features, it is difficult to extract and separate features which share similar texture properties. The data provided by a Light Detection And Ranging (LIDAR) system contain dense spatial information for terrain and non-terrain objects, but feature extraction poses difficulties in separating different features sharing the same height information. The thesis objective is to introduce an automated urban classification technique using combined semantic and spatial information leading to the ability to extract different features efficiently.

RGB color channels are used to produce two color invariant images for vegetation and shadowy areas identification. Otsu segmentation is applied on these color invariant images to identify shadows and vegetation candidates from each other. An RGB image is transformed into two other color spaces, YC_bC_r and HSV . Luminance color channel is extracted from YC_bC_r color space, while hue and saturation color channels are extracted from HSV color space. Global thresholding is applied on these color channels individually and collectively for detecting sandy areas.

Wavelet transform is used for detecting building boundaries from LIDAR height data. Final building candidates are identified after removing vegetation areas from the resulting image of extracted buildings from LIDAR data. After successful building extraction using wavelets and vegetation, sandy and shadowy areas from an RGB, remaining features will be the roads. This new filter combination introduces a highly efficient automatic urban classification approach from combined LIDAR/RGB data.

The proposed urban classification algorithm will introduce classified libraries for several features and in order to use this output efficiently an independent search algorithm is required. An efficient texture and boundary search algorithm is introduced for automatic object recognition of buildings using both wavelet transform and Monte Carlo simulation. The classification is achieved using a minimum distance classifier. The new search algorithm is applied on objects and libraries' descriptors with the same scale and orientation as on a distorted (scale and orientation) descriptor. The technique showed efficiency on both types of data.

The investigated feature classification technique is automated and efficient and presents a suitable method for extracting all different features while overcoming most of the problems in situations of similarity existing in texture or height information accompanied by fast and reliable results.

The Monte Carlo simulation descriptor succeeded in recognizing objects regardless their shape or orientation. It is considered as an efficient descriptor for shape recognition application, so it can be used for different object types (as buildings and vehicles) with different scales and orientations.

Preface

This is an unaltered version of the author's Doctor of Philosophy thesis of the same title.

The supervisor of this work is Dr. Naser El-Sheimy and the supervisory committee are

Dr. Ayman Habib, Dr. Mohamed Elhabiby and Dr. Ismail Abdel Ghafar Ismail Farag.

Acknowledgements

I would like to express my gratitude and appreciation to my supervisor Dr. Naser El-Sheimy for his professional supervision, strong support, guidance and abundant cooperation. I was very pleased to attend to his great graduate course. His professional teaching is amazing and makes the scientific material interesting and easy for understanding. Really, I was very glade to work under his supervision.

I would like to give great thanks and extend my appreciation to Dr. Mohamed Elhabiby for his vision, proposed ideas, valuable discussion and constructive suggestions during my studies. His great effort and valuable time in revising my thesis and published papers are highly appreciated. Indeed, I had great professional skills for researches, scientific article writing and presentations during my work with Dr. Mohamed Elhabiby.

I wish to extend my appreciation to my supervising committee Dr. Ayman Habib for his professional teaching and valuable discussion. He helped me to get the LIDAR and RGB images that represent the base for my research thesis. I would like to extend my gratitude and appreciation to Gen. Dr. Ismail Abdel Ghafar for his unwavering support.

I would like to thank Ana Paula B. Kersting, graduate students in Digital Photogrammetry Research Group, for her help and discussion about the LIDAR data and RGB images for my research.

I would like to give a great thanks to my friend Ahmed Mohsen, graduate student in Plane Group, for his amazing friendship during my staying in Calgary. We came together from Egypt at the same time for PhD and we leave together at the same time to Egypt. He was an excellent guide for many issues especially for programming language. Indeed, Ahmed is not only a friend but also same like brother for me and I'm very lucky to have a friend same like Ahmed.

I wish give many thanks to my colleagues in my research group, Mobile Mapping Sensor Systems, for their wonderful friendship, support and valuable discussion: Dr. Sameh, Dr. Walid, Dr. Zainab, Dr. Crius, Xing (Bob) Zhao, Dr. Wes, Ahmed El-Ghazouly, Dr. Yigiter, Sara, Adel, Abdelrahman, Mohamed Ali, Mazen, Naif, Bassem, Siddharth, Dr. Dina and Hsiu-Wen Chang (Killy). I was pleasure to work with them in this wonderful working environment.

I would like to thank my dear friend Ahmed Shawky for his cooperation and wonderful friendship feelings and for the time we shared together, individually or with our families. We are friends for a long time, but our shared time here in Calgary makes this friendship stronger than previous, thanks Ahmed so much.

I would like to thank all the faculty and staff members of the Geomatics department, Schulich School of Engineering for providing a wonderful educational environment.

I cannot find words to express my gratitude and thanks to my wife, Amira, for her continuous support. She came with me to Calgary and give me a wonderful and suitable environment that is helping me to concentrate on my work. Amira, thanks so much for your help, time, sacrifice and your prayers for me. My great son, Ahmed, and my lovely daughter, Engy, thank you so much for the happiness that I get it from the shared time with you.

I owe my life to my great mother, Amina, thank you so much for your prayers and support. I cannot express my feeling to you. Indeed, I love you so much and I pray for your asking ALLAH to give you a good health. My mother, continue pray for me, I cannot imagine the life without your prayers. Your prayers are the secret and the key of my success. I love you and thank you for everything that I remembered or not. My brother, Ahmed, and my sisters, Samira and Doaa, thanks so much for your continuous support and encouragements.

I give a great thanks to my government and my country, Egypt, for funding and supporting my research thesis. I'm sponsored by the Egyptian Armed Forces, so I would like to extend my gratitude and appreciation my sponsor for unwavering support.

Finally, I give great thanks to who give me the life, ALLAH, thanks so much for everything, faith, health, power, brain, thinking, ...etc, I cannot count your givens. I pray to you to give me a good health and make me in continuous success. I pray to you to make my country in safe and peace and same for all worldwide countries.

Dedication

To My Beloved Family

My Mother,

My Brother and My Sisters,

My Wife, My Son and My Daughters

(Thank You So Much)

Table of Contents

Approval Page.....	i
Abstract.....	ii
Preface.....	iv
Acknowledgements.....	v
Dedication.....	viii
Table of Contents.....	ix
List of Tables.....	xi
List of Figures and Illustrations.....	xii
List of Symbols, Abbreviations and Nomenclature.....	xvii
CHAPTER ONE: INTRODUCTION.....	1
1.1 Motivations.....	1
1.2 Thesis Objectives.....	2
1.3 Thesis Outline.....	4
CHAPTER TWO: AIRBORNE RGB/LIDAR DATA.....	6
2.1 Airborne RGB Data.....	6
2.1.1 Image Capturing.....	6
2.1.2 Input RGB Images.....	7
2.2 Airborne LIDAR Data.....	8
2.2.1 LIDAR System Components.....	9
2.2.2 Data Acquisition.....	10
2.2.3 Data Processing.....	12
2.2.4 LIDAR Data for Study Area.....	16
CHAPTER THREE: URBAN AREA SEGMENTATION AND CLASSIFICATION FROM SEMANTIC INFORMATION.....	21
3.1 Literature Review.....	21
3.2 Feature Classification Methodology.....	26
3.2.1 Color Invariant Image Production.....	26
3.2.2 Otsu Segmentation for Vegetation and Shadow Identification.....	27
3.2.3 Color Transformation for Feature Detection.....	28
3.3 New Implementation for Feature Extraction from RGB Image Using a Combination of Traditional Filtering Methods.....	32
3.4 Feature Extraction from Different RGB Images for Urban Classification.....	33
CHAPTER FOUR: BUILDING EXTRACTION FROM SPATIAL DATA FOR EFFICIENT URBAN AREA CLASSIFICATION FROM COMBINED LIDAR/RGB DATA.....	57
4.1 Literature Review.....	58
4.2 Literature Review for Wavelet Feature Extraction from LIDAR Data - Efficiency and Limitations.....	64
4.3 Multi-Resolution Analysis for Building Extraction from LIDAR Data.....	69
4.3.1 Wavelet Analysis for LIDAR Data Applications.....	70
4.3.2 Wavelets versus Fourier Transform.....	70

4.3.3 One Dimensional Continuous Wavelet Analysis	75
4.3.4 One Dimensional Discrete Wavelet Analysis	76
4.3.5 One Dimensional Multi-Resolution Analysis	77
4.3.6 Two Dimensional Multi-Resolution Analysis.....	78
4.4 New MRA Building Extraction Implementation from LIDAR Data	79
4.5 Multi Resolution Analysis of LIDAR Height Image.....	80
4.6 Data Fusion for 3-D RGB Building Modelling	84
4.7 Feature Classification from Combined RGB/LIDAR Data.....	95
4.8 Quantitative Assessment for the Automatic Classification Technique from Combined RGB/LIDAR Data Using Wavelets and Statistical Filters.....	102
 CHAPTER FIVE: MONTE CARLO SIMULATION DESCRIPTOR FOR OBJECT SHAPE RECOGNITION.....	112
5.1 Different Object Descriptors, Production and Assessment, Literature Review	114
5.2 Object Description and Classification Methodologies	117
5.2.1 Monte Carlo Simulation for Object and Image Processing Applications	117
5.2.2 Monte Carlo Simulation Descriptor for Object Recognition	118
5.2.3 Minimum Distance Classifier.....	119
5.3 Library Preparation from Building Extracted Image.....	120
5.3.1 Boundary and Texture Object Representation	123
5.3.2 Wavelet Object Description	126
5.3.3 MCS Object Description	128
5.4 Recognition Descriptors Assessment.....	129
5.4.1 Object Recognition from Original LIDAR Height Image.....	131
5.4.2 Object Recognition from Scaled and Oriented Original LIDAR Height Image.....	135
 CHAPTER SIX: SUMMARY, CONCLUSIONS, CONTRIBUTIONS AND RECOMMENDATIONS.....	142
6.1 Summary.....	142
6.2 Conclusions and Contributions.....	143
6.3 Recommendations.....	147
 REFERENCES	149
 APPENDICES.....	156

List of Tables

Table 2-1: Flight Configuration and Accuracies	17
Table 2-2: University of Calgary LIDAR Project Summary	17
Table 3-1: Color Ranges for Different Features Suggested by Bong et al. [2009].....	24
Table 4-1: Confusion Matrix of the Classification Results Versus GIS Data	107
Table 4-2: Confusion Matrix after Adding the Classification Results of Shadowy Areas	110
Table 5-1: MCSD for Library Objects – First Trial.....	129

List of Figures and Illustrations

Figure 1-1: Research Workflow.....	3
Figure 2-1: Airborne Image Production and Applications [Sandau, 2010]	7
Figure 2-2: Input RGB Data	8
Figure 2-3: LIDAR System Components	10
Figure 2-4: Linear Scanning Pattern.....	11
Figure 2-5: Zigzag Scanning Pattern	11
Figure 2-6: Nutating Scanning Pattern	12
Figure 2-7: Ellipsoidal Scanning Pattern	12
Figure 2-8: LIDAR Coordinate System.....	13
Figure 2-9: Range Image	15
Figure 2-10: Intensity Image.....	15
Figure 2-11: Optech ALTM 3100 Laser Scanner System	16
Figure 2-12: Nearest Neighbor Interpolation Technique.....	18
Figure 2-13: Height Image.....	19
Figure 2-14: Intensity Image.....	20
Figure 3-1: RGB Color Space.....	29
Figure 3-2: HSV Color Space.....	30
Figure 3-3: HSI Color Space.....	31
Figure 3-4: Algorithm of Feature Extraction from RGB [Elhifnawy et al., 2011c]	32
Figure 3-5: Input RGB - Image I [Image Courtesy of Twisted Sifter]	34
Figure 3-6: Color Invariant Image (Vegetation) - Image I.....	34
Figure 3-7: Otsu Segmentation Vegetation Image - Image I.....	35
Figure 3-8: Vegetation Candidates - Image I.....	35
Figure 3-9: RGB Vegetation Image - Image I	36

Figure 3-10: Invariant Image (Shadows) – Image I.....	37
Figure 3-11: Otsu Segmentation Shadow Image – Image I.....	37
Figure 3-12: Shadow Candidates – Image I.....	38
Figure 3-13: RGB Shadow Image – Image I.....	38
Figure 3-14: Input Image after Removing Vegetation and Shadows – Image I.....	39
Figure 3-15: Input Image in YCbCr Color Space – Image I.....	40
Figure 3-16: Input Image in HSV Color Space – Image I.....	40
Figure 3-17: Luminance Color Channel – Image I.....	41
Figure 3-18: Saturation Color Channel – Image I.....	42
Figure 3-19: Hue Color Channel – Image I.....	42
Figure 3-20: Road Candidate Image after Luminance Thresholding - Image I.....	43
Figure 3-21: RGB Road Image after Luminance Thresholding - Image I.....	44
Figure 3-22: Non-Road Candidate Image after Hue Thresholding - Image I.....	45
Figure 3-23: Non-Road Image after Hue Thresholding - Image I.....	45
Figure 3-24: Sandy and Unhealthy Vegetation Candidates after All Color Thresholding - Image I.....	46
Figure 3-25: RGB Sandy and Unhealthy Vegetation Areas Image after All Color Thresholding - Image I.....	47
Figure 3-26: Road Candidates from RGB Image - Image I.....	48
Figure 3-27: RGB Road Image - Image I.....	48
Figure 3-28: Building Candidates from RGB Image - Image I.....	49
Figure 3-29: RGB Building Image - Image I.....	50
Figure 3-30: Features Classification for Urban Area – Image I.....	51
Figure 3-31: RGB Input Image – Image II.....	52
Figure 3-32: RGB Vegetation Image – Image II.....	52
Figure 3-33: RGB Shadow Image – Image II.....	53

Figure 3-34: Input Image after Removing Vegetation and Shadows – Image II.....	53
Figure 3-35: RGB Sandy and Unhealthy Vegetation Areas Image - Image II	54
Figure 3-36: RGB Road Image - Image II	54
Figure 3-37: RGB Building Image - Image II	55
Figure 3-38: Features Classification for Urban Area – Image II	55
Figure 4-1: Fourier Transform Image	72
Figure 4-2: After Log Transformation	73
Figure 4-3: Samples of Different Wavelet Base Functions	75
Figure 4-4: Mallat Algorithm.....	78
Figure 4-5: Building Extraction Algorithm	80
Figure 4-6: LIDAR Height Image of the Area of Study	81
Figure 4-7: SYM4 Wavelet Base Function.....	81
Figure 4-8: Wavelet Analysis	82
Figure 4-9: Feature Edges of the Area of Study	83
Figure 4-10: Feature Image of the Area of Study	84
Figure 4-11: Building Candidates of the Area of Study	85
Figure 4-12: MOSAIC Includes the Area of Study	86
Figure 4-13: RGB Image of the Area of Study.....	86
Figure 4-14: Building Extraction Image of the Area of Study (Intensity Information)....	87
Figure 4-15: Building Extraction Image of the Area of Study (RGB)	88
Figure 4-16: Building Extraction Image of the Area of Study (Height Information).....	88
Figure 4-17: 3-D Building Model (View 1).....	89
Figure 4-18: 3-D Building Model (View 2).....	90
Figure 4-19: Input LIDAR Height Image	91
Figure 4-20: Building Extraction Image (Height)	92

Figure 4-21: Building Extraction Image (RGB)	92
Figure 4-22: 3-D Building Model (View1).....	93
Figure 4-23: 3-D Building Model (View1).....	93
Figure 4-24: Input LIDAR Height Image (British Columbia).....	94
Figure 4-25: 3-D Building Models (Second Data – View1).....	94
Figure 4-26: 3-D Building Models (Second Data – View2).....	95
Figure 4-27: RGB Input Image	96
Figure 4-28: Feature Classification Algorithm from RGB/LIDAR Data	97
Figure 4-29: Building Extraction Image after Removing Vegetation and Shadows	98
Figure 4-30: Sandy Areas after Eliminating Building Candidates	99
Figure 4-31: Road Image	100
Figure 4-32: Road Extraction Image after Considering Sandy Areas as Roads	101
Figure 4-33: Final Feature Classification Image [Elhabiby et al., 2011].....	101
Figure 4-34: Final Building Candidates from the Proposed Classification Technique using Combined RGB/LIDAR Data	103
Figure 4-35: GIS Data for the University of Calgary Including Buildings, Roads and Vegetation	104
Figure 4-36: Building Candidates from GIS Data	105
Figure 4-37: Road Candidates from GIS Data.....	105
Figure 4-38: Vegetation Candidates from GIS data	106
Figure 4-39: Building Classification Assessment versus GIS Data.....	108
Figure 4-40: Road Classification Assessment versus GIS Data	108
Figure 4-41: Vegetation Classification Assessment versus GIS Data	109
Figure 4-42: Classification Technique Assessment versus GIS Data for Building, Road and Vegetation Classification.....	109
Figure 5-1: Generated Random Values Segmentation.....	119

Figure 5-2: Library Preparation Algorithm.....	121
Figure 5-3: Building Image.....	122
Figure 5-4: Building Separation.....	123
Figure 5-5: Object Boundary Images.....	124
Figure 5-6: Object Texture Representation.....	125
Figure 5-7: Wavelet Function for Object Descriptors	127
Figure 5-8: Object Shape Recognition Algorithm	130
Figure 5-9: Object Recognition Results from Original Input Data using Wavelet Descriptors	132
Figure 5-10: Object Recognition Results from Original Input Data using MCSD.....	134
Figure 5-11: MCSD Recognition Successful Percentage form Original Input Data	134
Figure 5-12: Scaled and Oriented LIDAR Height Image	135
Figure 5-13: Scaled and Oriented RGB Image	136
Figure 5-14: Recognition Results for Scaled and Oriented Objects using Wavelet Descriptors	138
Figure 5-15: Wavelet Descriptors Recognition Successful Percentage for Scaled and Oriented Objects	139
Figure 5-16 Recognition Results for Scaled and Oriented Objects using MCSD	140
Figure 5-17: MCSD Recognition Successful Percentage for Scaled and oriented Objects	140

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
RGB	Color image with Red (R), Green (G) and Blue (B) color channels
LIDAR	Light Detection And Ranging
RGB	RGB color space
MCS	Monte Carlo Simulation
MRA	Multi- Resolution Analysis
AOSR	Automatic Object Shape Recognition
MCS D	Monte Carlo Simulation Descriptor
row, col	Image dimension (rows, columns)
GPS	Global Positioning System
INS	Inertial Navigation System
IMU	Inertial Measurement Unit
\vec{V}_p	Ground coordinates of point of interest
\vec{V}_a	Ground Coordinates of antenna phase center
\vec{V}_{bs}	Vector of lever arm between antenna phase center and center of laser scanner coordinate system with respect to IMU coordinate system
\vec{V}_r	Vector of range of point of interest with respect to laser beam coordinate system
R_{gi}	rotation matrix between ground coordinate system and IMU coordinate system
R_{il}	Rotation matrix between IMU coordinate system and laser scanner coordinate system
R_{lb}	Rotation matrix between laser scanner coordinate system and laser beam coordinate system
NN	Nearest Neighbor
AWT	A trous Wavelet Transform
KLT	Karhounen-Louve Transfrom

<i>RC</i>	Road Candidates
<i>CIE L * a * b *</i>	Commission International de l'Eclairage LAB color space
JSEG	J value SEGgmentation
<i>BI</i>	Building Index
<i>Y</i>	Luminance color channel
C_b	Difference between blue channel and a reference value
C_r	Difference between red channel and a reference value
<i>H</i>	Hue color channel
<i>S</i>	Saturation color channel
<i>V</i>	Value color channel
RADAR	Radio Detection And Ranging
<i>RI</i>	Ratio Index
<i>VARI</i>	Visible Atmospherically Resistant Index
<i>VI</i>	Vegetation Index
<i>SI</i>	Shadow Index
<i>r</i>	Gray value of image pixel
g_r	Number of pixels that have gray level r
<i>G</i>	Total number of image pixels
P_r	Probability of existence of gray level r
<i>L</i>	Total number of gray values
σ_B^2	Between-class variance
<i>k</i>	Gray thresholding value
<i>I</i>	Intensity color channel
<i>RCY</i>	Road Candidates from luminance color channel
<i>NRCH</i>	Non-Road Candidates from Hue color channel
<i>SC</i>	Sand areas Candidates
<i>BC</i>	Building Candidates
RANSAC	RANdom SAmple Consensus
DSM	Digital Surface Model

NDVI	Normalized Difference Vegetation Index
DTM	Digital Terrain Model
PTS	Planar Terrain Surface
DEM	Digital Elevation Model
GIS	Geographic Information System
ALSWave	Airborne Laser Scanner Wavelet
TIN	Triangular Irregular Network
CHM	Canopy Height Model
VWF	Variable Window Filter
SWA	Spatial Wavelet Analysis
DWT	Discrete Wavelet Transform
EMT	Expectation Maximization Technique
NCC	Normalize Cross Correlation
T	Period of Time
f	Signal or function
x	Space variable
A_p	Amplitude of periodic function number p
e	Exponential
i	Imaginary part = $\sqrt{-1}$
λ_p	Wavelength of periodic function number p
ω	Frequency
$A(\omega)$	Amplitude spectrum of f at frequency ω
IS	New intensity value after log transform
$\varphi(x, y)l$	Log transform constant
$\mathbb{G}\{f\}$	G transformation of f using window function g
t	Transition parameter
g	Window function
CWT	Continuous Wavelet Transform
W	Wavelet Transformation
ψ	Wavelet function

m	Scale parameter
n	Translation parameter
Z	Domain of integers
λ_o	Scale
t_o	Translation
φ	Scaling function
M	Total number of signal sampling
c	Scaling coefficients
d	Detail coefficients
N	Total number of levels of decomposition
\odot	Tensor product
$\varphi(x, y)$	Two dimensional scaling function
$\psi^H(x, y)$	Two dimensional horizontal wavelet function
$\psi^V(x, y)$	Two dimensional vertical wavelet function
$\psi^D(x, y)$	Two dimensional diagonal wavelet function
x, y	Coordinates of the output image
w, z	Coordinates of the input image
s_x, s_y	Scale parameters in x and y directions
MMS	Mobile Mapping System
ATR	Automatic Target Recognition
AMCR	Active Monte Carlo Recognition
OCE	Object oriented Contour Extractor
CDCA	Curvature Dependent Contour Approximation
KNN	K-Nearest Neighbour
PDF	Probability Density Function
$P_z(z)$	Probability distribution of random variable z
a, b	Probability distribution parameters
m	Mean
σ^2	Variance
$df_1, df_2, df_3, \dots, df_{DF}$	Decision functions

<i>DF</i>	Total Number of classes
<i>MDCT</i>	Minimum Distance Classifier Technique
<i>ED</i>	Euclidean Distance
<i>OD</i>	Object of interest descriptor
<i>LD</i>	Library of classes descriptors
<i>BF</i>	Boundary function
<i>q</i>	Order of boundary pixel
<i>b_q</i>	Total number of boundary pixels
<i>NRANV</i>	Total number of random values
<i>u</i>	The order of the descriptor in MCSD vector

Chapter One: Introduction

The hybrid data (Light Detection And Ranging (LIDAR) and Red, Green and Blue (RGB)) contains many features. Feature extraction is the process of separating features in individual data sets. Each data set contains all information from the input data for specific feature. Feature classification is the process of defining all features in the input data to classes. Object recognition is the process of identifying an object of interest to specific class. Feature extraction, classification and recognition are the basis and target for many civil and military applications such as urban planning and development, forest monitoring, road network extensions, battle field monitoring and target detection. Feature extraction and classification applications sometimes have many obstructions depending on data availability and quality. This research thesis introduces an automatic urban area classification and recognition technique from hybrid data.

1.1 Motivations

The availability of high quality hybrid data, namely digital images and LIDAR data captured in the same flight missions was the main motivation for introducing a combined algorithm for efficient urban area classification using the properties of these two sources of information. Although the RGB image provides rich descriptive information for ground and non-ground objects, detecting and separating different features sharing the same texture or color properties remains difficult especially in the *RGB* color space. In contrast, the LIDAR data provides rapid and dense information for surface elevation and non-terrain objects however efficient feature extraction from LIDAR data can face the same problem detecting and separating different objects sharing the same height values or located close proximity to each other. These two difficulties, from both RGB and LIDAR

data, represent a motivation for combining semantic and spatial information to classify and separate all different features efficiently. The proposed research thesis introduces an automatic urban area classification technique using combined RGB/LIDAR data. The proposed technique is composed of several color transformation and segmentation methods combined together and applied with RGB images. The semantic transformation techniques are aided by wavelet transformation of the LIDAR height data for building extraction. The new combination of color transformation-RGB and wavelet-LIDAR introduces an automatic urban classification technique for extracting roads, green areas, buildings and sand areas.

The availability of detailed classified objects from the extraction process was the motivation behind having one more step toward an efficient change detection algorithm for buildings. The objects identified and classified from the extraction process are then used in creating an object recognition library that can be used in change detection search algorithms. Extracted objects are represented generally by their boundaries. Monte Carlo Simulation (MCS) technique is used to get each object area based on generating random values inside each object boundary. The ratio between the numbers of random values inside an object boundary to that covering the object image, represents the ratio between the object area and the complete image area. This research thesis assesses the use of different descriptors and their corresponding efficiency for feature recognition applications.

1.2 Thesis Objectives

The main objective of the thesis is to implement an efficient fully automatic feature classification and recognition technique from combined RGB and LIDAR data for urban

areas. Figure 1-1 shows the schematic diagram of an overview about the thesis workflow. The thesis work begins with the extraction and classification of different features in an urban area. The extracted features are then used to build a library of different classes used for recognizing any object of interest. The object of interest can be identified from RGB, LIDAR, intensity, satellite or any other imaging data captured at any time.

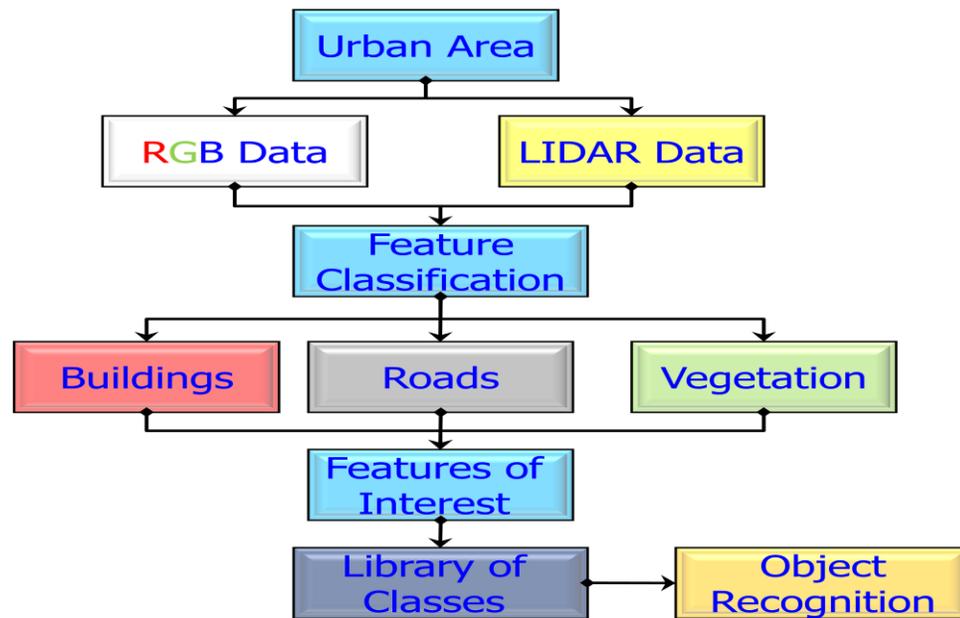


Figure 1-1: Research Workflow

Consequently, the thesis has two sub-objectives. The first objective is to introduce and implement an efficient feature extraction and classification technique from RGB and LIDAR data. This objective is achieved by creating a combination between color segmentation, color transformation, global color thresholding techniques and image Multi-Resolution Analysis (MRA) using wavelet transform. The second objective is to introduce a new spatial domain object descriptive for Automatic Object Shape Recognition (AOSR) for change detection. The investigated spatial domain object

descriptor is based on the Monte Carlo Simulation (MCS) application for object area determination. The thesis will include the following sub-objectives that lead to the main objectives as follows:

- Implementation of an automated feature classification technique for urban areas from RGB image.
- Implementation of an automated Multi Resolution Analysis (MRA) building extraction technique from LIDAR data.
- Investigation of a new spatial domain descriptor based on Monte Carlo Simulation (MCS) application.
- Assessment of Monte Carlo Simulation Descriptor (MCSD) versus other object descriptors for shape recognition application.
- Developing a complete software package for automatic feature extraction and recognition for urban applications, with minimum need of human intervention. The package can deal with RGB only for image classification, LIDAR data only for building extraction and combined hybrid data for complete urban classification of complex images and data.

1.3 Thesis Outline

The thesis is arranged into six chapters. Chapter One is for the introduction and overview of the research target and contents.

Chapter Two introduces an overview of the different available data. It introduces an RGB image, components and properties. Also, it explains the basics of the LIDAR system components and LIDAR data, collecting and processing.

Chapter Three introduces a feature classification technique from RGB image. It explains the feature classification algorithm and all methodologies. It shows the implementation of the proposed algorithm through urban classification results of urban area images.

Chapter Four introduces a building extraction technique from LIDAR data. It explains the building extraction algorithm and mathematical models of wavelet analysis for signal and image processing. It shows the implementation of the proposed algorithm through building extraction results from LIDAR data for urban areas.

Chapter Five introduces a new spatial domain object descriptive. It explains the production of an object descriptive from a boundary image. It also introduces the assessment of the new descriptor versus wavelet descriptors for object recognition using different objects with random shapes, orientations and scales.

Chapter Six provides the summary and conclusion of this research. It shows the contribution from all proposed techniques and mentions some of interesting research topics to extend and develop this research thesis as recommendations for future work.

Chapter Two: Airborne RGB/LIDAR Data

As mentioned in Chapter One, the thesis target is the investigation of feature classification and recognition techniques using available data. Two different types of data are prepared for this research. RGB and LIDAR data are produced for the same study area representing an urban area located in city of Calgary, Alberta in Canada. By considering any of these data as complimentary to one another, they can be used to verify an efficient urban classification technique using RGB/LIDAR data. This chapter introduces an overview of airborne RGB and LIDAR data and the images used for verifying and assessing the proposed techniques.

2.1 Airborne RGB Data

An RGB image is the best descriptive data for ground and non-ground objects. It is composed of three color channels red, green and blue. All color channels have the same dimension *row*, number of rows, and *col*, number of columns. Mathematically, an RGB image is represented as a matrix of dimension $row \times col \times 3$.

2.1.1 *Image Capturing*

There are two methods to produce an RGB digital image. The first method is classified as a post mission by scanning an analogue image with a digital scanner. This is not a preferable method since the resultant image may contain scanning errors in addition to raw production errors such as lens distortion and atmospheric refraction effects. The second method is preferable however, because an RGB image is captured directly, using a digital camera. Figure 2-1 explains the difference between the two methods and their applications [Sandau, 2010].

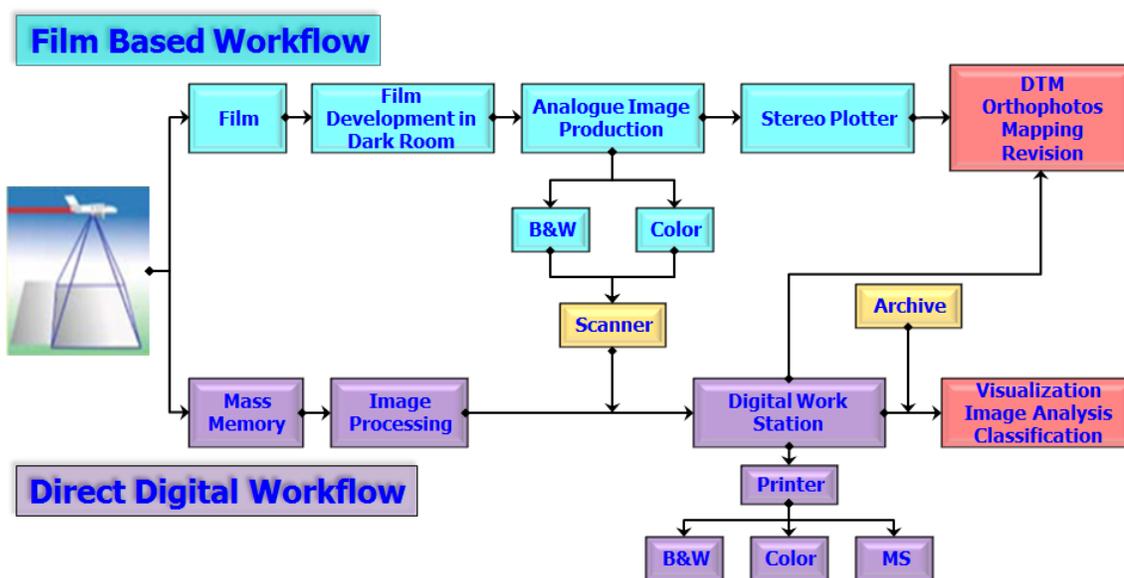


Figure 2-1: Airborne Image Production and Applications [Sandau, 2010]

2.1.2 *Input RGB Images*

Two different RGB images are used for verifying the proposed feature classification algorithm from RGB data. The first RGB image for urban area (Housing Subdivision, Arizona) contains buildings, vegetation and roads as shown in Figure 2-2(a). The selected image has a high differentiable and distinguishable contrast between different features. Different features can be efficiently identified and classified from this image.

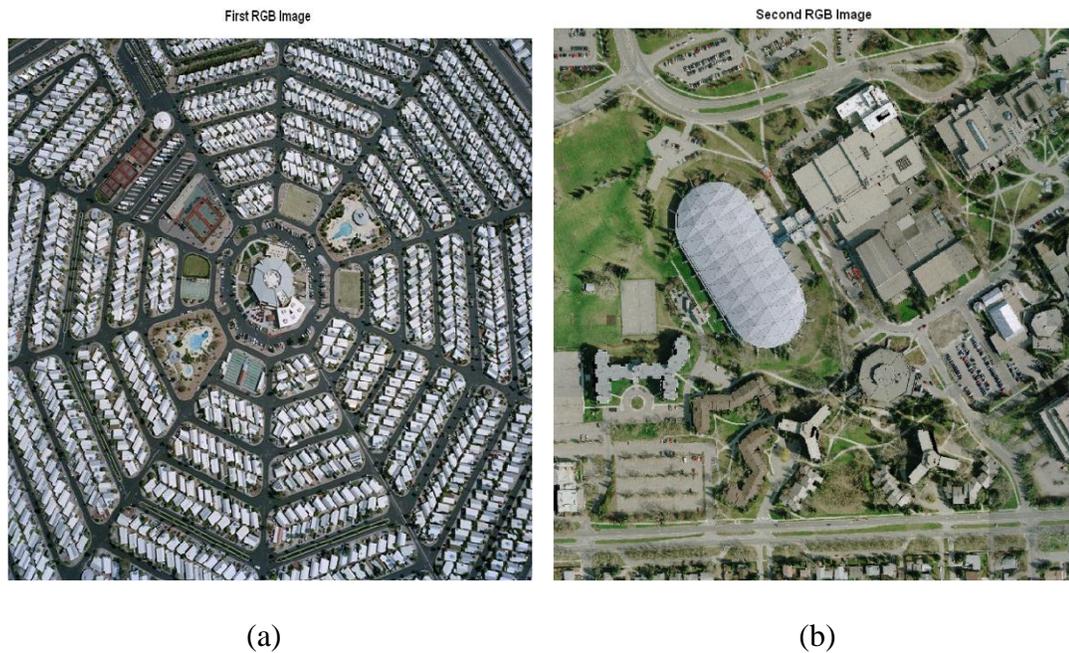


Figure 2-2: Input RGB Data

The second input RGB image is a part of the University of Calgary campus as shown in Figure 2-2(b). Building and road features share texture properties. Features cannot be correctly classified from this image based on the shared texture properties of different objects. This is a motivation for using additional information such as spatial data to solve the problems that arise from these shared texture properties and effectively extract and classify different features. This will be explained clearly through the results discussed in Chapters 3 and 4.

2.2 Airborne LIDAR Data

A LIDAR system provides dense positioning information for surface topography and all features for the study area. LIDAR systems provide not only positioning information but also property information, referred to as intensity values, for all scanned points. Two types of images can be produced from this information, a range image and an intensity

image. A range image is based on height values while an intensity image is based on the intensity values for all scanned points [Shan and Toth, 2009]. The LIDAR system is composed of many sensors to collect this data with high density and accuracy.

2.2.1 *LIDAR System Components*

A LIDAR system consists of two main components, a data collection component and a navigation component. For the data collection component, a laser scanner is the main sensor as it collects both dense range and intensity information for the scanned area [El-Sheimy et al., 2005; Al-Durgham, 2007; Habib et al., 2009]. Newer versions of the LIDAR system contain additional data collection sensors such as a digital camera and/or an infrared sensor. Not only do these systems collect positional information, they also collect semantic information for each scanned point. The LIDAR system consists of a laser scanner and a digital camera, as an additional data collection sensor, used for producing LIDAR, height and intensity, and at the same time, an RGB image for the same scanned area. The classification results using the data produced from this system are a significantly better quality than those using each data separately for any given study area. This will be illustrated in more detail in Chapters 3 and 4.

As for the navigation component, it consists of two systems; Global Positioning System (GPS) and Inertial Navigation System (INS). Figure 2-3 identifies the main components of the LIDAR system; laser scanner that provides range and intensity data, GPS that provides 3D positioning and timing parameters, while Inertial Measurement Unit (IMU) that provides orientation parameters.



Figure 2-3: LIDAR System Components

2.2.2 Data Acquisition

The laser scanner contains a rotating mirror that reflects the laser beam and produces the required scanning pattern. The laser scanner may contain one mirror; in this case, the scanning pattern will be linear, zigzag or nutating. The scanner pattern assumes an ellipsoidal shape when the laser scanner contains two rotating mirrors. A linear scanning pattern (Figure 2-4) is undesirable since no data is collected from the areas between the lines. Alternatively, a zigzag scanning pattern (Figure 2-5) provides a continuous collection of data of the scanning area. A nutating scanning pattern (Figure 2-6) provides dense collection of data with a low flying speed. Amongst all the scanning patterns listed, the elliptical scanning pattern (Figure 2-7) is preferable because it gives a dense collection of data from any flying speed since it duplicates the data collected from the scanned area, however, it does require more complicated calculations as a result of the two existing mirrors that rotate at different scanning angles [El-Sheimy et al., 2005].

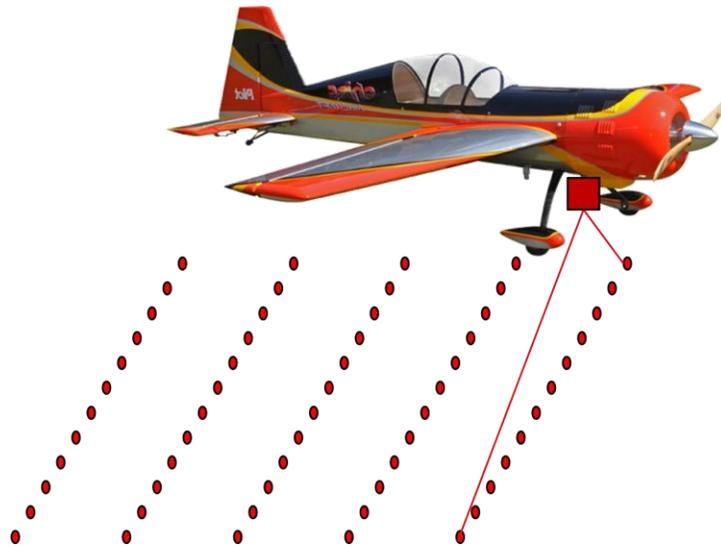


Figure 2-4: Linear Scanning Pattern



Figure 2-5: Zigzag Scanning Pattern

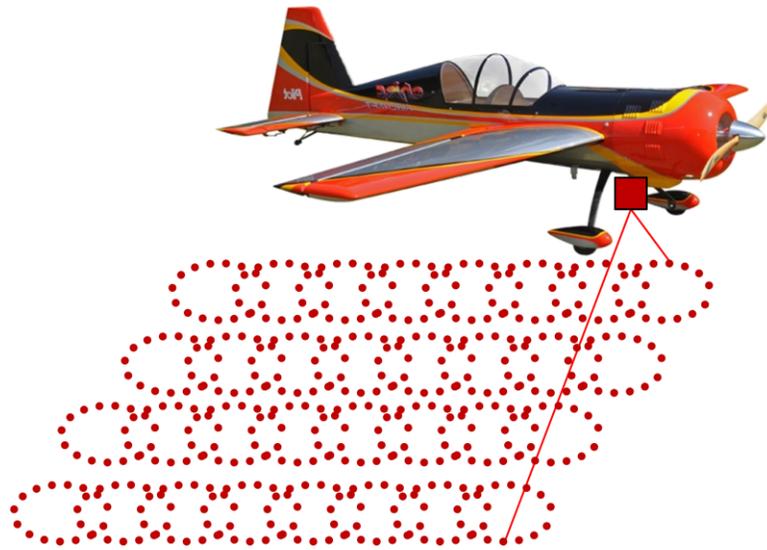


Figure 2-6: Nutating Scanning Pattern

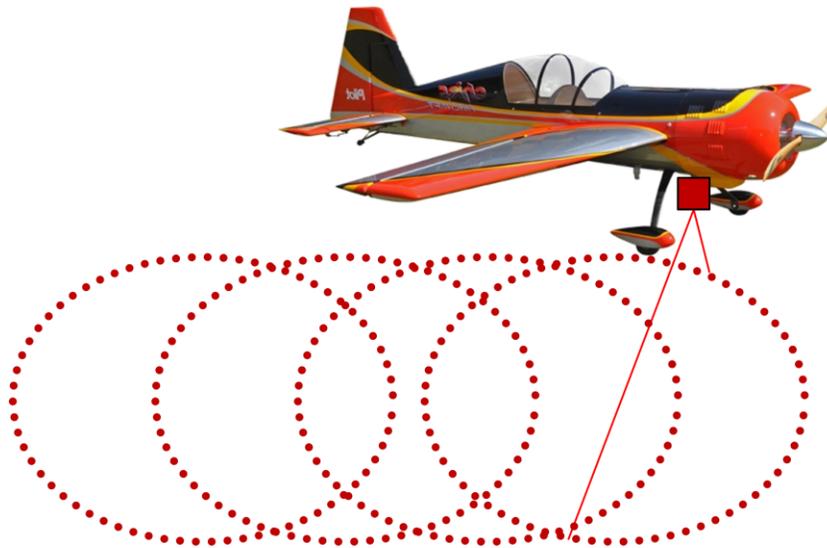


Figure 2-7: Ellipsoidal Scanning Pattern

2.2.3 Data Processing

An Inertial Measurement Unit (IMU) collects navigation data for flying missions while a GPS antenna collects coordinates of flying directions referenced to the GPS coordinate system. The laser scanner collects range and intensity information for each scanned point. An IMU and GPS are working together to provide integrated navigation data; data which

is used to calculate the coordinates of the flying direction referenced to the ground coordinate system. These coordinates are used in collaborating with the LIDAR system calibration parameters and range information to calculate ground coordinates for each scanned point. A LIDAR system is a direct georeferencing system, and as such the ground coordinates of the point of interest are calculated as a vector summation, shown in Figure 2-8, that represents a LIDAR equation [El-Sheimy et al., 2005]. Equation 2.1 [Habib et al., 2009] represents a mathematical model of LIDAR equation for calculating ground coordinates of point of interest.

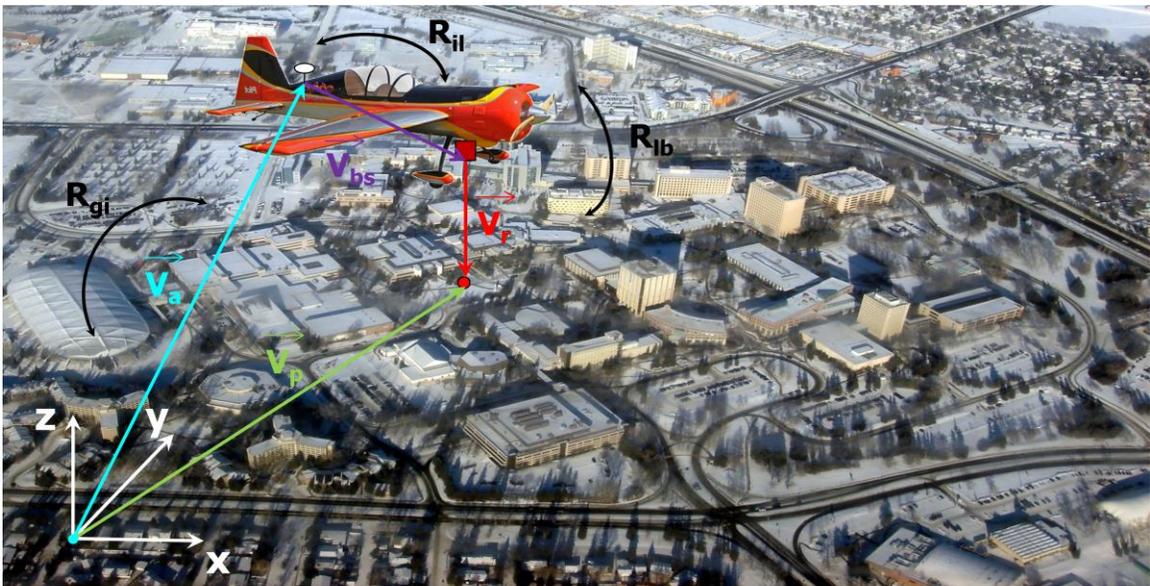


Figure 2-8: LIDAR Coordinate System

$$\vec{V}_p = \vec{V}_a + R_{gi}\vec{V}_{bs} + R_{gi}R_{il}R_{lb}\vec{V}_r \quad 2.1$$

- where: \vec{V}_p = ground coordinates of point of interest
- \vec{V}_a = ground coordinates of antenna phase center
- \vec{V}_{bs} = vector of lever arm between antenna phase center and center of laser scanner coordinate system with respect to IMU coordinate system
- \vec{V}_r = vector of range of point of interest with respect to laser beam coordinate system
- R_{gi} = rotation matrix between ground coordinate system and IMU coordinate system
- R_{il} = rotation matrix between IMU coordinate system and laser scanner coordinate system
- R_{lb} = rotation matrix between laser scanner coordinate system and laser beam coordinate system

Positional, spatial and intensity information are gridded using an interpolation technique to produce LIDAR images, a range image and an intensity image. The grid remains the same for both images, however while a range image requires the pixel values to represent the heights of scanned points, an intensity image requires the pixel values to represent intensity information [Al-Durgham, 2007]. Figure 2-9 and Figure 2-10 show a range (height) image and an intensity image for the same area, and both images are of the same area depicted in Figure 2-2b.

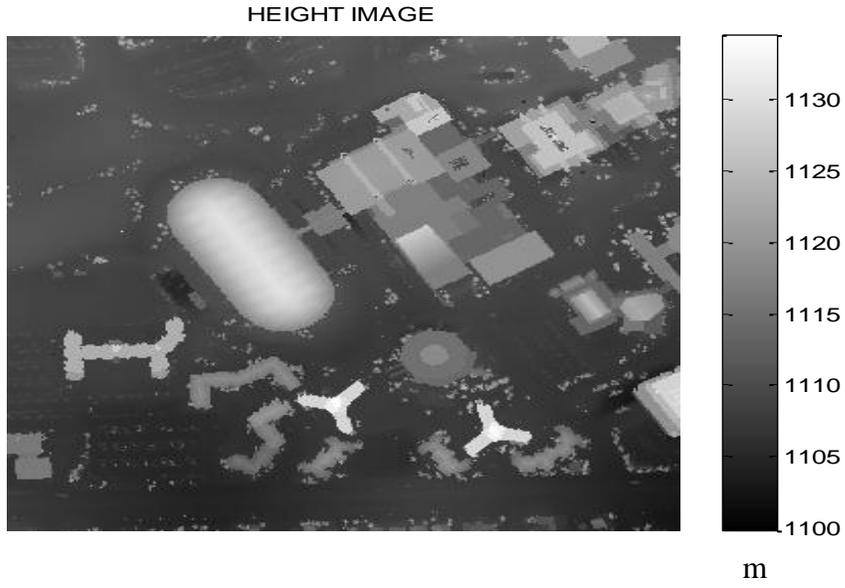


Figure 2-9: Range Image

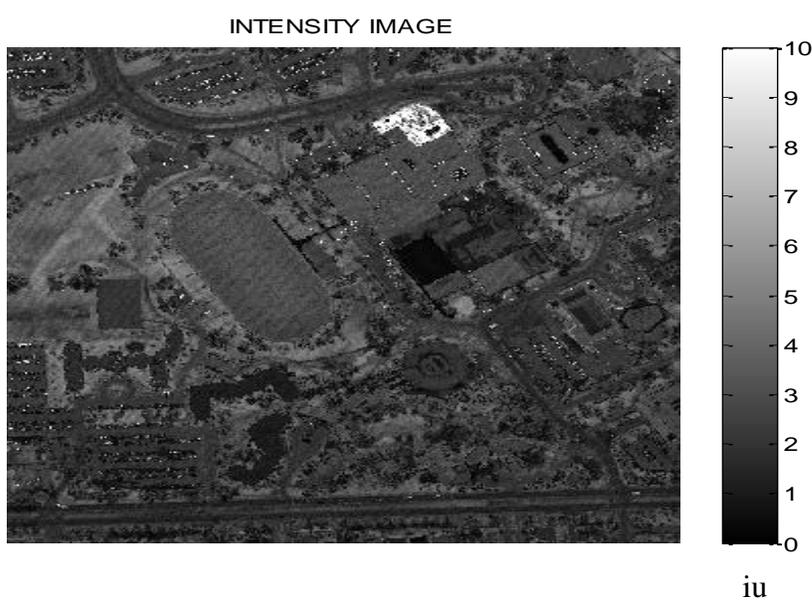


Figure 2-10: Intensity Image

2.2.4 LIDAR Data for Study Area

LIDAR data is used for building extraction as mentioned in Chapter One, so the selected area of study has to be focused on a region rich with buildings. This condition is provided by the selected area of study that is located in Calgary, Canada and covers part of the University of Calgary campus. Part of the selected area covers the same area as the RGB image shown in Figure 2-2b. The area of study was scanned by Optech ALTM 3100 LIDAR system (Figure 2-11).

Table 2-1 lists the flight configuration and accuracy specifications of this mission and Table 2-2 outlines the summary of this LIDAR project produced by Airborne Imaging Company on May 22, 2007.



Figure 2-11: Optech ALTM 3100 Laser Scanner System

Table 2-1: Flight Configuration and Accuracies

Vertical Accuracy	Horizontal Accuracy		2 Surveying	1 st Day	6 strips @1000m
15 cm	50cm@1000m	70cm@1400m	Days	2 nd Day	4 strips @1400m

Table 2-2: University of Calgary LIDAR Project Summary

Project Info:	Project Name:	U of C Campus	Client:	Univ. Of Calgary	
	Project Number:	8077	Acquisition:	Spring 2007	
General Information	Project Type:	Wild Area			
	Project Location:	Calgary, Ab, Canada			
	Project Size:	5.8 sq km			
Acquisition Parameters	Flying Height AGL:	1000/1400 m	Pulse Rate Rep:	50 kHz	
	Flying Speed:	160 kts	Scan Frequency:	33/38 Hz	
	Side Lap:	50 %	Scan Angle:	$\pm 25^\circ$	
	Point Spacing:	0.75 m	LIDAR System:	Optech 3100	
Geodetic Control	Horizontal Datum:	Nad83 CSRS	Vertical Datum:	CGVD28	
	Geoid Model:	HT2.0	UTM Zone:	11	
	Notes: local geodetic network was established fixed to the following governmental control:				
	Control Station (ASCM):	Sta_ID	Lat	Lon	MSL Hgt
	77x1058	51 5 47.10080	-114 22 24.08720	1198.909 m	
Observed Accuracy	Ground Truth RMS:	9 cm			

This project provides ten LIDAR strips for the whole scanned area. The area of study is the fourth strip from this project; it is bounded with X-coordinates from 699,915.05583963 m to 701,414.63341027 m and Y-coordinates from 5,661,929.8682168 m to 5,662,769.353172 m. The LIDAR system scanned this area with a zigzag scanning system with a total amount of scanned points equaling 982,956. The distribution of scanned points is irregular and it is therefore necessary to regulate this distribution to produce height and intensity images. Redistribution of these points using an interpolation technique generates an equally spaced mesh and calculates height and intensity information for each node. The building extraction is based on detecting its edge, so a gridding or an interpolation method used for producing a height image must not change or distort height information of building edges. This condition is only fulfilled when using the Nearest Neighbor (NN) interpolation technique. Figure 2-12 shows the mathematical principal for NN interpolation technique that is applied to generate height and intensity images over a 0.5-m equally spaced mesh.

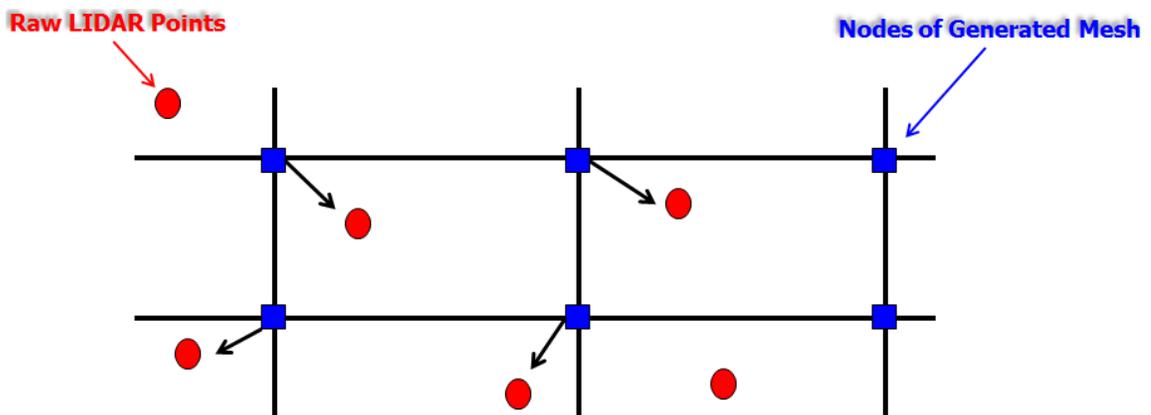


Figure 2-12: Nearest Neighbor Interpolation Technique

Figure 2-13 shows the height image after the interpolation process. The pixel values represent the height information for objects. Figure 2-14 shows the descriptive data of features produced by the LIDAR data referred to as the intensity image. The pixel values in the intensity image are calculated based on the transmitted and received energy of the laser pulse. These values depend on the object material and its efficiency for absorbing laser energy [Vain et al., 2010].

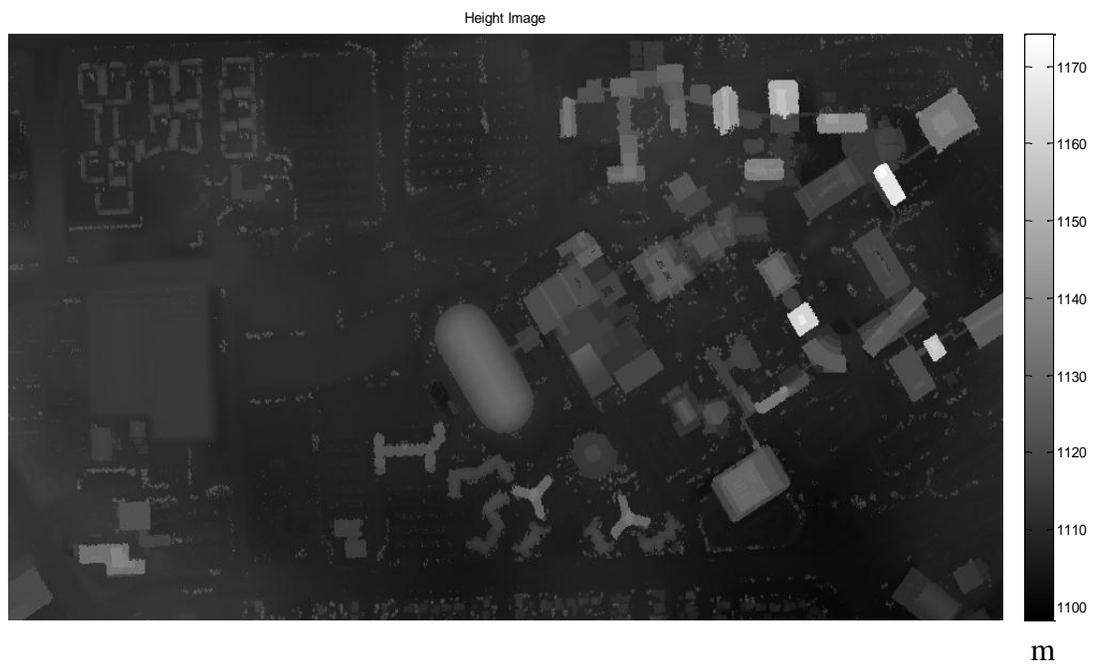


Figure 2-13: Height Image



Figure 2-14: Intensity Image

The building extraction algorithm from LIDAR data is applied to the height image by detecting the locations of height changes using a Multi-Resolution Analysis (MRA) tool as will be mentioned later in Chapter 4. All buildings can be extracted, but other features sharing height properties such as trees located close to the buildings cannot be separated. This is the motivation for using combined height image (Figure 2-9) and an RGB image (Figure 2-2b) to enable an efficient separation of these features. The results discussed in Chapter 4 will help to explain this in further detail.

Chapter Three: Urban Area Segmentation and Classification from Semantic Information

As mentioned in the previous chapter, an RGB image provides rich semantic information for ground and non-ground objects [Ghanma, 2006]. Semantic information is a powerful feature descriptive because different objects can be identified visually from RGB images. In this thesis, detailed investigation of a new fully automated technique for feature extraction from RGB data is introduced based on the combinations of supervised and unsupervised segmentation methods. The proposed technique is used for extracting vegetation, shadows, sandy areas and roads with only the buildings left as a result of these successive removals. This is done through the use of different color spaces indexing based on statistical assumptions and empirical formulas corresponding to certain color candidates.

3.1 Literature Review

There are several researchers who have worked on the extraction of different features from digital, aerial or satellite, images such as Laptev et al. [2000], Mena and Malpica [2005], Ooi and Lim [2006], Dong et al. [2007], Li et al. [2007b], Mokhtarzade and Zoej [2007], and Tuncer [2007] .

Tuncer [2007] introduced a fully automated technique for road extraction from satellite images. The research paper began with transforming the satellite image to a gray scale image while wavelet filtering based on A trous Wavelet Transform (AWT) algorithm is applied on this gray scale image. This application employed two wavelet base functions, Haar and Daubechies (db8), while stopping at the fourth level of decomposition. The two results are fused in one image using the Karhounen-Louve Transform (KLT). Fuzzy logic

and Hough transform are used for building a fuzzy logic interface algorithm to detect road candidates. Equation 3.1 shows a mathematical model for detecting road candidates, RC . This equation is applied for each window after dividing the fused image into small windows.

$$RC = \begin{cases} 1 & \text{if (Mean is Average) and (Std is Low) and (Hough is Line)} \\ 0 & \text{Otherwise} \end{cases} \quad \mathbf{3.1}$$

This technique provides acceptable results in the case of non-urban areas, but for urban areas some building pixels are extracted as road candidates.

Song and Shan [2008] investigated a building extraction technique from high resolution RGB images. The input image is transformed to Commission International de l'Eclairage LAB ($CIE L * a * b *$) color space, where L is a vertical axis that represents lightness, a axis is circular, normal to L axis, and runs horizontally from red (positive values) to green (negative values), and b axis is also circular and runs from yellow (positive values) to blue (negative values). The output image is de-noised using an anisotropic diffusion technique that is based on Laplacian filtering of the Gaussian smoothing image. This diffusion is applied on each color channel of the output image and active contour image segmentation is applied to detect building boundaries. The J value segmentation (JSEG) frame worm is used to construct building polygons and a 3-D wired frame. This technique has succeeded in case of red rooftop buildings that provide high contrast between buildings and background, but has presents issues when dealing with different rooftop colors.

Sirmacek and Unsalan [2008] detected building pixels based on calculating building index (BI) from red and green color channels as shown in Equation 3.2, where R and G represent red and green color channels in an input RGB image.

$$BI = \frac{4}{\pi} \arctan\left(\frac{R - G}{R + G}\right) \quad 3.2$$

This technique succeeded in cases of red rooftop buildings only, but it is unsuitable for other rooftop colors.

Bong et al. [2009] investigated color channel ranges that represent a road texture property after transforming an RGB image from RGB color space to YC_bC_r color space, where Y refers to a luminance color channel that represents gray scale information, C_b and C_r components represent color differences between blue and red channels and reference values respectively, and HSV color space, where H represents the hue color channel that represents a true color, S represents the saturation color channel that represents the degradation measurements for diluting a true color by a white light and V represents the value color channel corresponding to brightness color spaces. This research produced color ranges that represent different features concentrating on luminance, hue and saturation color channels as shown in Table 3-1.

Table 3-1: Color Ranges for Different Features Suggested by Bong et al. [2009]

Color Channel	Features		
	Road	Sand	Building
Luminance (Y)	110 – 160	100 – 200	50 – 150
Saturation (S)	0.1 – 0.25	0.3 – 0.5	0.1 – 0.5
Hue (H)	0.05 – 0.2	0.05 – 0.1	0 – 0.1

Road image is produced by applying global color thresholding on color channels using specific ranges as shown in Equation 3.3, where RC represents road candidate and Y , S and H represent luminance, saturation, and hue color channels.

$$RC = \begin{cases} 1 & \text{if } (100 < Y) \text{ and } (S < 0.3) \text{ and } (H \geq 0.05) \\ 0 & \text{Otherwise} \end{cases} \quad \mathbf{3.3}$$

This technique is acceptable for high resolution satellite images, but with clearly differentiable and distinguishable semantic information between buildings and roads; such as red rooftop buildings that avoid extracting building pixels as road candidates.

Shorter and Kasparis [2009] investigated an automatic technique for building extraction from an RGB aerial image. The proposed technique started with a color quantization process. The output image is segmented, and small areas are removed based on a specific area thresholding value since small areas are classified as ground areas. The segmented image is used for producing a color invariant vegetation image based on green and blue color channels by calculating the vegetation index. The raw RGB image is used for producing another color invariant image for shadow detection based on all RGB color channels by calculating the shadow index. The watershed segmentation technique is

applied on a raw RGB image followed by calculating solidity properties for all segmented regions to investigate building and non-building candidates in the input image. The output image is investigated after removing vegetation and shadowy areas from the building candidate image. This technique succeeded in detecting vegetation, shadows and building areas from the RGB image, but it depends on the input image and color varieties in all color channels. The main drawback of this approach is that roads with texture properties similar to buildings are identified as building candidates.

Feature classification resulting from RGB images is challenging when there are features sharing similar texture properties. Wegner et al. [2009] used an RGB image aided by Radio Detection And Ranging (RADAR) data to overcome the problem of building extraction, however, this aid could not provide an acceptable solution for extracting all buildings.

These feature extraction problems are the motivation behind this research thesis for investigating a combined technique that maximizes the benefits of different feature extraction algorithms. This thesis will introduce an automated feature classification technique from an RGB image based on the combination of two different algorithms, gray scale segmentation and color thresholding [Gonzalez and Woods, 2002].

The proposed feature extraction technique from an RGB image is aided by spatial information using the power of Multi-Resolution Analysis (MRA) for height information for building extraction from LIDAR data. This will create an efficient automatic feature classification technique that overcomes all problems from the existence of different features sharing texture properties. The algorithm improvement by aiding with LIDAR data will be illustrated later in the next chapter in detail.

3.2 Feature Classification Methodology

A feature extraction technique from an RGB image is a well-known topic and there are many techniques that have been used such as color invariant images production, color segmentation, color space transformation and global color thresholding [Sirmacek and Unsalan, 2008; Song and Shan, 2008; Bong et al., 2009]. The proposed feature classification technique from an RGB image will be investigated by using and combining several image processing techniques, color invariant image production, Otsu segmentation, color transformation and global color thresholding. This section introduces an overview of different color image processing techniques in addition to their respective mathematical representations used while investigating the proposed feature classification technique from RGB images.

3.2.1 Color Invariant Image Production

Color invariant image represents a ratio between radiometric measurements among different color channels. Basically, if this ratio is between green and blue color channels it is defined as Ratio Index (*RI*) [Jensen, 2005], if it is calculated using all color channels it is referred to as Visible Atmospherically Resistant Index (*VARI*) [Gitelson et al., 2002]. Equations 3.4 and 3.5 show the mathematical models for calculating *RI* and *VARI*.

$$RI = \frac{G}{B} \quad 3.4$$

$$VARI = \frac{G - R}{G + R - B} \quad 3.5$$

Two of these types of radiometric measurement ratios are the vegetation index and shadow index. Shorter and Kasparis [2009] used a Vegetation Index (*VI*) proposed by Boyer and Unsalan [2005] to investigate a color invariant image that is effective in

detecting vegetation areas after quantizing all the *RGB* color channels to sixteen color ranges. Also, the Shadow Index (*SI*) introduced by Sirmacek and Unsalan [2008] is used to identify shadow regions. Equations 3.6 and 3.7 show mathematical models for calculating *VI* and *SI*.

$$VI = \frac{4}{\pi} \arctan\left(\frac{G - B}{G + B}\right) \quad 3.6$$

$$SI = \frac{4}{\pi} \arctan\left[\frac{R - (R^2 + G^2 + B^2)^{\frac{1}{2}}}{R + (R^2 + G^2 + B^2)^{\frac{1}{2}}}\right] \quad 3.7$$

3.2.2 Otsu Segmentation for Vegetation and Shadow Identification

The Otsu thresholding process is an unsupervised segmentation process using a global thresholding technique to separate specific features. This method is called the gray thresholding technique because it is based on an image gray values histogram. The image histogram is normalized as shown in Equation 3.8 [Otsu, 1979].

$$P_r = g_r / G \quad 3.8$$

where: g_r = number of pixels that have gray level r

G = total number of image pixels

$P_r \geq 0$ = probability of existence of gray level r

r = 1,2,3,, L

L = total number of gray values

The selected gray thresholding k must maximize a specific value that is termed between-class variance (Equation 3.9) [Otsu, 1979; Gonzalez et al., 2004].

$$\sigma_B^2 = \omega_o(u_o - u_T)^2 + \omega_1(u_1 - u_T)^2 \quad 3.9$$

$$\text{where: } \omega_o = \sum_{r=1}^k P_r$$

$$\omega_1 = \sum_{r=k+1}^L P_r$$

$$u_o = \sum_{r=1}^k rP_r / \omega_o$$

$$u_1 = \sum_{r=k+1}^L rP_r / \omega_1$$

$$u_T = \sum_{r=1}^L rP_r$$

3.2.3 Color Transformation for Feature Detection

Although an RGB image produces rich semantic information for ground and non-ground objects, RGB color channels are not always suitable to use directly in feature extraction or classification. Consequently, the three channels data have to be transformed from *RGB* color space (Figure 3-1) to other color spaces. In this thesis, two color spaces are used for road and building extraction from the RGB image, YC_bC_r and *HSV* [Bong et al., 2009].

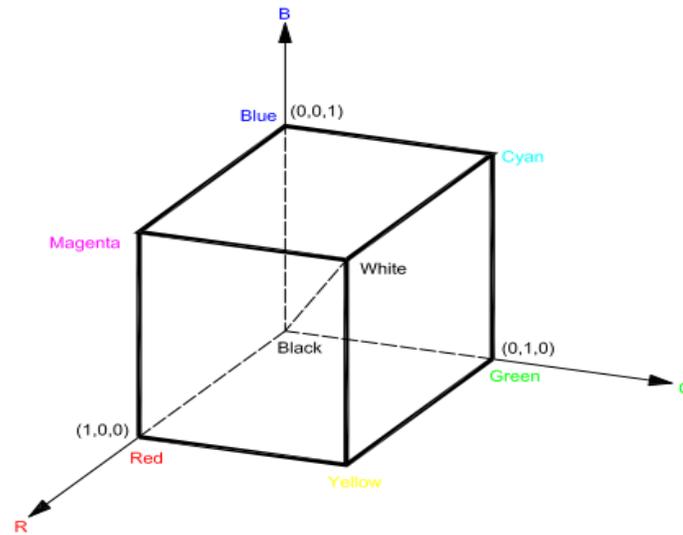


Figure 3-1: RGB Color Space

YC_bC_r color space is used in digital video applications; Equation 3.10 shows the mathematical model for color space transformation from RGB color space to YC_bC_r color space, where R, G and B representing red, green and blue color channels respectively.

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.699 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \mathbf{3.10}$$

HSV color space (Figure 3-2) is used for selecting colors. In HS , V is not suitable for human interpretation, so intensity value is used instead of the V color channel. This led to investigating HSI color space that is represented in triangular or circular shapes as shown in Figure 3-3. Equation 3.11 shows the mathematical model for color transformation from RGB color space to HSI color space [Gonzalez and Woods, 2002; Gonzalez et al., 2004].

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 & \text{if } B > G \end{cases}$$

$$\theta = \cos^{-1} \left[\frac{\frac{1}{2} [(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{\frac{1}{2}}} \right]$$

3.11

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

$$I = \frac{1}{3} (R + G + B)$$



Figure 3-2: HSV Color Space

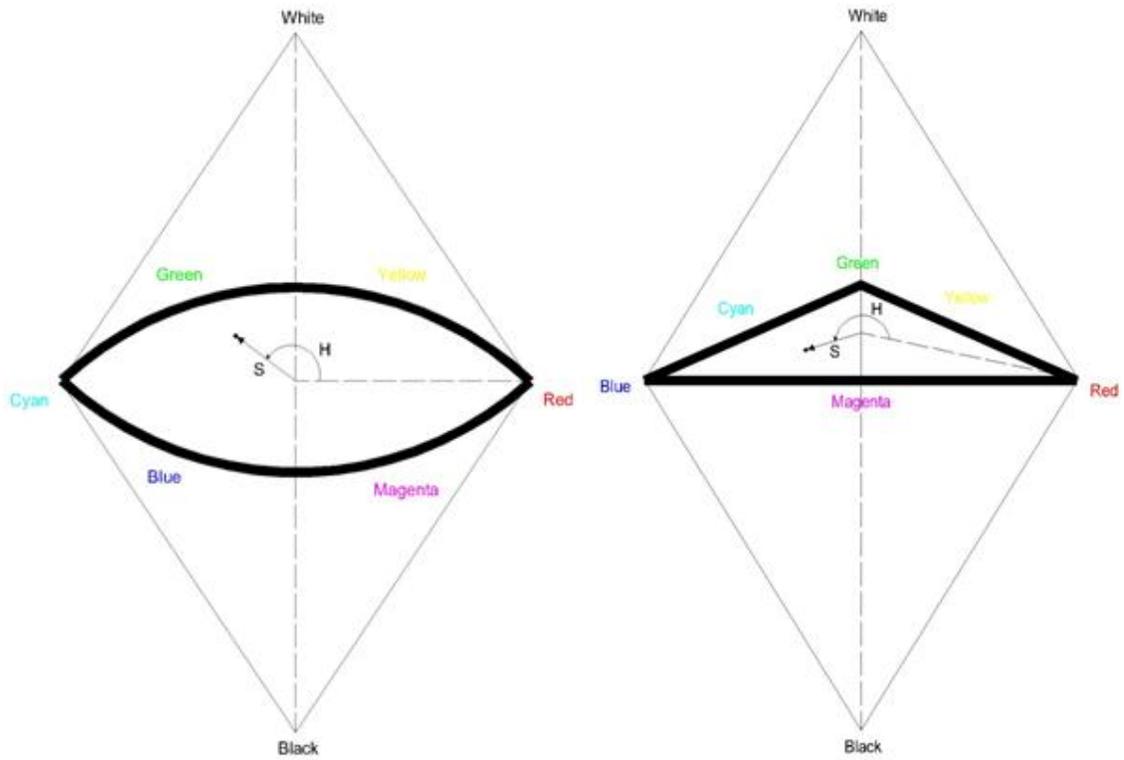


Figure 3-3: HSI Color Space

The proposed algorithm of feature classification from RGB data is a combination of various mentioned methods. Each method is used in a certain order to extract specific features to create an efficient urban classification from RGB images. First, vegetation and shadow area identification are done based on segmentation of color invariant images generated from RGB color channels. Second, roads are extracted using global color thresholding for luminance, hue and saturation color channels generated after transforming the data from RGB color space to YC_bC_r and HSV color spaces. The methodology for combining these techniques is discussed in the next section and through results represented in this chapter.

3.3 New Implementation for Feature Extraction from RGB Image Using a Combination of Traditional Filtering Methods

RGB images contain descriptive information, so feature extraction is based on identifying color ranges suitable for each feature. RGB color channels are used to calculate index values that are used for generating color invariant images for vegetation and shadowy area detection. The investigated technique relies on a combination of these techniques to be able to extract specific features from an RGB image. Figure 3-4 shows a schematic diagram of the proposed algorithm that consists of two main stages. The first stage, bounded by a red frame, is the color segmentation for generating color invariant images for vegetation and shadows identification. The second stage, bounded by a green frame, is the color thresholding for three color channels extracted after transforming the input image into two different color spaces to identify roads, subsequent to the removal of vegetation and shadows.

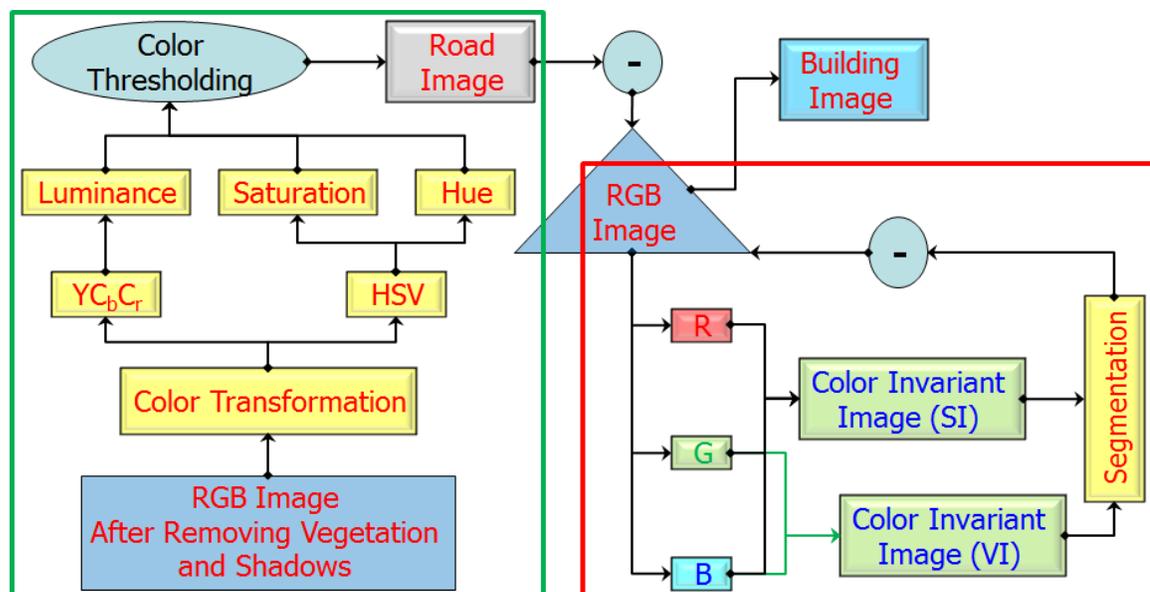


Figure 3-4: Algorithm of Feature Extraction from RGB [Elhifnawy et al., 2011c]

Vegetation, shadows, sandy areas and roads are detected at the end of the second stage. After removing successive extracted features the remaining features in the input image are classified as buildings.

3.4 Feature Extraction from Different RGB Images for Urban Classification

The algorithm is applied on different images with varying characteristics. The first image, Image I, shown in Figure 3-5 has no buildings with red rooftops, but with high differentiable texture properties among different features. Figure 3-6 shows the color invariant image generated after calculating the vegetation index (VI), as explained in section 3.2.1 using Equation 3.6, for vegetation identification. Figure 3-7 and Figure 3-8 show the segmented image after applying the Otsu segmentation technique on a color invariant image to detect the vegetation candidates, as explained in section 3.2.2. The candidate image is of logical class with ones in candidate pixels and zeros for others. Figure 3-9 shows a vegetation image in RGB color space after fusing RGB color channels with the vegetation candidate image.



Figure 3-5: Input RGB - Image I [Image Courtesy of Twisted Sifter]

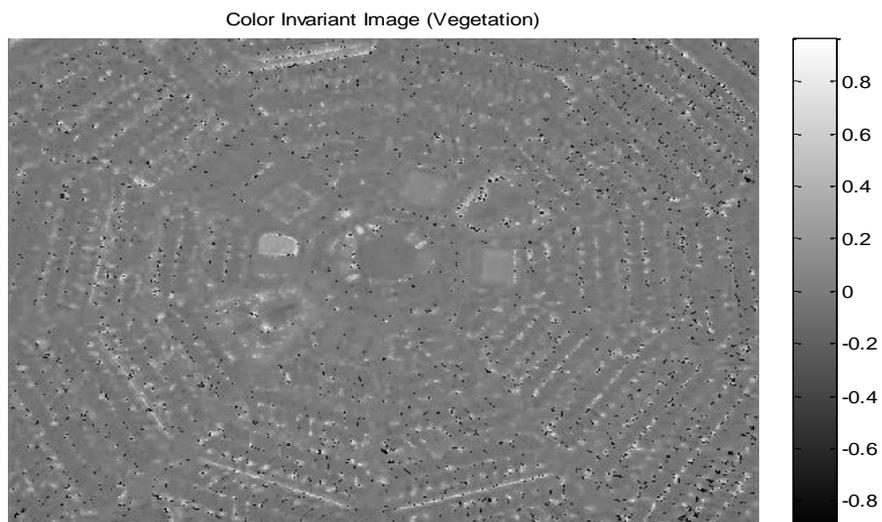


Figure 3-6: Color Invariant Image (Vegetation) - Image I

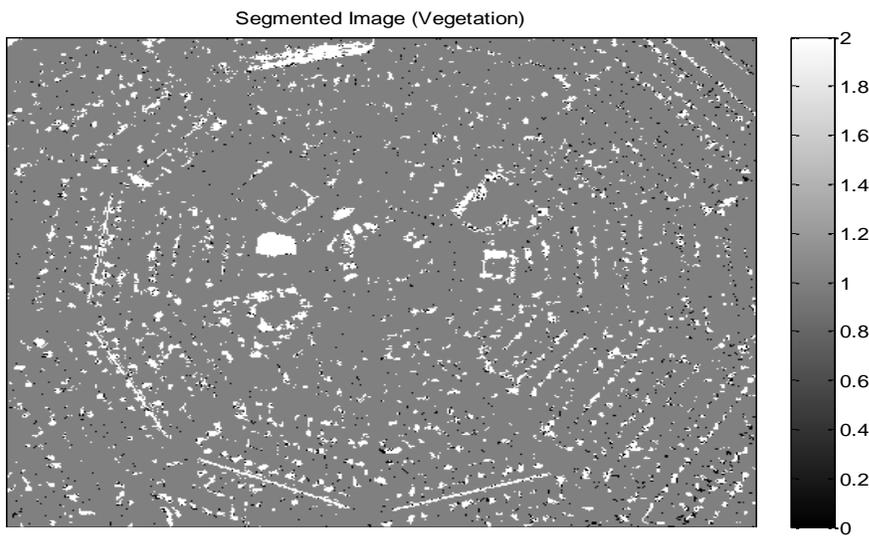


Figure 3-7: Otsu Segmentation Vegetation Image - Image I

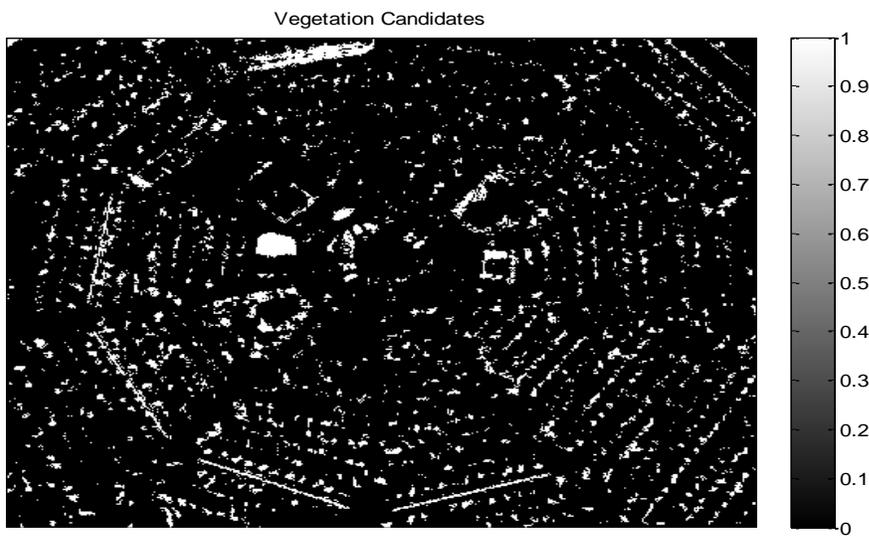


Figure 3-8: Vegetation Candidates - Image I

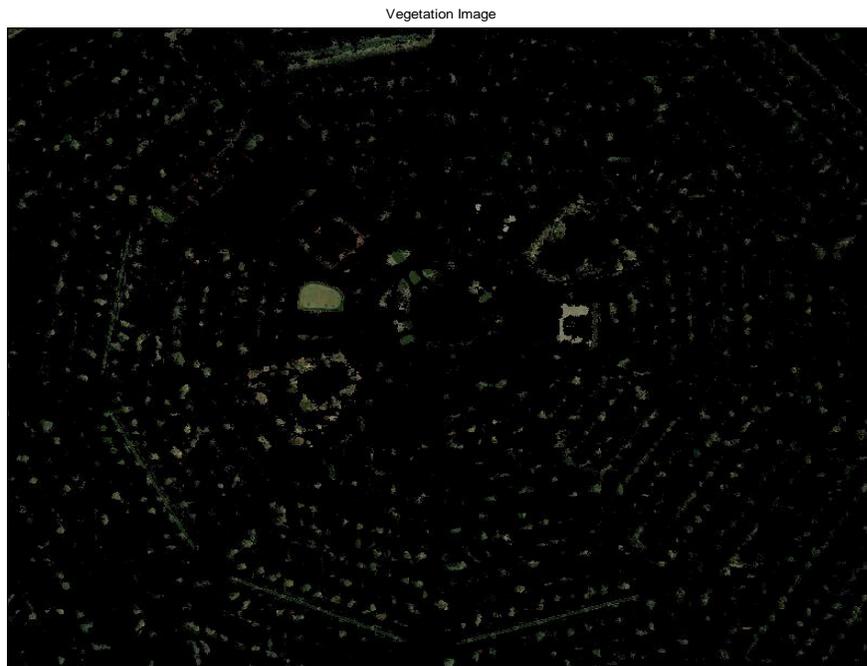


Figure 3-9: RGB Vegetation Image - Image I

For shadowy areas identification, color invariant image (Figure 3-10) is generated after calculating the shadow index (SI), as explained in section 3.2.1 using Equation 3.7. Figures from Figure 3-11 to Figure 3-13 show the steps of shadowy areas detection as mentioned in the vegetation identification process.



Figure 3-10: Invariant Image (Shadows) – Image I

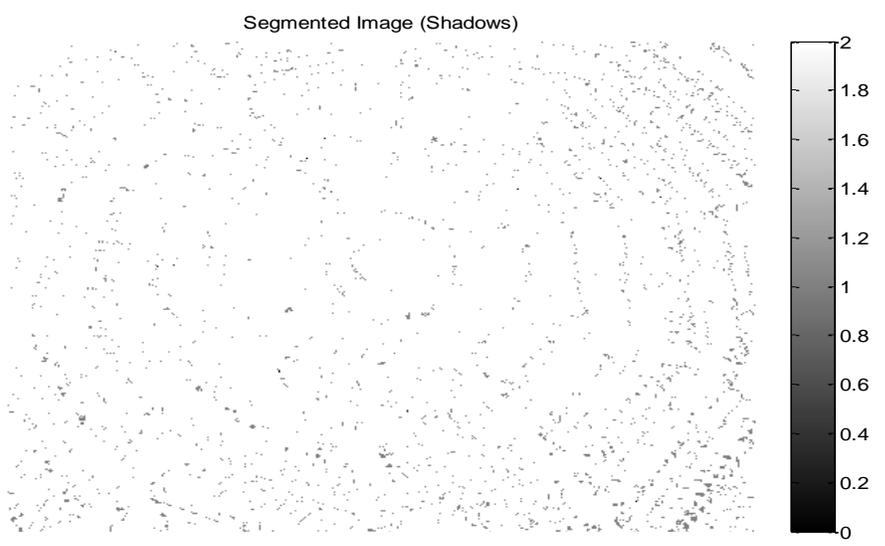


Figure 3-11: Otsu Segmentation Shadow Image – Image I

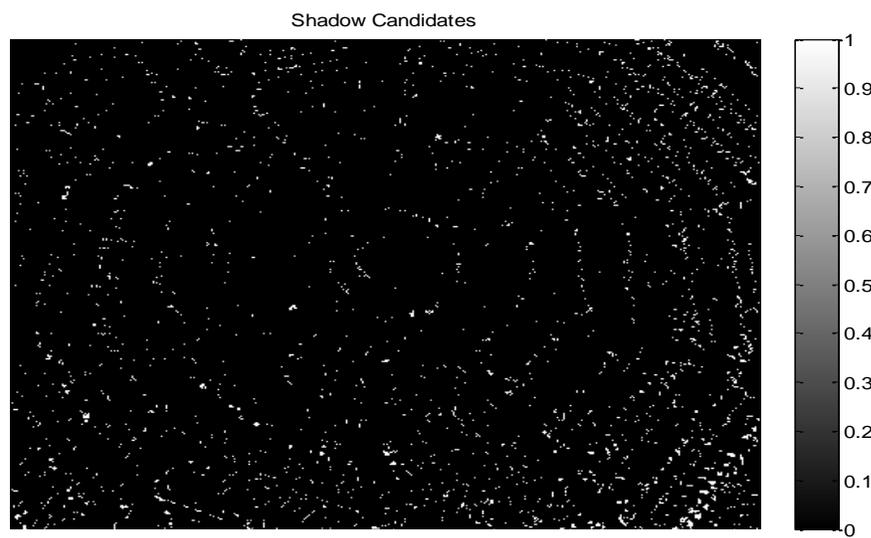


Figure 3-12: Shadow Candidates – Image I

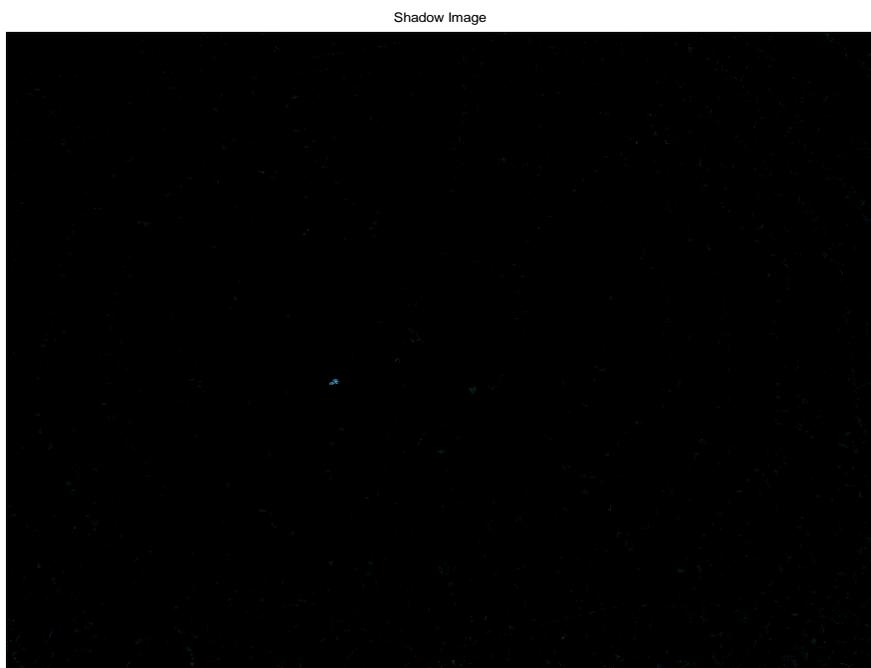


Figure 3-13: RGB Shadow Image – Image I

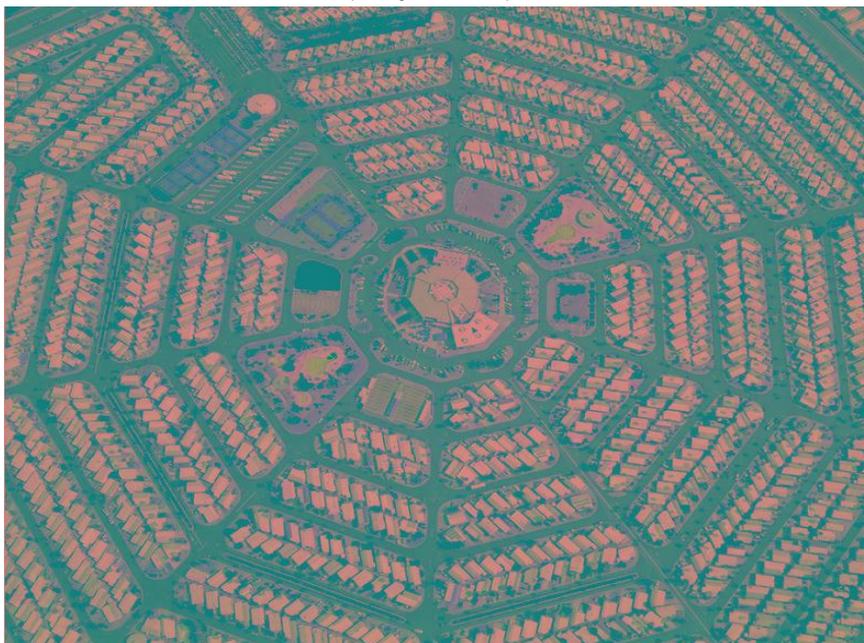
Figure 3-14 shows the output of the first stage of the research algorithm, bounded by the red frame in Figure 3-4, which is an RGB image after removing vegetation and shadowy areas. This image is input data for the second stage of the algorithm, bounded by the green frame in Figure 3-4 [Elhifnawy et al., 2011c].



Figure 3-14: Input Image after Removing Vegetation and Shadows – Image I

The first process of the second stage is the color transformation of the input image into two color spaces, YC_bC_r and HSV . Figure 3-15 shows the image after color transformation to YC_bC_r color space and Figure 3-16 after color transforming to HSV color space.

Input Image YCbCr color space

**Figure 3-15: Input Image in YC_bC_r Color Space – Image I**

Input Image HSV color space

**Figure 3-16: Input Image in HSV Color Space – Image I**

This is followed by applying global color thresholding using three color channels from both of the color spaces. The first color channel is the luminance color channel (Figure 3-17), the first layer of YC_bC_r color space, and the second, saturation (Figure 3-18), and third, hue (Figure 3-19), color channels are extracted from HSV color space.



Figure 3-17: Luminance Color Channel – Image I



Figure 3-18: Saturation Color Channel – Image I



Figure 3-19: Hue Color Channel – Image I

Bong et al. [2009] succeeded in identifying color ranges representing different features from the specific high resolution satellite image, listed in Table 3-1 for luminance, hue and saturation color channels. This research thesis uses these color ranges as a guide or starting point and applies a combination of them for investigating final color thresholding ranges suitable for different color image types to successfully extract specific features (sand and roads). In another words, global color thresholding is applied on all color channels combined and individually to end up with the extraction of road candidates. Equation 3.12 represents the investigated formula for road candidates from luminance color channel only, where RCY represents road candidates from luminance color channel and Y represents the luminance color channel. Figure 3-20 and Figure 3-21 show the road candidates image and the RGB road image.

$$RCY = \begin{cases} 1 & \text{if } Y \leq 100 \\ 0 & \text{Otherwise} \end{cases} \quad 3.12$$

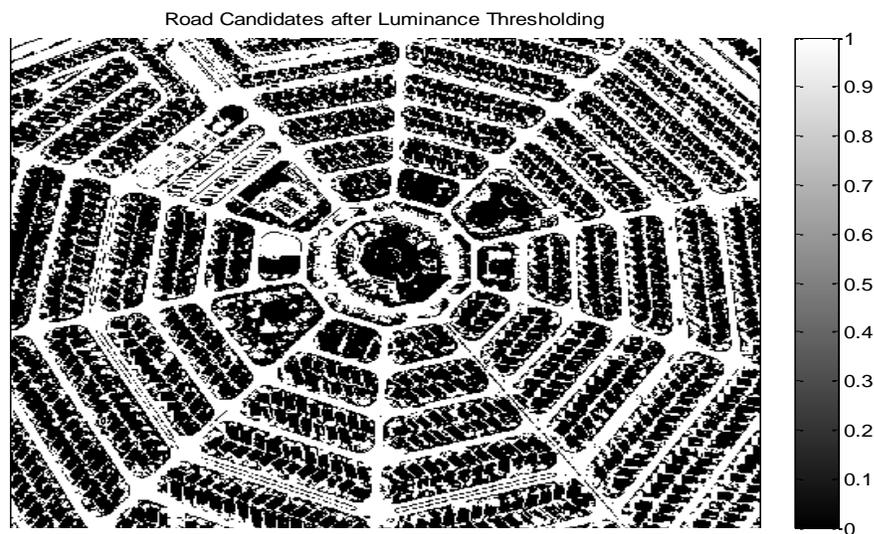


Figure 3-20: Road Candidate Image after Luminance Thresholding - Image I

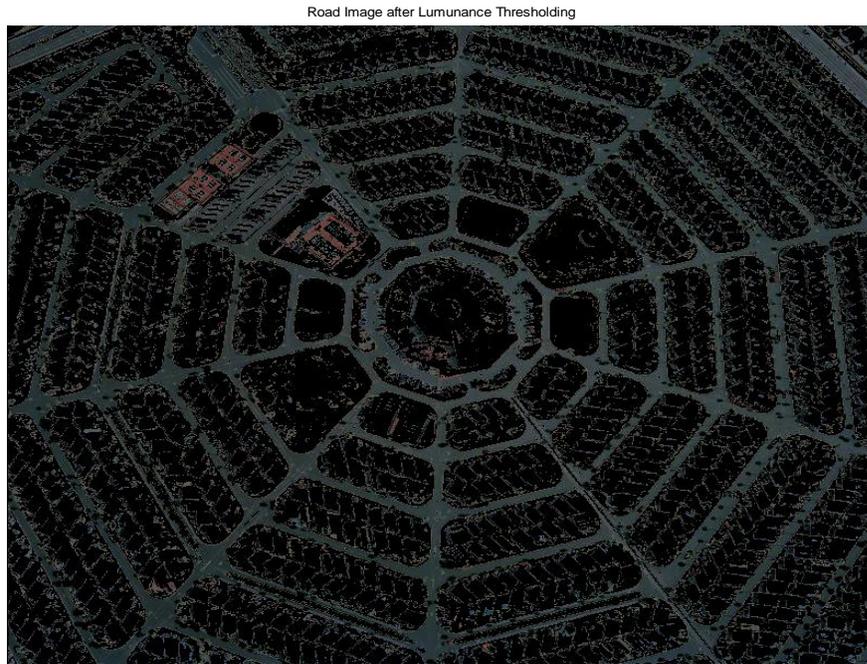


Figure 3-21: RGB Road Image after Luminance Thresholding - Image I

Many areas other than roads still exist in the image and extracted as a road candidates. This thesis took into consideration the luminance color channel as a basis for road extraction from an RGB image. The next step is the removal of any non-road feature from the image of extracted roads from luminance color channel. This is done by applying global color thresholding on hue color channel using Equation 3.13, where $NRCH$ represents the non-road candidates from hue color channel and H represents the hue color channel. Figure 3-22 and Figure 3-23 show the image of these non-road candidates and the same candidate image after fusion with the input RGB image.

$$NRCH = \begin{cases} 1 & \text{if } H < 0.05 \\ 0 & \text{Otherwise} \end{cases} \quad 3.13$$

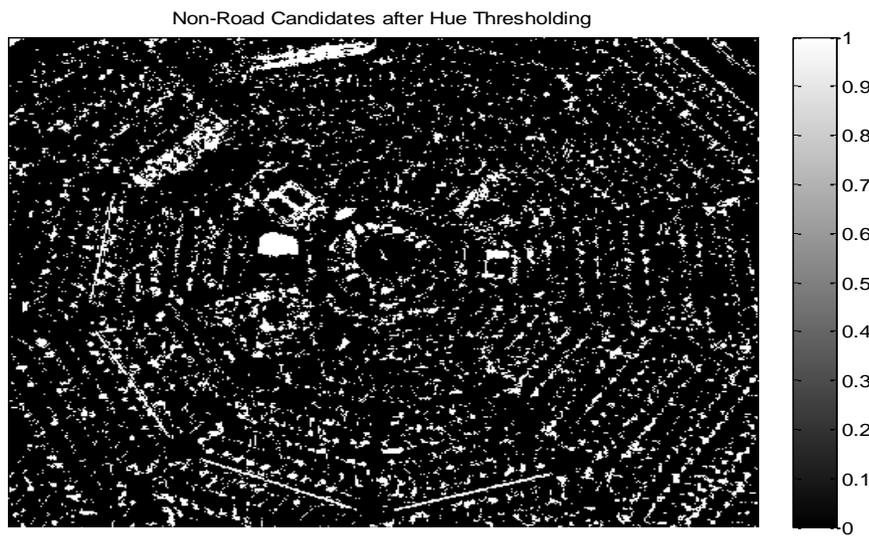


Figure 3-22: Non-Road Candidate Image after Hue Thresholding - Image I

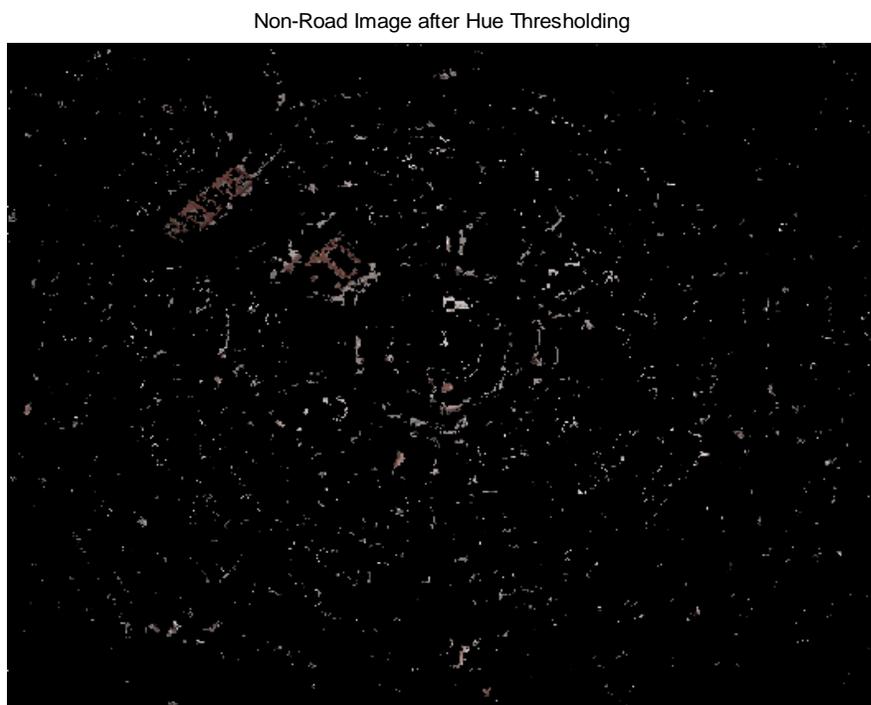


Figure 3-23: Non-Road Image after Hue Thresholding - Image I

After testing the color ranges representing sandy soil from Table 3-1, it is found that sandy areas can be detected from an RGB image after applying global color thresholding on all color channels using the investigated mathematical model shown in Equation 3.14, where SC represents candidates of sand areas and Y , S and H represent luminance, saturation and hue color channels respectively.

$$SC = \begin{cases} 1 & \text{if } 110 \leq Y \leq 160 \text{ and } 0.10 \leq S \leq 0.25 \text{ and } 0.05 \leq H \leq 0.20 \\ 0 & \text{Otherwise} \end{cases} \quad 3.14$$

Figure 3-24 shows sandy soil candidates and Figure 3-25 shows RGB image of sandy areas. After visual inspection and comparison between an RGB image of sandy areas and the original input RGB image, this mathematical model can be used not only for sandy areas identification but also for detecting unhealthy vegetation areas.

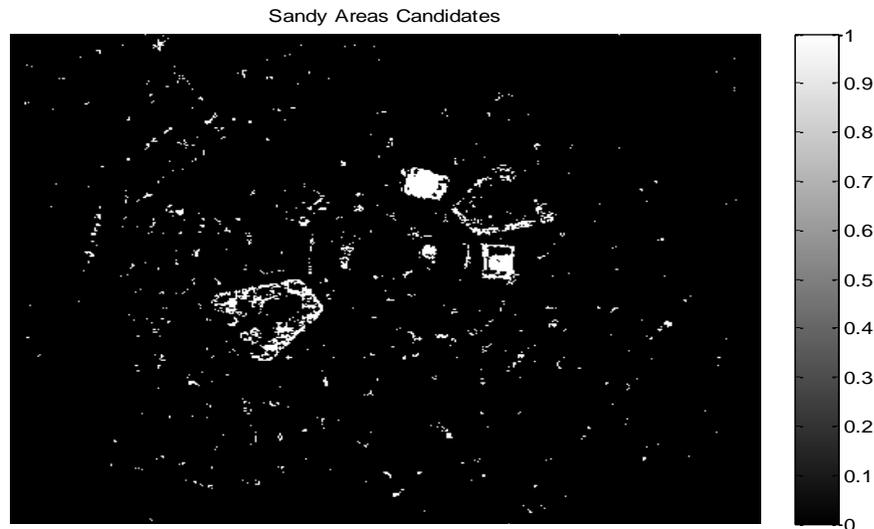


Figure 3-24: Sandy and Unhealthy Vegetation Candidates after All Color Thresholding - Image I



Figure 3-25: RGB Sandy and Unhealthy Vegetation Areas Image after All Color Thresholding - Image I

Now it is possible to extract road candidates from all previous thresholding by taking RCY (Equation 3.12) as main road candidates in an RGB image and eliminating any other candidates that are extracted from $NRCH$ (Equation 3.13) and SC (Equation 3.14). The investigated mathematical model is shown in Equation 3.15, where RC represents the road candidates from RGB image. It equals to RCY and the mathematical model shows the updating process applied on this candidate image. Figure 3-26 shows the extracted road candidates and Figure 3-27 shows the extracted RGB road image.

$$RC = \begin{cases} 0 & \text{if } NRCH \neq 0 \text{ or } SC \neq 0 \\ RCY & \text{Otherwise} \end{cases} \quad 3.15$$

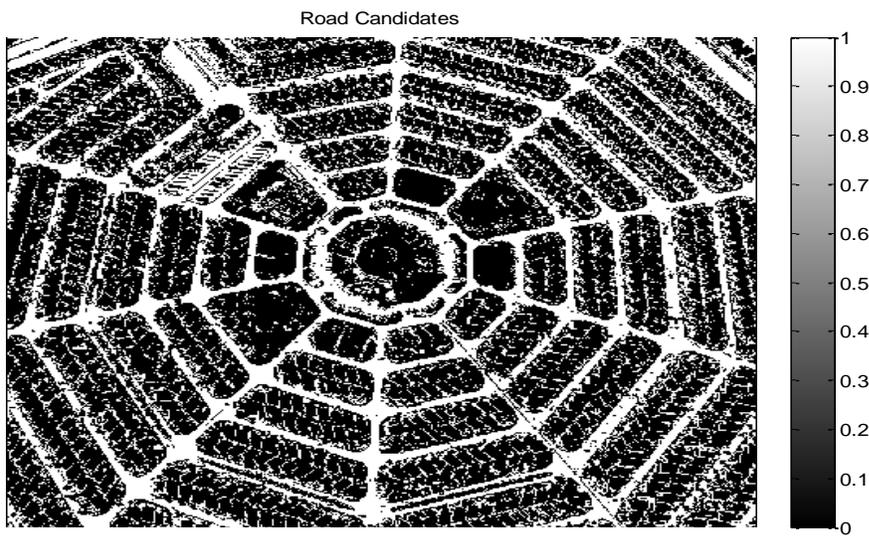


Figure 3-26: Road Candidates from RGB Image - Image I

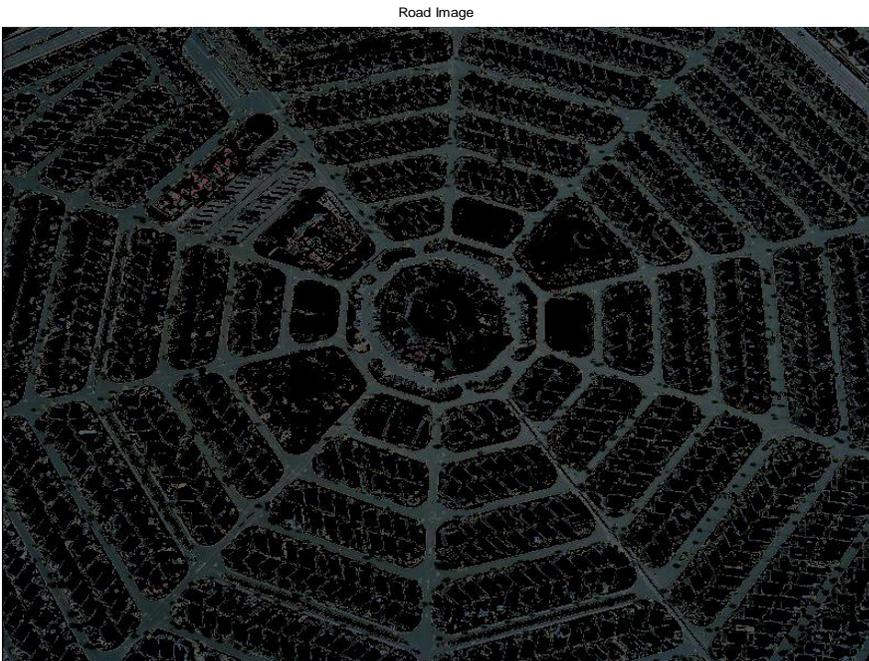


Figure 3-27: RGB Road Image - Image I

For urban areas, any pixel that is not classified as vegetation, shadow, sandy soil or road is classified as a building candidate. So, building candidates are detected after eliminating all successive extracted features, vegetation, shadows, sandy areas and roads, as shown in Equation 3.16, where BC represents building candidates. Figure 3-28 shows a building candidate image and Figure 3-29 shows an RGB building image extracted from the RGB input image.

$$BC = \begin{cases} 0 & \text{if } VI \neq 0 \text{ or } SI \neq 0 \text{ or } SC \neq 0 \text{ or } RC \neq 0 \\ 1 & \text{Otherwise} \end{cases} \quad 3.16$$

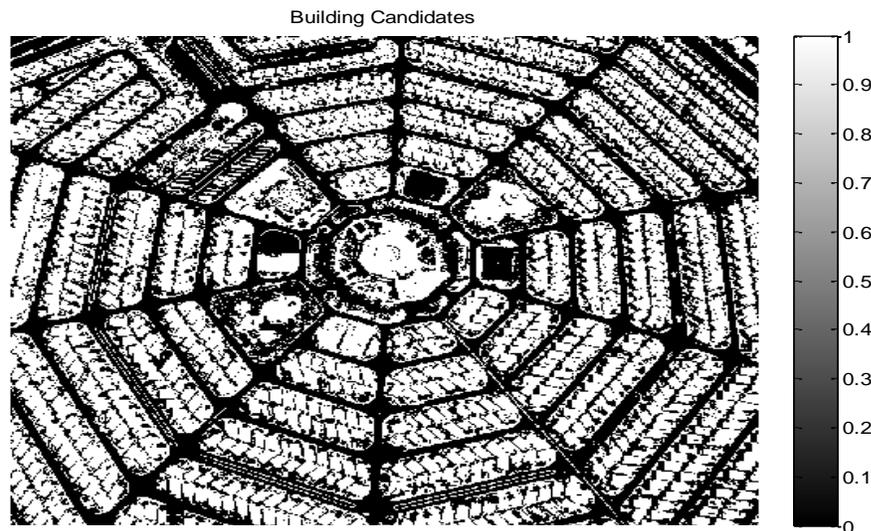


Figure 3-28: Building Candidates from RGB Image - Image I

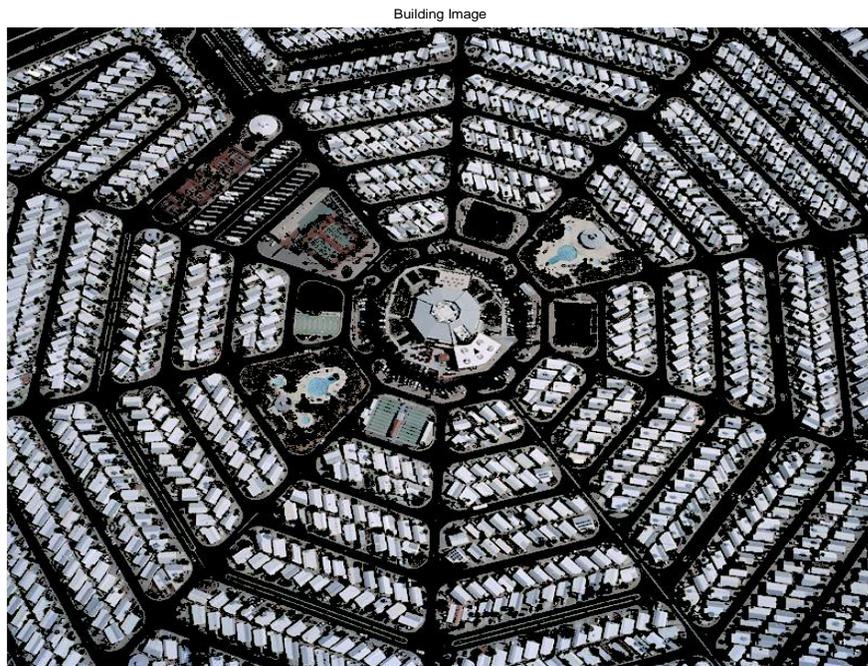


Figure 3-29: RGB Building Image - Image I

Figure 3-30 shows the classification results for this urban area where red color represents buildings, green color represents vegetation areas, black color represents roads, gray color represents shadows and yellow color represents sandy and unhealthy vegetation areas.

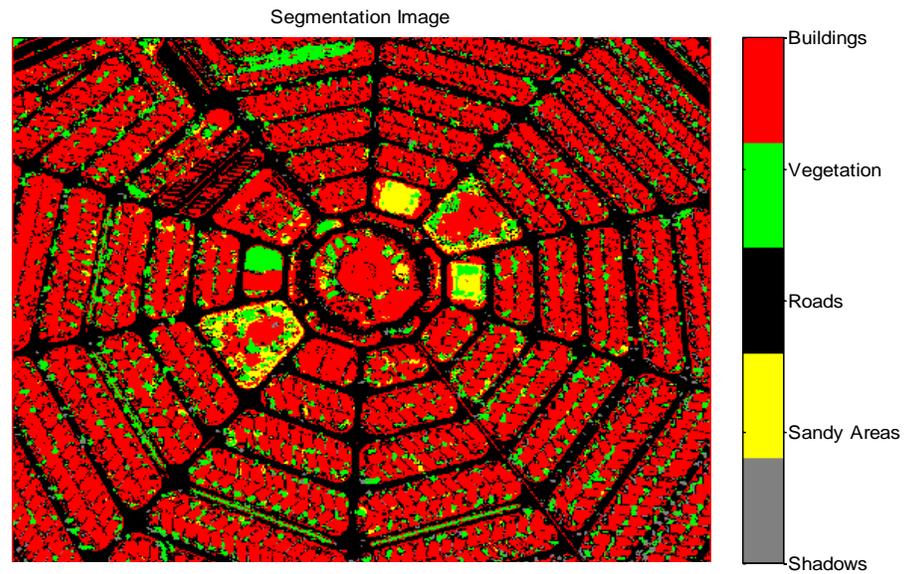


Figure 3-30: Features Classification for Urban Area – Image I

To assess this technique, it is applied on another RGB image, Image II, shown in Figure 3-31. This image shows similarities in texture properties for different features, buildings and roads. Figure 3-32 to Figure 3-38 show feature extraction and classification results following the same steps that are applied on Image I.

Input Image



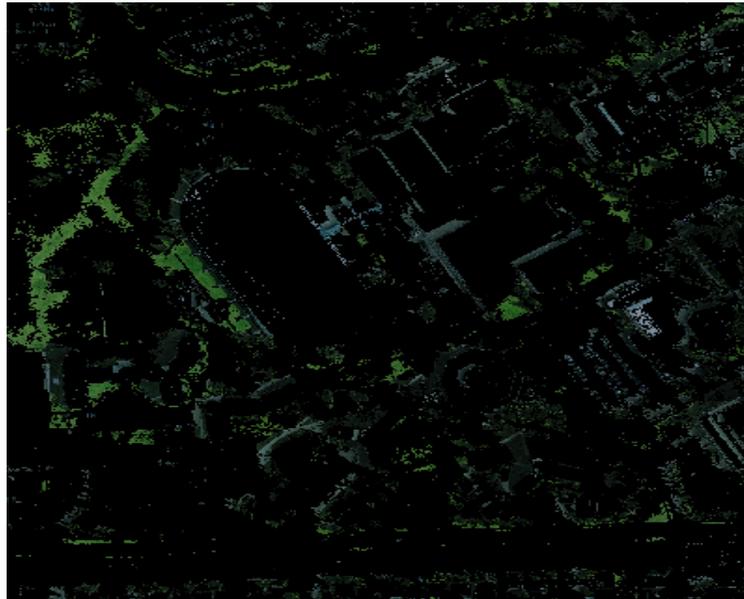
Figure 3-31: RGB Input Image – Image II

Vegetation Image



Figure 3-32: RGB Vegetation Image – Image II

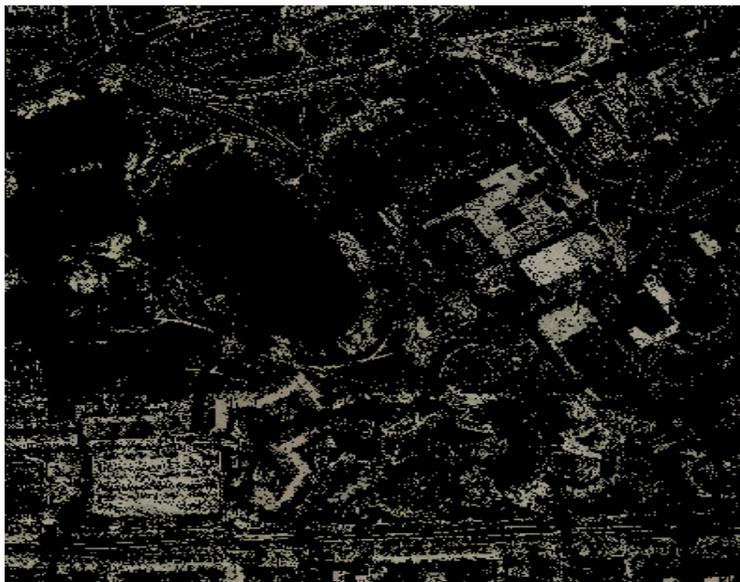
Shadow Image

**Figure 3-33: RGB Shadow Image – Image II**

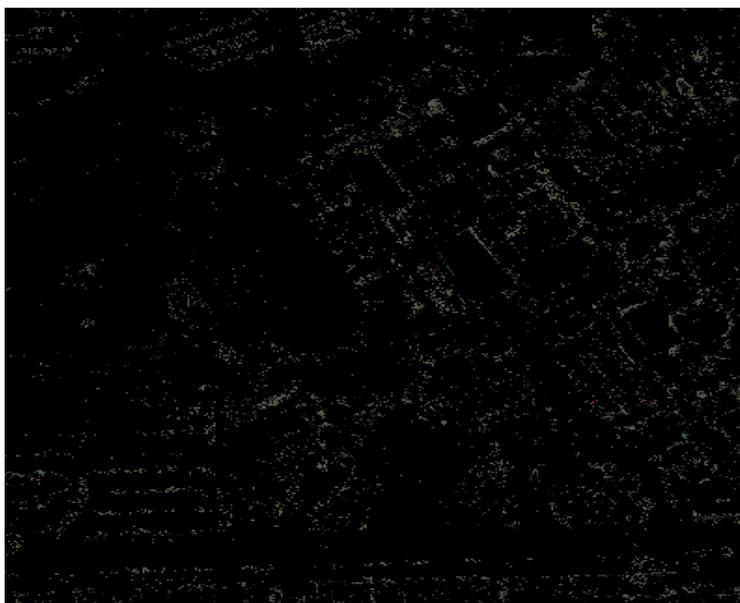
Input image after removing vegetation and shadows

**Figure 3-34: Input Image after Removing Vegetation and Shadows – Image II**

Sandy Areas Image

**Figure 3-35: RGB Sandy and Unhealthy Vegetation Areas Image - Image II**

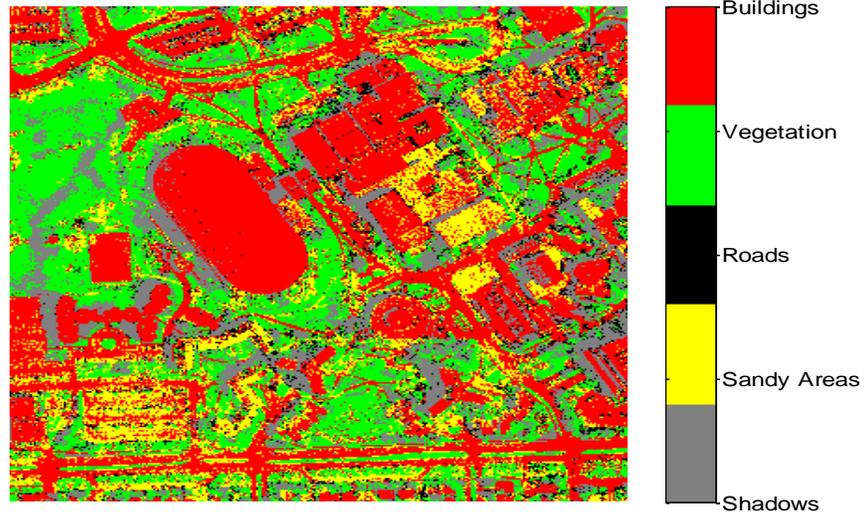
Road Image

**Figure 3-36: RGB Road Image - Image II**

Building Image

**Figure 3-37: RGB Building Image - Image II**

Segmentation Image

**Figure 3-38: Features Classification for Urban Area – Image II**

This technique succeeded in extracting vegetation and shadowy areas, but it could not separate buildings, roads and sandy areas because of the similarity existing in roads, buildings and sandy areas texture properties. This classification problem is the motivation for aiding RGB data with spatial information, as well as height from LIDAR data to be able to separate buildings and roads, features sharing in texture properties, efficiently. The proposed classification technique using combined RGB/LIDAR data will be explained and implemented in more detail in the next chapter.

Chapter Four: Building Extraction from Spatial Data for Efficient Urban Area Classification from Combined LIDAR/RGB Data

As mentioned in the previous chapter, an RGB image is not enough to classify different features which share the same texture properties as roads and buildings. As a result, spatial information such as height from LIDAR must be used to aid an RGB image in effectively separating various features. Since LIDAR data provide full spatial information for ground and non-ground objects it produces accurate height information for different features. These features can be extracted by segmenting LIDAR points related to certain grouping shapes and height information for differentiating between ground and non-ground points.

In the case of RGB images, features that have the same texture properties as buildings and roads cannot be separated as shown in Figure 3-38. However, the use of LIDAR data will help in differentiating buildings from roads using the sudden change in height information between ground and non-ground points (buildings). This chapter focuses on building extraction using spatial information provided by LIDAR data to facilitate the RGB urban classification algorithm illustrated in Chapter 3.

The buildings can be extracted by identifying the positions of height changes for all features in the LIDAR height image. These positions represent the locations of feature (buildings) boundaries. The extraction process is executed based on an image Multi-Resolution Analysis (MRA) operation using wavelet transform. Wavelet, as a MRA tool, is a powerful technique used for de-noising and compressing signals and images by applying several thresholding techniques [Jansen and Bultheel, 1999; Cheng et al., 2004]. Wavelet transform is very effective for detecting feature edges by monitoring the

changing positions of spatial or intensity properties. Wavelet has many applications with the LIDAR system output and can be applied to LIDAR signals, data and images investigated by LIDAR data. In addition, it can de-noise LIDAR signals [Yin and Wang, 2006] and compress data to solve LIDAR storage problems while posting no significant effect on the derived surface elevation model [Pradhan et al., 2005]. Wavelet is also used for filtering and smoothing LIDAR data when separating between high and low frequency signal content. And the surface elevation derived from the LIDAR data can be filtered using discrete wavelet analysis [Tate et al., 2005].

This chapter introduces a literature review of feature extraction algorithms from LIDAR data by discussing the conditions associated with verifying each algorithm and the subsequent advantages and disadvantages for each one. Furthermore, the mathematical background of wavelet transform will be explained in details. At the end of this chapter, the process of combining RGB and LIDAR data resulting in efficient urban area classification results will be introduced and tested.

4.1 Literature Review

Many algorithms have been used when performing feature extraction from LIDAR data, such as the research works of Nardinocchi et al. [2001], Rottensteiner and Briese [2002], Schwalbe et al. [2005] and Li et al. [2007a]. Nardinocchi et al [2001] used a last return of LIDAR pulses to generate a LIDAR image with a 1-m x 1-m grid. Terrain points were removed by analyzing residuals through spline interpolation, while the connected height regions were grouped using the growing technique. Roof plane segmentation was performed using the RANdom SAMple Consensus (RANSAC) algorithm which seeks to fit data on planes and then used to build contour lines. Once a 3-D vector representation

of these planes has been derived, the roof can then be obtained. External and internal nodes can be derived by establishing the relationship between every two adjacent planes. Least-squares estimation is then used to produce a 3-D approximation for the building models. The roof planes are modeled in 3-D space given the orthogonality constraints of the boundary lines. However, this algorithm is only suitable for orthogonally edged roof shapes.

Combined 2-D/3-D Hough transform technique is introduced for detecting planar patches [Vosselman and Sander, 2001]. While 3-D Hough transform is used for line detection to form planar patches from irregular LIDAR point clouds, 2-D Hough transform is applied after projecting the points of lines on vertical planes to get the roof faces. The splitting and merging technique for the segmented planes is applied when detecting two types of lines: the intersected lines and the height jump lines. The 3-D model is then formed by combining segments related to the same planar faces. This technique is also suitable for planar surfaces with rectangular shapes.

Rottensteiner and Briese [2002] generated a polyhedral building model from LIDAR data. This algorithm originated from a point classification applied when identifying off-terrain object points. Tree removal is achieved by evaluating terrain roughness based on texture homogeneity. An opening morphological filter is applied when removing vegetation areas that are connected to the buildings. Seed growing is then added to connect adjacent planar patches while a label image is introduced for grouping these planar surfaces and extracting polygon borders to create the final polyhedral models. On one hand, this algorithm consists of a variety of complex methods and on the other hand, it is appropriate for extracting objects with rectangular roofs only.

Rottensteiner et al. [2003] succeeded in extracting buildings from three different data sets. The first two are from LIDAR data, Digital Surface Models (DSM) for the first pulse and the last pulse. All DSM's are gridded using linear prediction. The third data set is the Normalized Difference Vegetation Index (NDVI) and is produced by infrared and green channels from a multi-spectral image. A Digital Terrain Model (DTM) is generated after applying a morphological gray scale opening algorithm on the DSM for the last pulse. The opening operation is executed in a hierarchical manner with carefully selected structuring elements to completely remove all non-ground objects. The height difference between the DSM and DTM with specific threshold is used to detect a building candidate and the morphological opening is applied with a small structuring element (3x3) to remove small objects. Segmented buildings image is produced once tree removal has been accomplished using NDVI. The drawback of this technique is that additional data such as multi-spectral sources are needed to be able to remove vegetation pixels that are extracted as building candidates and first and last LIDAR pulses must be available.

Sohn and Dowman [2003] used combined satellite image and LIDAR data for building extraction. The first step is the DTM generation from the LIDAR data and is based on fragmenting the LIDAR points to Planar Terrain Surfaces (PTS) which are triangular in shape with three vertices. The points that have continuity and homogeneity criterions are classified as on-terrain points. This analysis is executed on a LIDAR Digital Elevation Model (DEM) in a hierarchical manner from coarse to fine using the Delaunay Triangulation method for defragmenting PTS. The final DTM is generated from finer PTS. By using the height threshold from the DTM the off-terrain points are identified. The tree removal is performed with NDVI that can be identified by IKONOS imagery

using red and near infrared channels. The boundary lines are also extracted from an IKONOS image. LIDAR off-terrain points are used with the help of extracted lines to investigate points for building boundaries. Final building boundaries are extracted based on specific length and angle thresholding, which are referred to intensity lines. Virtual lines are generated from intensity lines because they form virtual boxes. These boxes are expanded in two directions to contain building points, without terrain points and have what are termed virtual lines. The geometric lines projected afterwards onto the image space are adjusted using gradient weighted least-squares method. The building polygon is generated by the intersection between intensity and virtual lines. Building polygon grouping is performed using the binary space partitioning method. This technique is complicated because many segmenting, partitioning and vegetation removing methods are used. This technique failed to extracting buildings from only LIDAR data, but succeeded when aided by a multi-spectral satellite image used for initial building boundary identification and producing NDVI image for vegetation removal. This technique is suitable for extracting building roofs with rectangular shape only.

Schwalbe et al. [2005] constructed 3-D building models using neighboring planes formed by line tracking to respective boundary points after projecting LIDAR point clouds to a 2-D orthogonal projection aided by 2-D Geographic Information System (GIS) data to identify plane orientation. The reconstruction algorithm consists of many steps; the first step is the elimination of the ground points using height histogram analysis followed by determination of the primary roof orientation by height histogram bin analysis or ground plane analysis to get the building azimuth. The segmented point clouds are projected to the XZ plane after rotating by the building azimuth and to the YZ plane after adding 90°

to the building azimuth. The roof planes appear as lines in these projections and roof planes are formed from the detected lines, and are ultimately combined to form roof structure. Walls are generated by projecting the roofs on the ground. A final building model is reconstructed after these walls are added to the roof structures. This technique used LIDAR data for detecting solely off-terrain points, not building edges or roofs. The extracted building is mainly based on GIS data for identifying roof and wall orientations. Kim et al. [2007] investigated a new approach for LIDAR data segmentation based on three main requirements, neighbourhood definition, deriving attributes for the neighbourhood points and clustering neighbourhood points with similar attributes. Points are defined as neighbourhoods if they are located on the same surface. Neighbourhood points located on the same surface are detected using the adaptive cylinder method. The magnitude of the surface normal from the origin to the plane of neighbourhood points is considered a derived attribute. Two origins are used to provide two attributes for each surface and the line joining two origins has to form an angle of 45° with the horizontal plane to minimize the possibility of segmentation ambiguities. The neighbouring points with the same attribute are then aggregated and the boundary for each group of points is detected and all points located within form one cluster. But the boundaries for all segmented clusters are irregular. The Douglas-Peucker algorithm is used to simplify the detected boundaries. Boundary points are fitted using least square to form four boundary lines. The intersection between these boundary lines forms the planar patch corners. This approach is complicated and involves mostly segment LIDAR data in planar patches. Although this is suitable for building extraction or any feature with planar surface-top

and regular boundaries it cannot be considered a general segmentation technique for extracting all features from LIDAR data.

Kim et al. [2008] investigated a new algorithm for generating a 3-D wire frame for building extraction from LIDAR data. This algorithm consists of four steps. First, non-ground points are identified by point classification and performed based on visibility analysis. This analysis is applied by looking for points which cause occlusion in the perspective projection aerial image in the area of interest. The second step, a generation of building hypothesis, is achieved by studying the spatial relationship between points to detect those related to the same planar surface. The third step is the segmentation of planar patches, a process consisting of three sub-steps; neighborhood definition using adaptive cylinder definition technique, attribute derivation and clustering based on attribute similarity [Kim et al., 2007]. The modified convex hull approach is used to detect the boundary for all planar patches. Finally, planar boundary refinement is achieved. By knowing the average elevation of points around the building, a 3-D wire frame is investigated to represent buildings in the area of study. This technique succeeded in extracting non-terrain LIDAR points after aiding LIDAR data with an aerial image. This technique is complicated and it is unsuitable for extracting features other than those with rectangular shapes.

Zeng [2008] used a planar filter on the point clouds to extract building surfaces by employing specific height, length and delineation thresholds. This filter studied the relationship between all adjacent points to detect features surfaces. Feature points are obtained after the ground points from LIDAR point clouds have been removed. Building-top points are classified as points with a continuous distribution. Finally, the 3-D building

points (surface top and walls) are extracted. This technique assumed that, non-ground points form areas smaller than that from ground points. This condition is not suitable for crowded urban areas and as a result this technique cannot be generalized for all study areas, rural and urban. However, this technique successfully extracted features with planar surface only but cannot extract other non-ground objects such as trees.

On one hand, the techniques previously mentioned concentrate more on objects extraction with regular boundaries, continuous distributed point clouds and/or uniformly sloped surfaces while on the other hand, they are unsuitable for objects with irregular boundaries such as trees and poles. Moreover, they will not work with circular or ellipsoidal surfaces structures. This research thesis introduces a technique for extracting objects with different shapes and boundary regularities using the power of wavelet transform for edge detection applications.

4.2 Literature Review for Wavelet Feature Extraction from LIDAR Data - Efficiency and Limitations

Feature extraction can rely mainly on height properties for different features, which help in the separation between ground and non-ground points. Non-ground features can be identified by detecting the positions of non-continuous height properties or the positions of a sudden change in height among different features. This feature extraction trend, based on sudden change in height, with the use of wavelet localization property is the motivation for analysing LIDAR height images by wavelet transform to detect locations of height differences and consequently boundaries of the feature corresponding to this sudden change in height. The extraction technique is not sensitive to the shape, boundary regularity or point cloud distribution.

There are many approaches and implementations for feature extraction using wavelet transform such as Vu and Tokunaga [2002], Wang and Hsu [2006] and Falkowski et al. [2006]. Vu and Tokunaga [2001] applied wavelet transform on LIDAR data to monitor feature properties through different scale ranges. The first step is gridding LIDAR data to regular representation forming the LIDAR image based on height information called a range image. This image is analysed using wavelet transform at different scales. The smoothed image (low frequency) and edge image (high frequency) are extracted at each scale. The clustering of LIDAR points is applied using the K-mean clustering method [Kanungo et al., 2002], which aims to partition a number of observations into a number of clusters, in which each observation belongs to the cluster with the nearest mean. The resulting clusters are formed in three main categories. The first category is for ground points, the second, for high rise buildings and the third for low buildings, trees and others. The final segmentation results are formed by the combination between clustering LIDAR points and features extracted after wavelet analysis. The segmented points are classified into ground points used to produce a DTM and object points used to produce a DSM for the area of study.

Vu and Tokunaga [2002] to produce a clustering technique based on wavelet smoothing LIDAR images at different successive scales corresponding to different frequencies. The first step is the production of LIDAR image after gridding the LIDAR data to regular representation resulting in a range image. Wavelet transform is then applied on a LIDAR image to produce four successive scales. The last smoothed image represents the higher smoothed objects in the LIDAR image and are considered buildings. The buildings points are extracted by masking the smoothed image to the original LIDAR image and are used

to reconstruct 3-D building models for the area of interest. This research was developed by [Vu and Tokunaga, 2004] to produce a modified clustering technique based also on wavelet analysis, which is referred to as Airborne Laser Scanner Wavelet (ALSwave) for LIDAR data. The first step of the new approach is the presentation of LIDAR data over a Triangular Irregular Network (TIN) using the planar interpolation technique, which minimizes the loss of information and gives an accurate interpolated image. The second step is the application of wavelet analysis. Wavelet analysis is executed using a cubic B-spline wavelet function with compact support. The wavelet multi-resolution framework is prepared based on a trous algorithm. The fuzzy edge points are detected along with ground points and object edge points. The Delaunay neighbors of points are used for classifying fuzzy edge points. Global thresholding and local thresholding are applied with the Delaunay neighbors to detect heights points. The ground points can be detected, and then used for generating the ground surface. This research is complicated and used a combination of wavelet transform, fuzzy edge pixels and Delaunay neighboring points to separate object points from ground points to produce the DSM and the DTM of the area of study. This research thesis will introduce a standalone wavelet technique for detecting object boundaries directly from LIDAR data without requiring any other technique to classify LIDAR points to ground and object points.

Wang and Hsu [2006] detected building boundaries using a spatial edge detection technique and analysing this edge image using a wavelet transform. The LIDAR point clouds are interpolated to a regular grid using the Kriging interpolation method. The edges are detected after applying a Canny edge detection technique. The resulting image contained edges for all non-ground points including trees and buildings. Wavelet

transform is applied on the resulting image up to three levels of decomposition to detect finer edges which represent building boundaries. Because of adjacent buildings or trees that are located close to the extracted buildings, these boundaries do not form closed lines or curves. A morphological closing operation is then applied to solve this issue and close the building boundaries and a 3-D building model is reconstructed from the extracted lines. This research paper used wavelet transform as an auxiliary tool to detect finer edges from the image of building edges produced by the Canny edge detection technique. This research thesis will detect building edges using only a wavelet transform without the aid of any other edge detection technique.

Falkowski et al. [2006] succeeded in extracting conifer trees from LIDAR data after generating a 2-m DEM and Canopy Height Model (CHM) using the natural neighbour interpolation method through ArcGIS software. The first tree detection technique used in this research paper is the Variable Window Filter (VWF). This filter is based on detecting local maxima (tree height) through the CHM within a specific window size and is required to have a prior knowledge of tree height and crown diameter. The major disadvantage of this technique is the estimation of the fixed relation between height and crown diameter. The second tree detection technique introduced by Falkowski et al. [2006] is Spatial Wavelet Analysis (SWA). Wavelet analysis is executed using the 2-D Mexican Hat wavelet function. 2-D Mexican Hat wavelet function is suitable when extracting this specific tree type since its shape looks like the shape of a conifer tree. This filter doesn't require a prior knowledge of tree height or crown diameter. The wavelet function has three main components, wavelet size, location and a goodness of fit metric. When the tree is fitted, the wavelet function, the crown diameter (wavelet size) and the

tree location are all recorded. The tree height is estimated by fitting the scaled wavelet function with the tree size. This technique explains the efficiency of wavelet analysis for extracting conifer trees, but is unsuitable for other types of trees or for building extraction.

Separation of ground and non-ground points in hilly terrain is not straight forward, especially when the slope of the terrain is not steep and uniform. Bartels and Wei [2006] overcame this problem by using wavelet analysis and extracted buildings from hilly terrain. The first step is gridding of LIDAR data to regularly spaced representation. Second order Daubechies wavelet filter is applied on the LIDAR image resulting in sub-images representing an approximation and three details. The approximation represents the ground points with regular distribution. This regularly distribution is formed by points with low degrees of inclination or flat areas. The details indicate locations of sharp inclination between adjacent points. The object boundary image is obtained by reconstructing wavelet coefficients using a 2-D Discrete Wavelet Transform (DWT) after replacing the approximation image with zeros.

Within the same year, this research was broadened in order to investigate a segmentation technique based on wavelet transform [Wei and Bartels, 2006]. The extended approach is composed of three main steps. The first step is the representation of LIDAR data as a gray scale image. This image is analysed using Gabor wavelet function to three scales with four different orientations using different windows. Finally, a statistical analysis is done and the standard deviation and the mean are calculated for each window. Windows with low standard deviation and low mean, and windows with high standard deviation and low mean are considered as ground areas. Windows with low standard deviation and

high mean are considered as buildings and vegetation areas. This technique succeeded in segmenting flat, ground and non-ground areas, but a problem arises in identifying buildings with large flat roofs, as they are classified as ground areas. Consequently, it is important to use additional classification information to segment all features in the input image efficiently.

The aforementioned feature extraction techniques based on wavelet analysis used additional aids to extract features such as Canny edge detection [Wang and Hsu, 2006], wavelet analysis on a trous algorithm method [Vu and Tokunaga, 2002] and necessary representation of LIDAR data in TIN format [Vu and Tokunaga, 2004; Wang and Hsu, 2006]. In this research thesis, an independent wavelet based building extraction technique, without any aid, is used in combination with a very simple and efficient interpolation method the Nearest Neighbour (NN) interpolation technique which maintains the spatial information of feature edges. The main advantage of the proposed approach is its simplicity and efficiency compared to the other previously mentioned approaches. The methodology and wavelet implementation will be described in detail in the following sections.

4.3 Multi-Resolution Analysis for Building Extraction from LIDAR Data

A Multi-Resolution Analysis (MRA) using wavelet transform is used for building extraction from the height information produced by the LIDAR system. The proposed MRA implementation is executed so as to benefit from the sudden change in height at the edges of all buildings. Because of the wavelet localization properties and its ability to detect singularities and sudden changes (high frequency), wavelet is used for building edge identification and extraction. The following section will introduce the mathematical

background of the wavelet transform for feature extraction without requiring further assistance.

4.3.1 *Wavelet Analysis for LIDAR Data Applications*

The proposed feature extraction technique from LIDAR data is based on wavelet analysis, so before introducing the wavelet implementation for feature extraction it is important to mention why wavelet is a suitable tool for this application. The following sub-sections explain the wavelet advantages compared to Fourier transform, which was a well-known signal processing technique for numerous applications for a significant amount of time [Elhabiby, 2007].

4.3.2 *Wavelets versus Fourier Transform*

In the 19th century, French mathematician Joseph Fourier succeeded in analyzing any periodic function as an infinite sum of periodic complex sinusoidal functions [Graps, 1995; Elhabiby, 2007]. Equation 4.1 [Keller, 2004] shows a Fourier series representation of a function over a period of $2T$. It shows that Fourier can represent any function f as a superposition of periodic functions with different wavelengths λ_p (Equation 4.2) with amplitude A_p (Equation 4.3).

$$f(x) = \sum_{n \in \mathbb{Z}} A_p e^{i \frac{2\pi}{2T} p x} \quad 4.1$$

$$\lambda_p = \frac{2\pi}{2pT} \quad 4.2$$

$$A_p = \frac{1}{2T} \int_{-T}^T f(x) e^{-i \frac{2\pi}{2T} p x} dx \quad 4.3$$

Fourier transform can express any non-periodic function when $T \xrightarrow{\text{yields}} \infty$ in a frequency domain as shown in Equation 4.4. The transition from f to A is called Fourier transformation (Equation 4.5), where A is an amplitude spectrum of f [Keller, 2004].

$$f(x) = \int_{-\infty}^{\infty} A(\omega) e^{i\omega x} d\omega \quad 4.4$$

$$A(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx \quad 4.5$$

The first disadvantage of the Fourier transform is deduced from these equations because an exponential form can be represented as a summation of sine and cosine functions (Equation 4.6), so Fourier transform is based only on two base functions.

$$e^{i\omega x} = \cos(\omega x) + i \sin(\omega x) \quad 4.6$$

Fourier transform can analyse any function to oscillations of different frequencies as mentioned before, but it cannot detect positions or occurrence time of each frequency or any change in frequencies. This is called lack of space localization and it represents the second disadvantage of Fourier transform [Keller, 2004]. To understand the lack of frequency localization problem, Figure 4-1 shows the Fourier transform of height image from LIDAR data shown in Figure 2-9. Due to a high dynamic range of Fourier spectrum, it is impossible to detect different frequencies in the image, but it becomes possible after applying an image enhancement using the log transform (Equation 4.7) [Gonzalez et al., 2004] as shown in Figure 4-2.

$$IS(r) = l * \log(1 + r) \quad 4.7$$

where: r = intensity value of pixel

IS = new intensity value of intensity r

l = Log transform constant

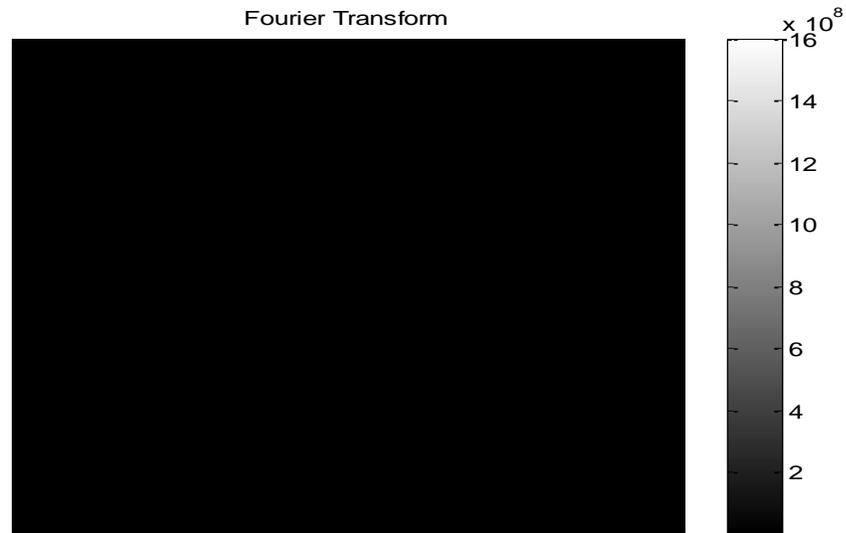


Figure 4-1: Fourier Transform Image

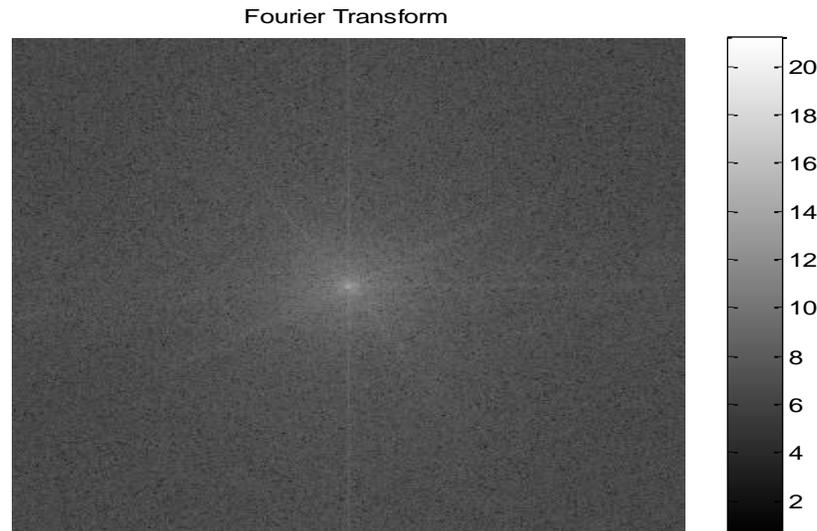


Figure 4-2: After Log Transformation

Fourier transform detects all different frequencies in the image, but it is not possible to detect or monitor the exact positions of certain frequency or positions of transient from any frequency to another because Fourier transform analyses a signal over the whole space span of the signal.

Short time Fourier transform, as a solution to the lack of space localization problem, is a localization technique that is applied by masking an input signal by a window function with compact support. This function has a specific frequency and space resolution, so it is called windowed Fourier transform technique. Equation 4.8 represents a mathematical model for windowed Fourier transform, where \mathbb{G} is a transformation function, g a window function and t a transition parameter [Keller, 2004; Elhabiby, 2007].

$$\mathbb{G}\{f\}(\omega, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)g(x - t)e^{-i\omega x} dx \quad 4.8$$

An input signal is divided into sub-signals with same space/frequency window; and then each window is analysed separately using a masking window function [Graps, 1995]. The main disadvantage of windowed Fourier transform is that, the window function has a constant space/frequency resolution and becomes unsuitable for different frequencies that exist in an input signal. High frequency resolution with low space resolution is suitable for localizing high frequencies and low frequency resolution with high space resolution is suitable for localizing low frequencies. So the space/frequency window for masking function must be coupled, in other words window size must change according to required frequency resolution; this is the idea of wavelet invention.

Wavelet transform overcomes the disadvantages of Fourier transform because it has many different base functions that give it flexibility to analyse and represent any signal or functions better than Fourier transform. The other advantage of wavelet transform is a compact support property for wavelet function that means it is bounded by limited range that can be tuned with respect to the space/frequency resolution. This property makes wavelet transform a powerful localization tool.

Figure 4-3 shows a sample of different wavelet functions. It can be based on a mathematical model such as HAAR or based on derived coefficients as Daubechies functions. Wavelet analysis has two main types, continuous and discrete. The selected analysis type is based on the nature of the input signal, wavelet function, and the application. This research thesis tested several wavelet base functions for building extraction. It is found that Symlets (sym4) is the most suitable wavelet base function for

building extraction from LIDAR data by analysing the LIDAR height image up to two levels of decompositions using 2-D discrete wavelet transform.

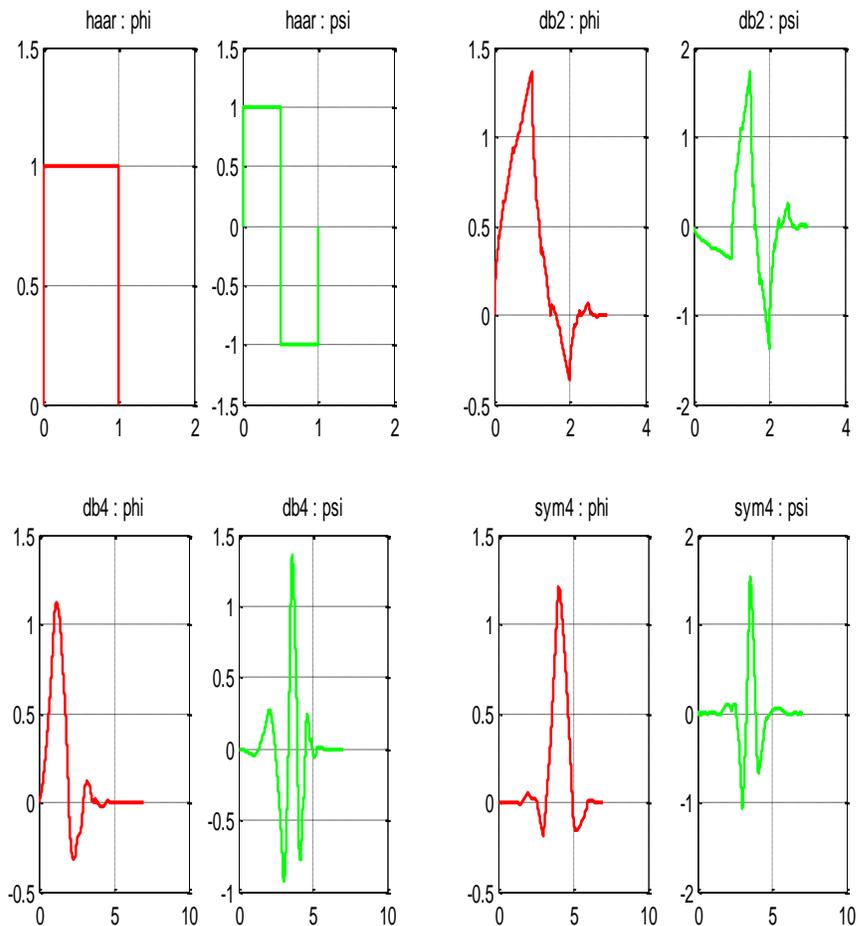


Figure 4-3: Samples of Different Wavelet Base Functions

4.3.3 *One Dimensional Continuous Wavelet Analysis*

Continuous Wavelet Transform (CWT) is the transform of continuous function with continuous variables that form a highly redundant function [Gonzalez et al., 2004]. When using continuous wavelet transform there are three properties that are taken into consideration. The first property is the high redundancy occurring because the wavelet transform is calculated by continuously shifting a continuously scalable function over a

signal and calculating the correlation between them. This tends to obtain the wavelet coefficients with high redundancy. The second property is the infinite number of wavelets that can be used. The high redundancy and infinite number of created wavelets make the calculation of the transformation highly time-consuming [Keller, 2004].

The transformation variables are the translation and scale parameters as shown in Equation 4.9 that represents a wavelet transformation function [Gonzalez et al., 2004].

$$W_{\psi}(m, n) = \int_{-\infty}^{\infty} f(x) \psi_{m,n}(x) dx \quad 4.9$$

where: m = scale parameter

n = translation parameter

4.3.4 *One Dimensional Discrete Wavelet Analysis*

When a variable x and its values $f(x)$ are all finite and discrete quantities, the wavelet function is discretized into a grid with dimensions (λ_o^m, nt_o) , where m is the scale parameter and n the translation parameter. This grid is not regular, and it is changed related to the scale and level of decomposition. Equation 4.10 outlines the mathematical model of Discrete Wavelet Transform (DWT) [Keller, 2004].

$$\{\psi_{m,n}^{(\lambda_o, t_o)}(x) = \lambda_o^{-m/2} \psi(\lambda_o^{-m} x - nt_o) \mid m, n \in \mathbb{Z}\} \quad 4.10$$

where: λ_o = scale

t_o = translation

A signal or function $f(x)$ can be analysed using an expansion function φ to obtain the function approximations. In wavelet analysis there is another function, ψ , that is used to

obtain the differences between adjacent approximations. φ is a scaling function and ψ a wavelet function. So, wavelet transform analyses the signal into two main parts; a low frequency part called approximation and a high frequency part called detail. Equation 4.11 shows the mathematical model for calculating DWT coefficients of function $f(x)$ [Gonzalez et al., 2004].

$$W_{\varphi}(m, n) = \frac{1}{\sqrt{M}} \sum_x f(x) \varphi_{m,n}(x) \quad 4.11$$

$$W_{\psi}(m, n) = \frac{1}{\sqrt{M}} \sum_x f(x) \psi_{m,n}(x)$$

where: φ = scaling function
 M = total number of samples
 m = scaling coefficients
 n = transition coefficients

Equation 4.12 shows the mathematical model of scaling and wavelet functions in cases where λ_o equals to two and t_o equals to one [Gonzalez and Woods, 2002; Gonzalez et al., 2004].

$$\varphi_{m,n} = 2^{-m/2} \varphi(2^{-m}x - n) \quad 4.12$$

$$\psi_{m,n} = 2^{-m/2} \psi(2^{-m}x - n)$$

4.3.5 *One Dimensional Multi-Resolution Analysis*

Multi-resolution analysis (MRA) technique is used for fast analysis of a signal into its frequency bands. As this process is iterated, the resulting approximation is decomposed into its finer frequency components with dyadic down sampling, so the signal is divided into many fine resolution components which is known as the Mallat algorithm shown in

Figure 4-4, where C is the indication of approximations and D is the indication of details. The final reconstructed signal is the result of the final approximation and the summation of all details as shown in Equation 4.13 [Keller, 2004].

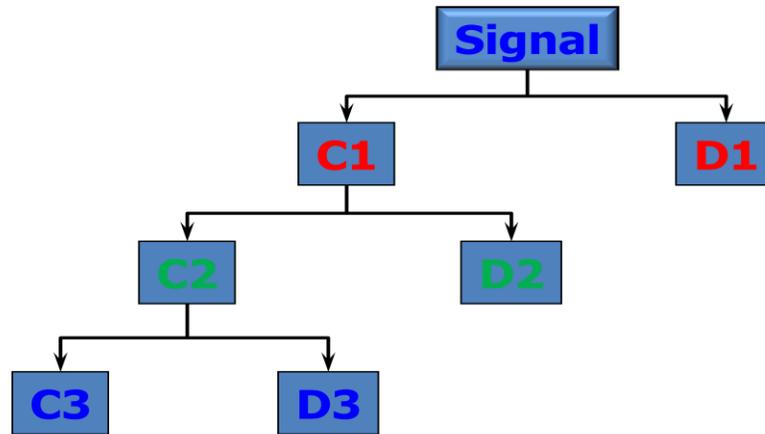


Figure 4-4: Mallat Algorithm

$$f(x) = \sum_n c_N(n) \varphi_{N,n}(x) + \sum_{m=1}^N \sum_{n \in \mathbb{Z}} d_m(n) \psi_{m,n}(x) \quad 4.13$$

where: c = scaling coefficient

d = detail coefficient

N = total number of level of decompositions

4.3.6 Two Dimensional Multi-Resolution Analysis

The image is a two dimensional function. For each direction X and Y, there are two analysing functions φ and ψ (Equation 4.12), Stefen Mallat used the tensor product between scaling and wavelet functions in two directions to investigate the two-dimensional scaling and wavelet functions (Equation 4.14). This process tends to result in four analysing functions, one scaling and three wavelet functions in three directions

(horizontal, vertical and diagonal). At each level of decomposition, image approximation is analyzed into one approximation and three details, horizontal, vertical and diagonal.

$$\begin{aligned}
 \varphi(x, y) &= \varphi(x) \odot \varphi(y) \\
 \psi^H(x, y) &= \psi(x) \odot \varphi(y) \\
 \psi^V(x, y) &= \varphi(x) \odot \psi(y) \\
 \psi^D(x, y) &= \psi(x) \odot \psi(y)
 \end{aligned}
 \tag{4.14}$$

where: $\varphi(x, y)$ = two-dimensional scaling function
 $\psi^H(x, y)$ = two-dimensional horizontal wavelet function
 $\psi^V(x, y)$ = two-dimensional vertical wavelet function
 $\psi^D(x, y)$ = two-dimensional diagonal wavelet function

Equation 4.15 outlines the mathematical model for scaled and translated two-dimensional scaling and wavelet functions where λ_o equals to two and t_o equals to one [Gonzalez et al., 2004; Keller, 2004].

$$\begin{aligned}
 \varphi_{m,n_x,n_y}(x, y) &= 2^{-m/2} \varphi(2^{-m}x - n_x, 2^{-m}y - n_y) \\
 \psi_{m,n_x,n_y}^i(x, y) &= 2^{-m/2} \psi(2^{-m}x - n_x, 2^{-m}y - n_y), i = H, V, D
 \end{aligned}
 \tag{4.15}$$

Building extraction from LIDAR data is executed by analysing the height LIDAR image using two-dimensional wavelet transform to detect building boundaries by monitoring height change positions which will be explained later in this chapter in further detail.

4.4 New MRA Building Extraction Implementation from LIDAR Data

A building extraction process is executed by detecting positions of height differences between adjacent objects using wavelet transform for urban areas. Figure 4-5 shows a

schematic diagram of a building extraction algorithm from LIDAR data [Elhifnawy et al., 2011a].

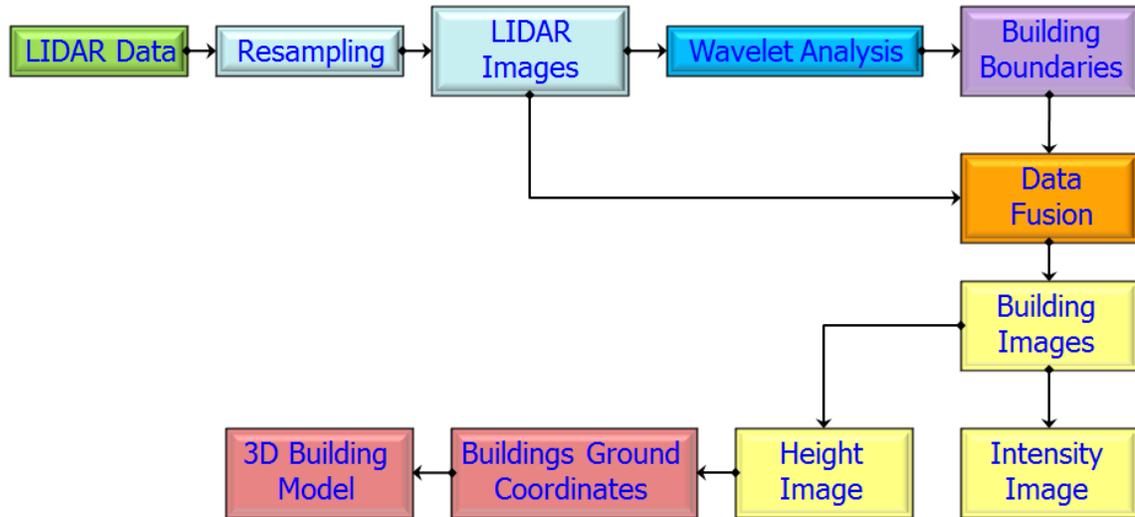


Figure 4-5: Building Extraction Algorithm

LIDAR data is a point cloud of irregular distribution. LIDAR point cloud is regularized using Nearest Neighbour interpolation technique to produce LIDAR images, height and intensity. LIDAR height image is analysed by wavelet transform to detect buildings' boundaries. Images of extracted buildings are produced after applying data fusion between image of buildings boundaries and LIDAR images resulting intensity building image and height building image. the height building image is used to produce a 3-D building model referenced to a ground coordinate system.

4.5 Multi Resolution Analysis of LIDAR Height Image

The LIDAR height image (Figure 4-6) is analysed by sym4 wavelet function (Figure 4-7) to two levels of decompositions. Wavelet function decomposes height image into approximation and six details wavelet coefficients sets, as shown in Figure 4-8 [Elhifnawy et al., 2011b].

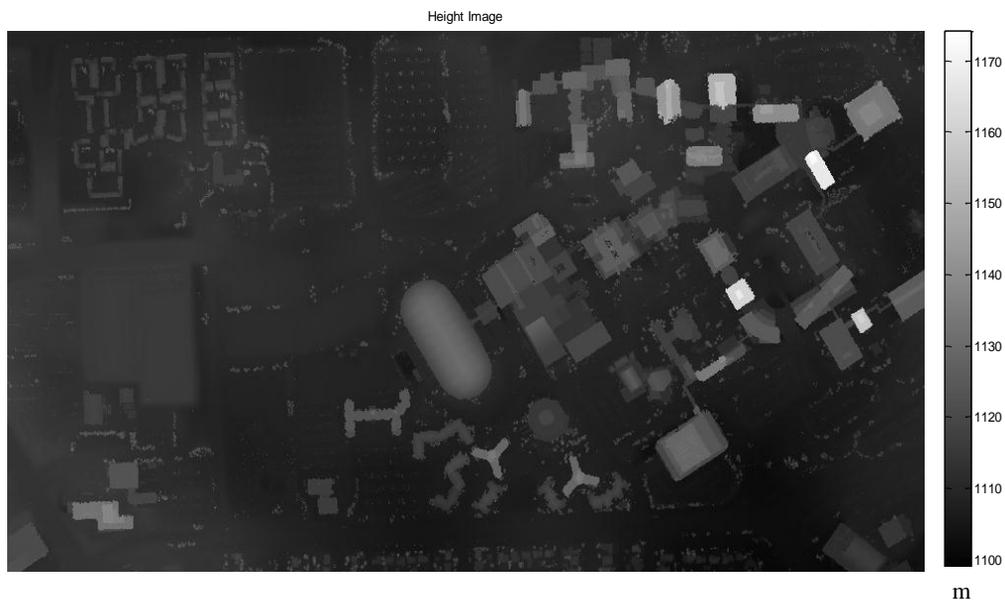


Figure 4-6: LIDAR Height Image of the Area of Study

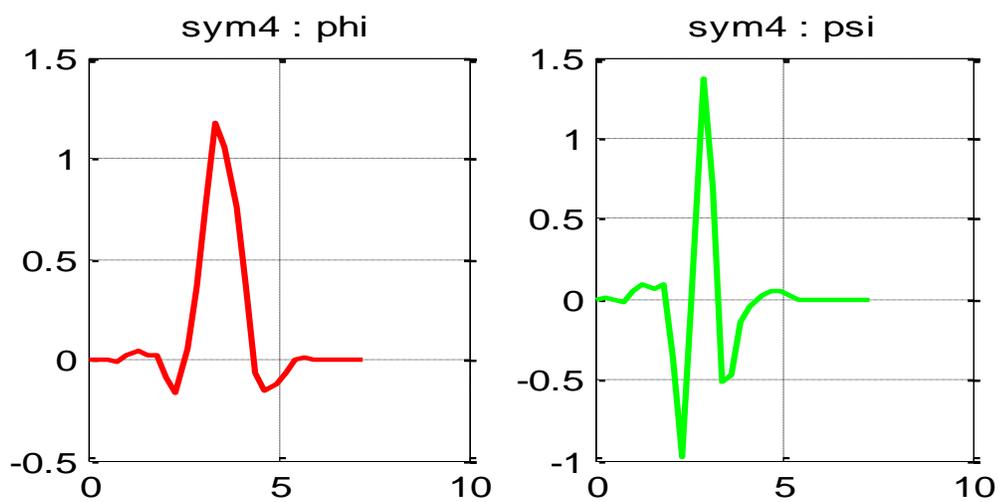


Figure 4-7: SYM4 Wavelet Base Function

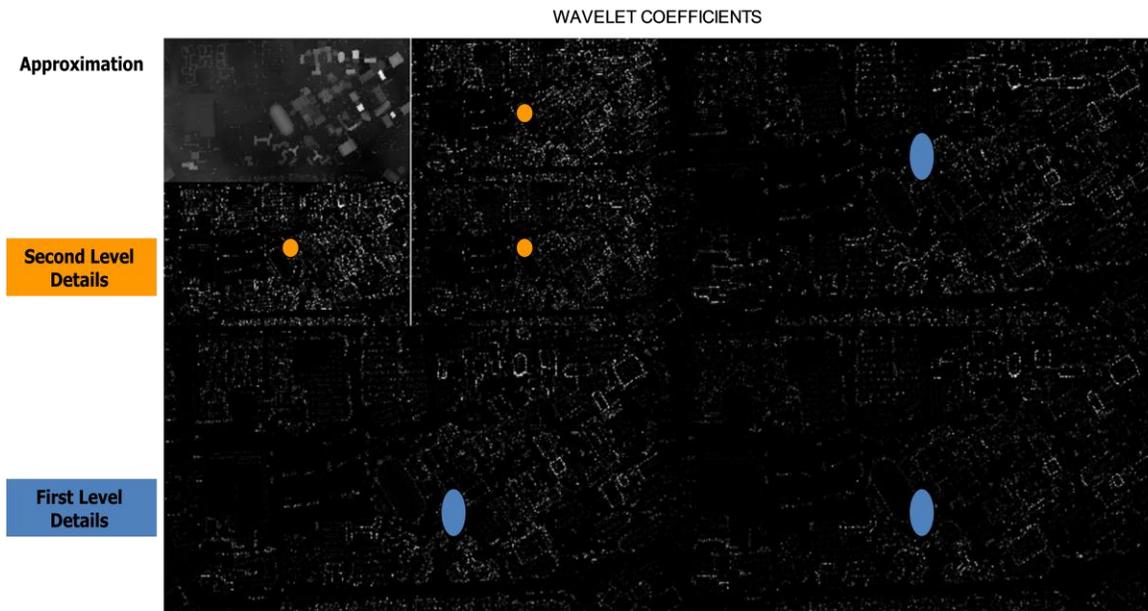


Figure 4-8: Wavelet Analysis

Feature edges are detected as high frequency components in the height image and can be extracted after reconstructing the three details from the second level of decomposition only. In other words, the three details coefficients sets from the first level of decomposition and the approximation set of coefficients are set to zero. Figure 4-9 shows the image of edges of all features in the area of study.



Figure 4-9: Feature Edges of the Area of Study

A morphological operator is applied to fill all feature edges to produce a binary image representing all features in the study area as shown in Figure 4-10, which represents all non-ground objects. Based on Figure 4-10, It can be easily concluded that the wavelet transform with two levels of decompositions and a simple thresholding technique (setting the details coefficients from the first level and approximation coefficients left after the second level to zero) can easily identify and extract all non-ground features from the LIDAR range image. The feature image contains all non-ground objects of urban areas, buildings, trees and cars. The next step involves removing trees and cars from the feature image resulting in building extraction image.



Figure 4-10: Feature Image of the Area of Study

4.6 Data Fusion for 3-D RGB Building Modelling

A feature resultant image from the wavelet transform step contains all non-ground objects including buildings, cars and trees. Building extraction from this feature image is performed using area thresholding to remove objects that have an area less than a specific value, such as cars and vehicles. The feature image is labelled and each label is reviewed with respect to its area. Any feature with an area less than a specific thresholding area is eliminated and considered as a non-building. The remaining labelled pixels represent the most probable building candidates. Figure 4-11 shows the image of building candidates after area thresholding.

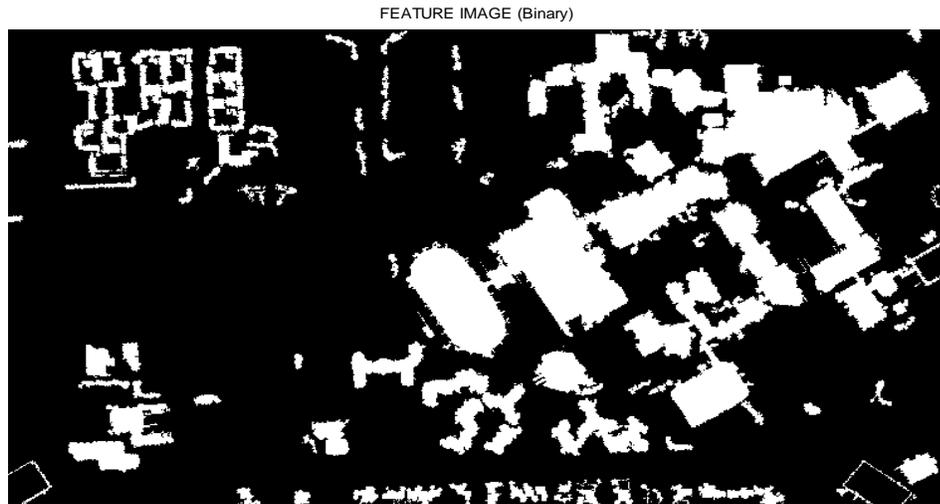


Figure 4-11: Building Candidates of the Area of Study

This image is in binary format meaning all building candidates are represented by ones and others are represented by zeros.

An RGB image for the same area of study is produced by MOSAIC, shown in Figure 4-12. This MOSAIC is a true orthoimage which includes the same area of study but is not the same LIDAR height image scale. Geometric corrections are applied to rescale this MOSAIC for it to be of the same LIDAR image scale using Equation 4.16 [Gonzalez et al., 2004].

$$[x \ y \ 1] = [w \ z \ 1] \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad 4.16$$

where: x, y = coordinates of the output image

w, z = coordinates of the input image

s_x, s_y = scale parameter in x and y directions

An RGB image with the same scale, orientation, and dimensions as a LIDAR height image is extracted using X and Y coordinates that are bounding the area of study (Figure 4-13).



Figure 4-12: MOSAIC Includes the Area of Study

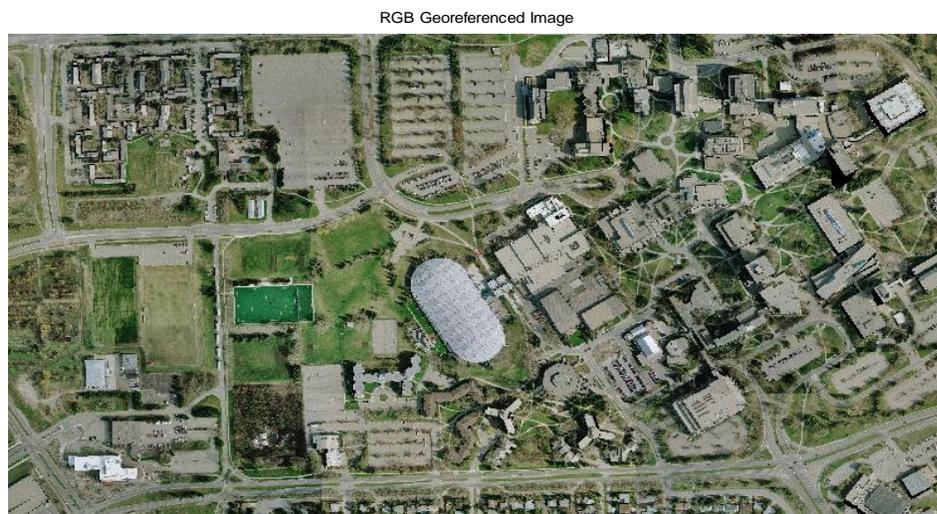


Figure 4-13: RGB Image of the Area of Study

Data fusion is applied between this building candidates' image and corresponding LIDAR intensity data as well as an RGB image for the same area of interest to produce a building image in different color domains (intensity and RGB). It is applied by replacing each one value by its intensity and RGB information. Figure 4-14 shows a building extraction image with intensity information and Figure 4-15 shows an RGB building extraction [Elhifnawy et al., 2011b].

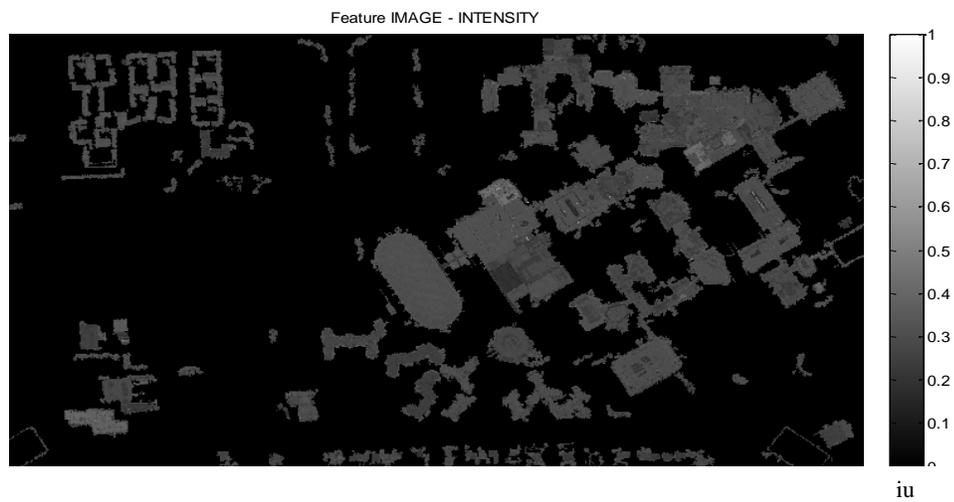


Figure 4-14: Building Extraction Image of the Area of Study (Intensity Information)

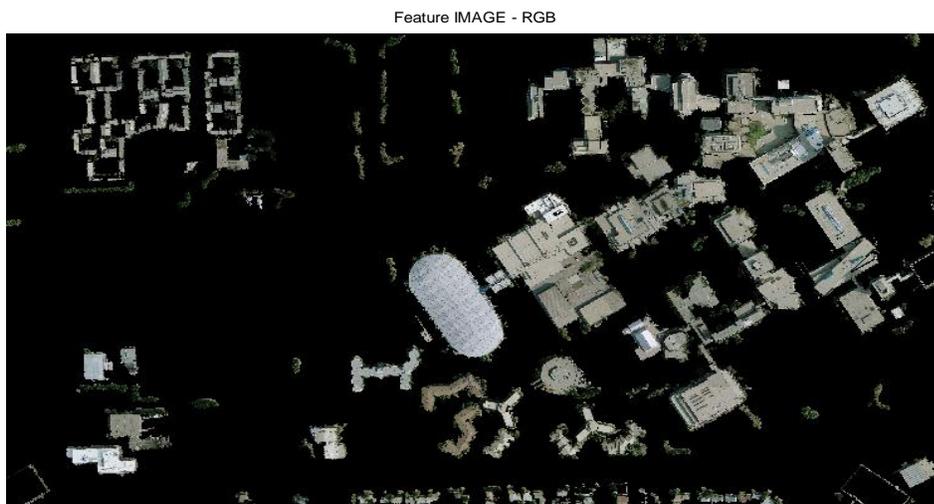


Figure 4-15: Building Extraction Image of the Area of Study (RGB)

Data fusion is applied between this building candidates image and corresponding LIDAR height data to represent an image of extracted buildings by their height information as shown in Figure 4-16.

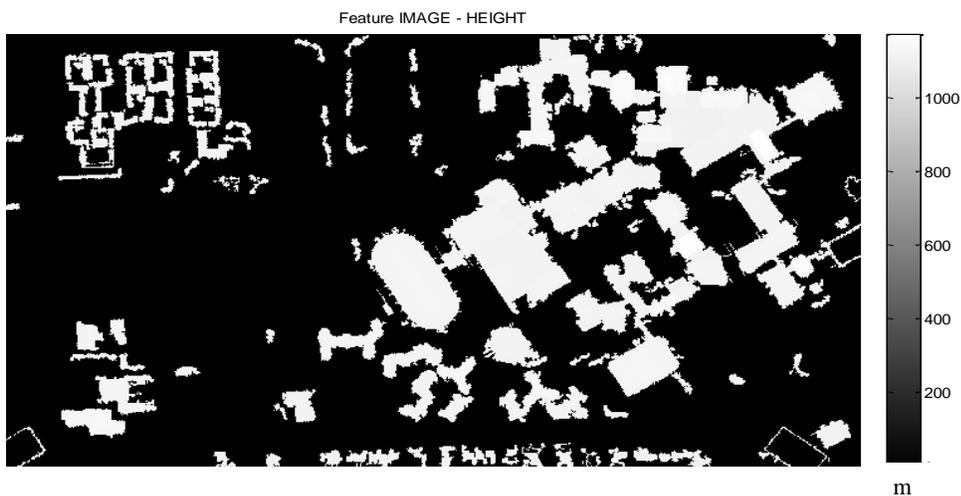


Figure 4-16: Building Extraction Image of the Area of Study (Height Information)

After transforming the image of extracted buildings from an image coordinate system to a ground coordinate system, the full ground coordinates of extracted buildings are investigated. The 3-D building model for the extracted buildings can be generated after calculating full ground positioning information for all extracted buildings. Figure 4-17 and Figure 4-18 depict, from different views referenced to the ground coordinate system, the 3-D building model of the extracted buildings.

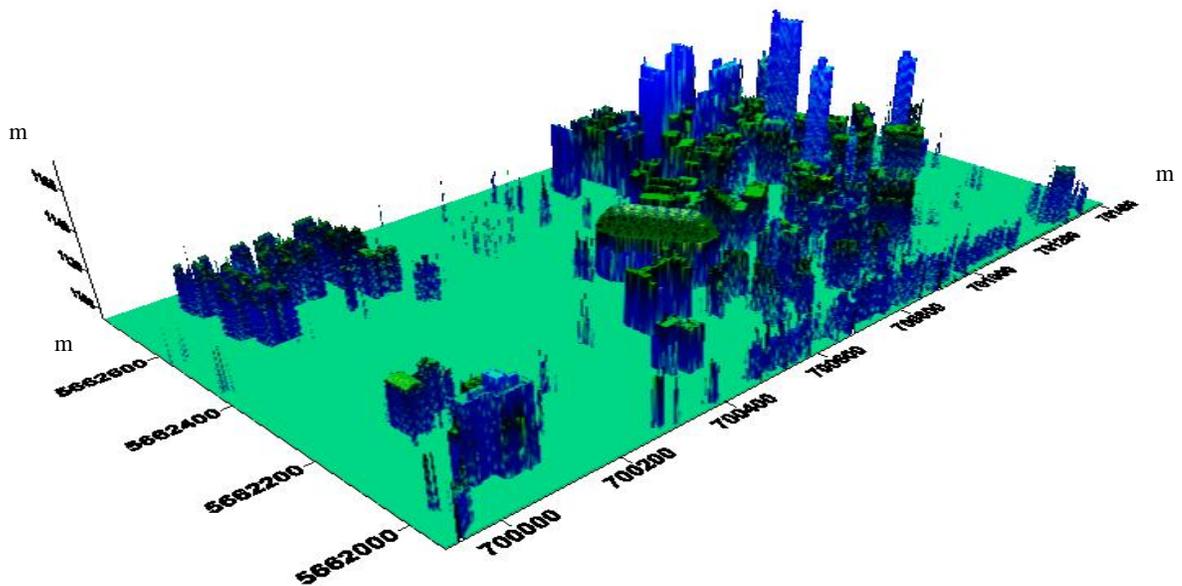


Figure 4-17: 3-D Building Model (View 1)

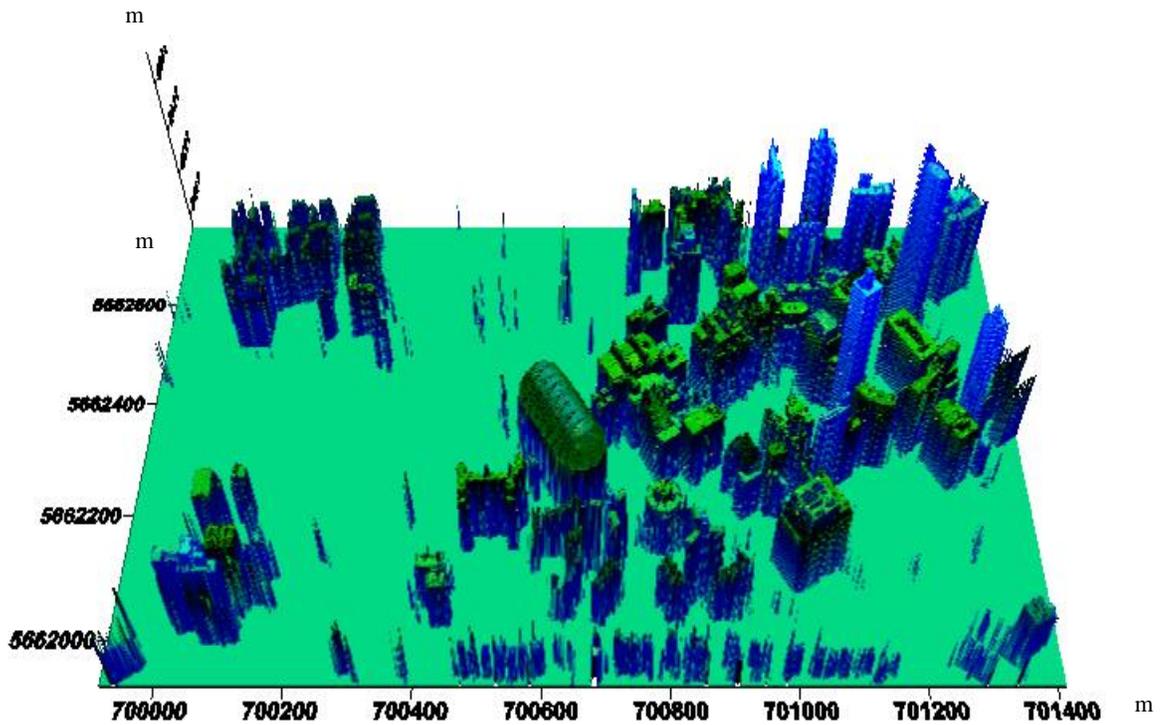


Figure 4-18: 3-D Building Model (View 2)

To assess this algorithm, it is applied on two other LIDAR data, the first is a part of the input LIDAR strip shown in Figure 4-19, while the second is of a residential area located in the province of British Columbia, Canada, shown in Figure 4-24.

The first LIDAR data contains a limited number of buildings and the second represents a crowded urban area. This investigated building extraction technique succeeded in efficiently extracting buildings from the first LIDAR data. Figure 4-20 and Figure 4-21 show building extraction images from the first LIDAR data with height and RGB information respectively. The buildings are extracted with full details as shown in the 3-D building model shown in Figure 4-22 and Figure 4-23. The investigated technique succeeded in extracting buildings from a crowded urban area with high efficiency as

shown in the 3-D building models for the LIDAR data in a residential area of British Columbia (Figure 4-25 and Figure 4-26).

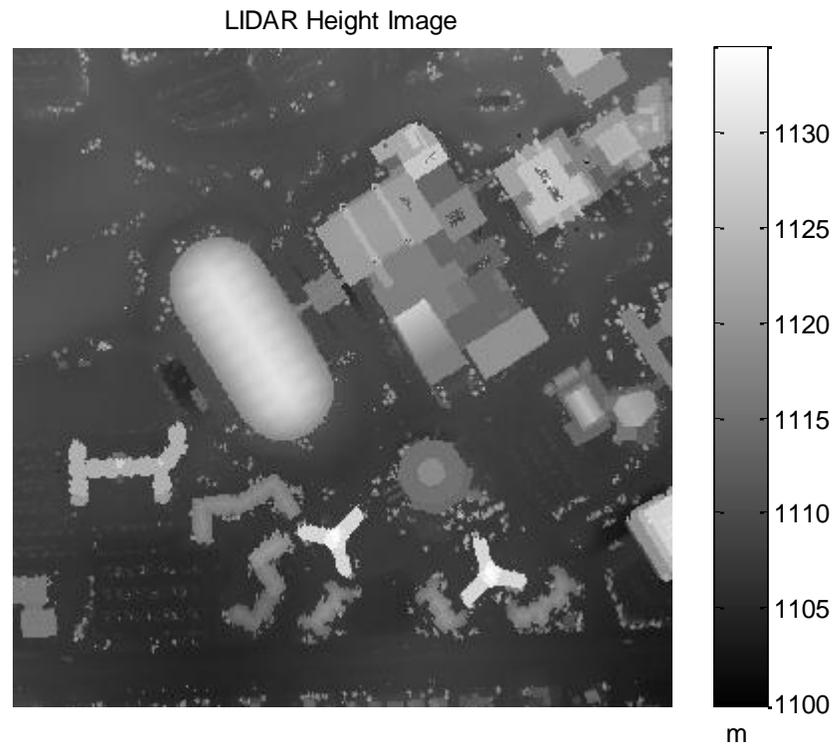


Figure 4-19: Input LIDAR Height Image

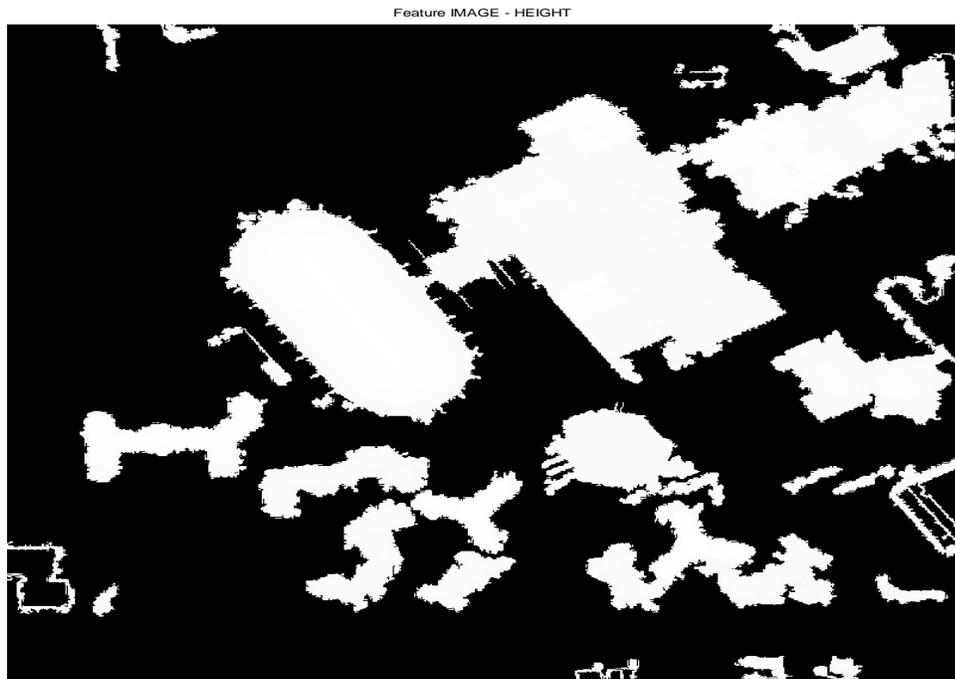


Figure 4-20: Building Extraction Image (Height)



Figure 4-21: Building Extraction Image (RGB)



Figure 4-22: 3-D Building Model (View1)

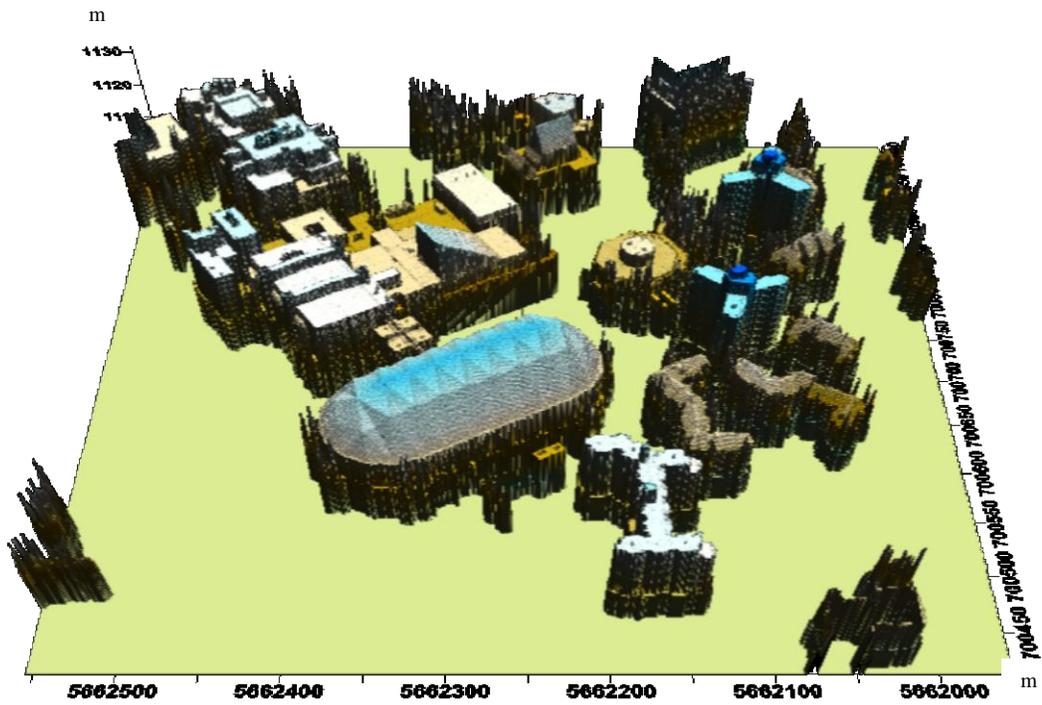


Figure 4-23: 3-D Building Model (View1)

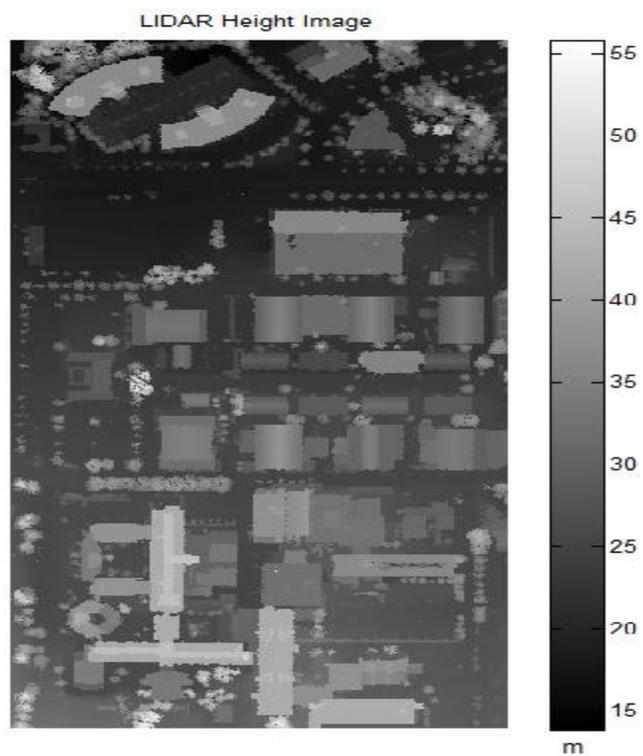


Figure 4-24: Input LIDAR Height Image (British Columbia)

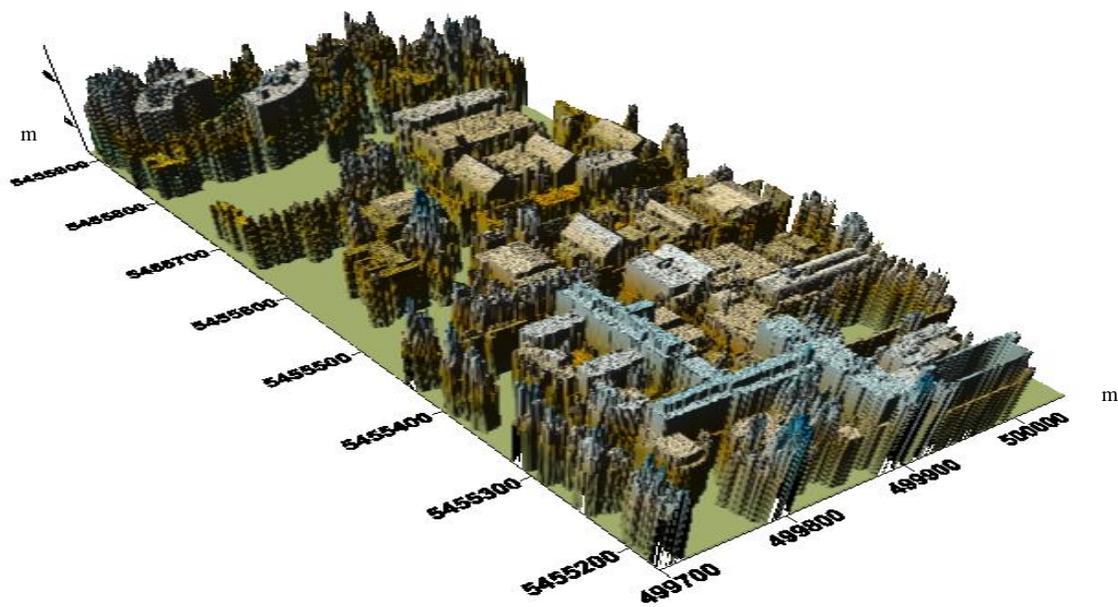


Figure 4-25: 3-D Building Models (Second Data – View1)



Figure 4-26: 3-D Building Models (Second Data – View2)

From the building extraction images, it is observed that there are some features other than buildings extracted as building candidates such as trees. Trees close to the buildings that have the same height information and trees located in a group close to each other forming large areas equal to or larger than the area thresholding value, mentioned before, are extracted as building candidates. This problem is the motivation for using additional information such as texture properties from an RGB image to aid the MRA approach and to be able to separate features sharing height properties.

4.7 Feature Classification from Combined RGB/LIDAR Data

As mentioned in Chapter 3, it is not possible to separate and classify different features sharing same the semantic and texture properties using RGB data only, and it is therefore recommended to use spatial data to aid RGB data in solving this problem. From the

results of the previous sections, it is not possible to separate features which share the same height properties using spatial data only. Subsequently, RGB data is recommended to aid LIDAR data as a solution to this problem. It is clear that both types of data (semantic and spatial) are crucial to complete the picture of an efficient and robust urban classification technique and accurate building extraction technique. This is the main reason and motivation behind the introduction of a new feature classification algorithm using both RGB and LIDAR data for urban classification. Figure 4-27 shows an RGB image that registered with the height image produced by LIDAR data shown in Figure 4-19.



Figure 4-27: RGB Input Image

The proposed technique applies a combination of these available data for extracting and classifying all features overcoming the problems that arise from features sharing same

texture or height properties. Figure 4-28 shows the algorithm for the automatic urban classification technique.

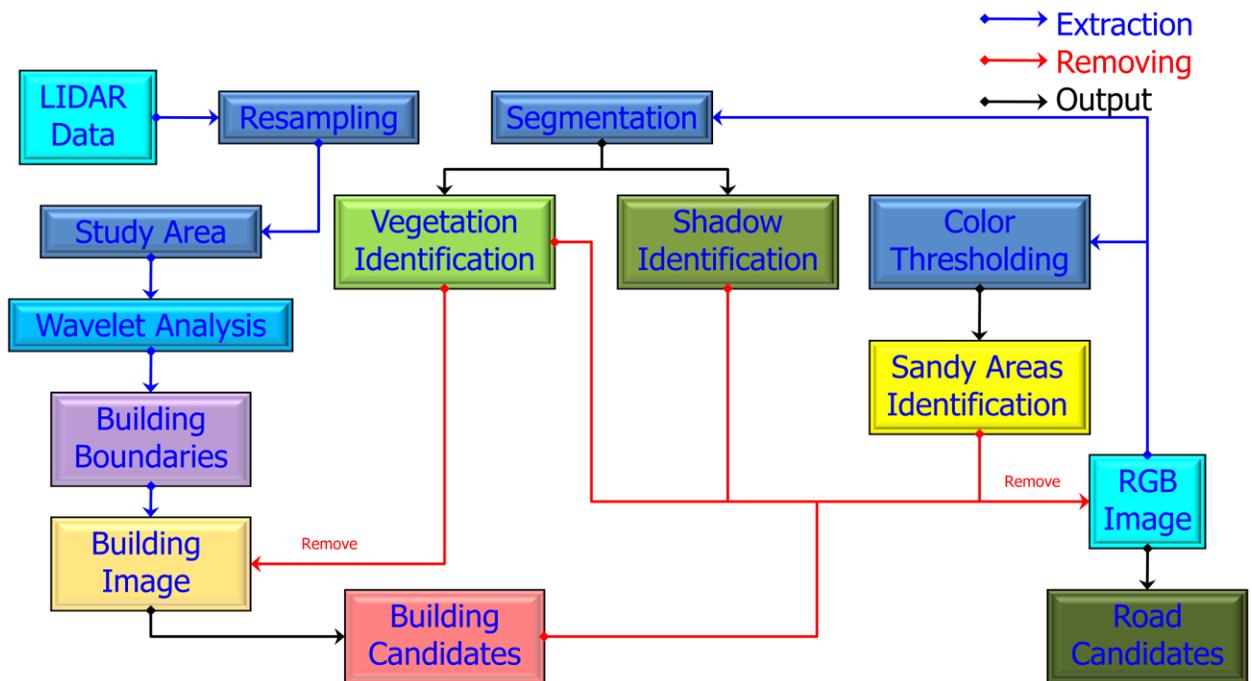


Figure 4-28: Feature Classification Algorithm from RGB/LIDAR Data

This algorithm consists of six steps. The first step is the building extraction from LIDAR data, this process is illustrated in detail earlier in this chapter (Sections 4.5 and 4.6). Figure 4-21 shows RGB building extraction images from LIDAR data after data fusion with the RGB image. The second step is the vegetation, shadows, and sandy area extraction from the RGB image. The results are previously shown in chapter 3 (Figure 3-32, Figure 3-33 and Figure 3-35). The third step is the removal of vegetation and shadows areas from building extraction image produced by LIDAR data (Section 4.6, Figure 4-21). This step removes any extracted green features that are sharing buildings' height information as shown in Figure 4-29. From Figure 4-29 and Figure 3-35, there are shared candidates in both images, some pixels identified as sandy candidates and building

candidates. The fourth step is the removal of building candidates that are identified as sandy areas as shown in Figure 4-30.

Building image after removing vegetation and shadows

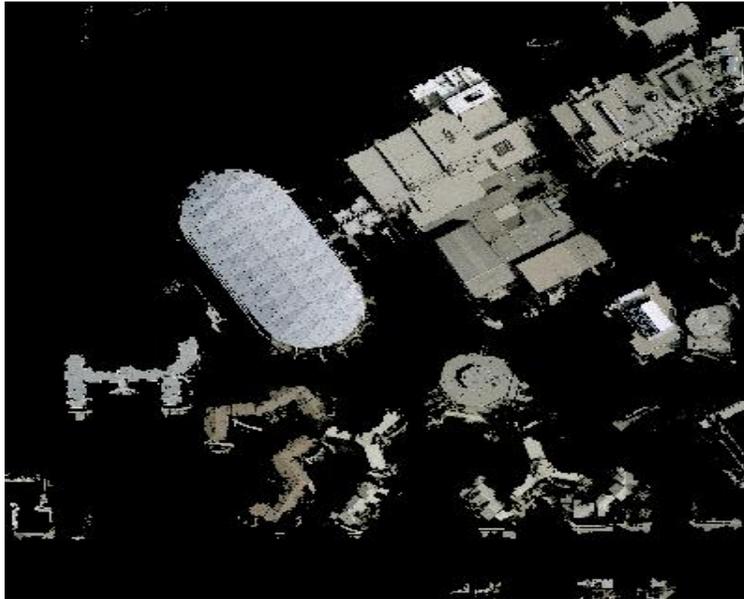


Figure 4-29: Building Extraction Image after Removing Vegetation and Shadows

Sandy Areas Image after Removing Building Candidates



Figure 4-30: Sandy Areas after Eliminating Building Candidates

The fifth step of feature extraction is the road identification. Road candidates can be identified by removing vegetation, shadows, buildings, and sandy areas from the input RGB image as shown in Figure 4-31.

Road Image

**Figure 4-31: Road Image**

From Figure 4-30 and Figure 4-31, it is found that the remaining sandy areas, after eliminating shared candidates with buildings, must be reclassified as roads, this is the last step of the urban area classification algorithm. Figure 4-32 shows the final RGB road image and Figure 4-33 is the final feature classification image [Elhabiby et al., 2011].

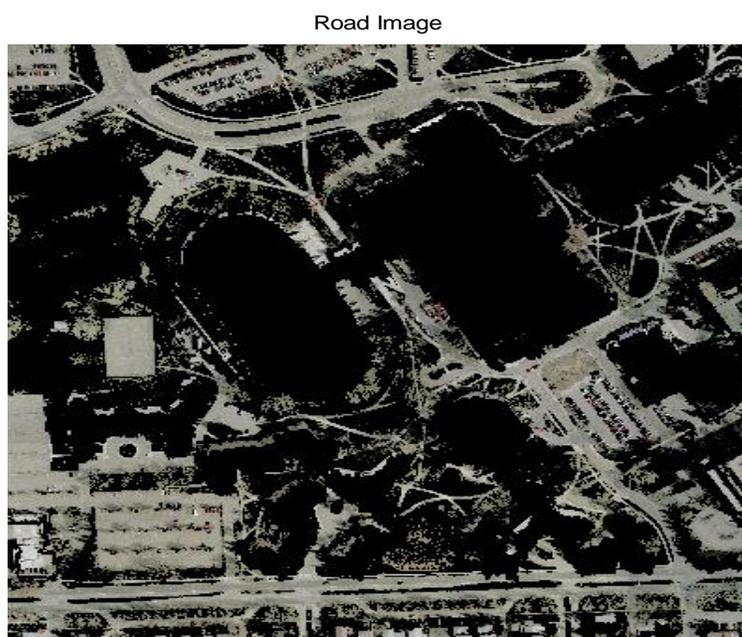


Figure 4-32: Road Extraction Image after Considering Sandy Areas as Roads

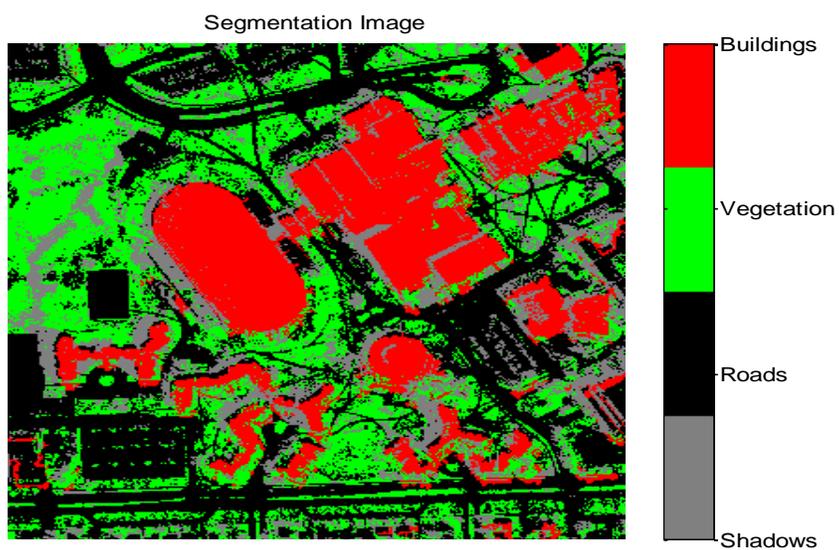


Figure 4-33: Final Feature Classification Image [Elhabiby et al., 2011]

4.8 Quantitative Assessment for the Automatic Classification Technique from Combined RGB/LIDAR Data Using Wavelets and Statistical Filters

The area of interest is located on the University of Calgary campus. The assessment of the classification results was done in two steps. The first step was a visual, object-based assessment, where the extracted objects were counted visually from the aerial image and on the ground. This step focuses on the buildings that are extracted from combined the LIDAR and RGB image. This collected data is considered as ground truth for buildings. The classification technique using combined LIDAR data with RGB imagery succeeded in extracting all the buildings inside the area of study with a success percentage equal to 100 % (13 buildings out of 13). The automatic classification technique succeeded in extracting all the edges of the existing buildings in the scene (complete building or parts of the buildings that exists in the scene image) with a success percentage approximately equal to 90%. Figure 4-34 shows the final building candidates that are extracted from the automatic classification technique presented in this research thesis.

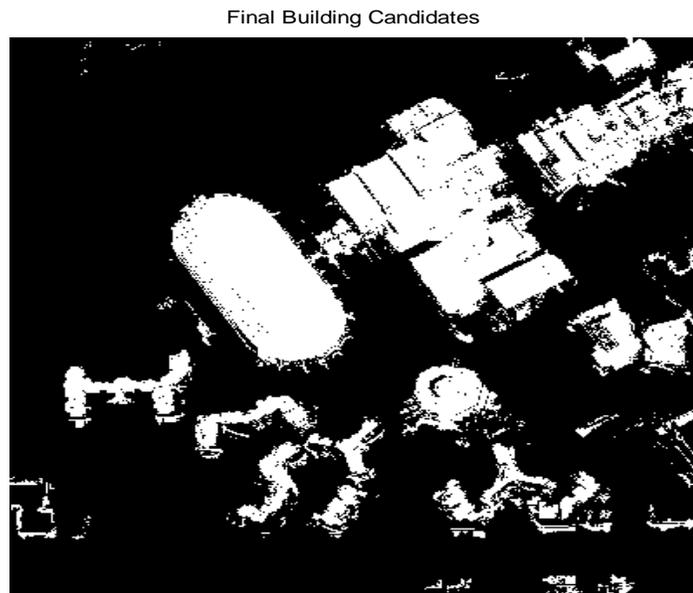


Figure 4-34: Final Building Candidates from the Proposed Classification Technique using Combined RGB/LIDAR Data

The second step is a comparative pixel-based assessment that was applied by checking the number of pixels that verify and compare the correct extracted objects with the GIS data from the area of study (created from an independent source of data). Figure 4-35 shows the available GIS data for the University of Calgary. The available data is a shape file including buildings, roads and vegetation areas.



Figure 4-35: GIS Data for the University of Calgary Including Buildings, Roads and Vegetation

The available GIS data was used to produce three different images (layers) for buildings, roads and vegetation areas from the same area of study. The images from Figure 4-36 to Figure 4-38 show the building, road and vegetation candidates from the GIS data.

Building Candidates from GIS Data

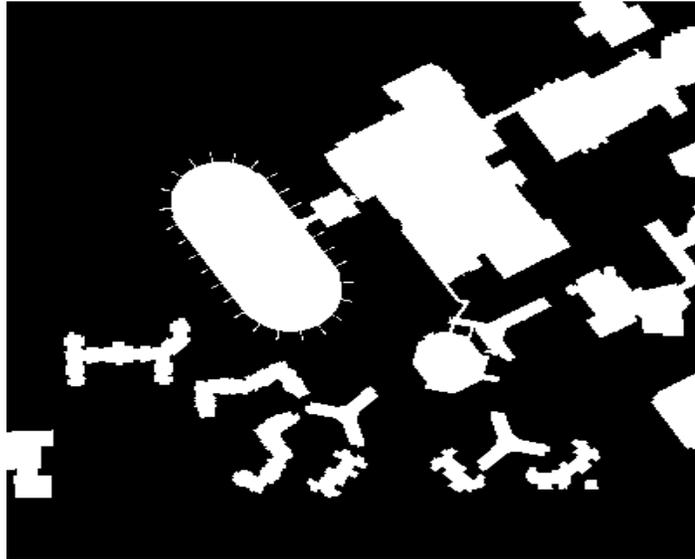


Figure 4-36: Building Candidates from GIS Data

Road Candidates from GIS Data

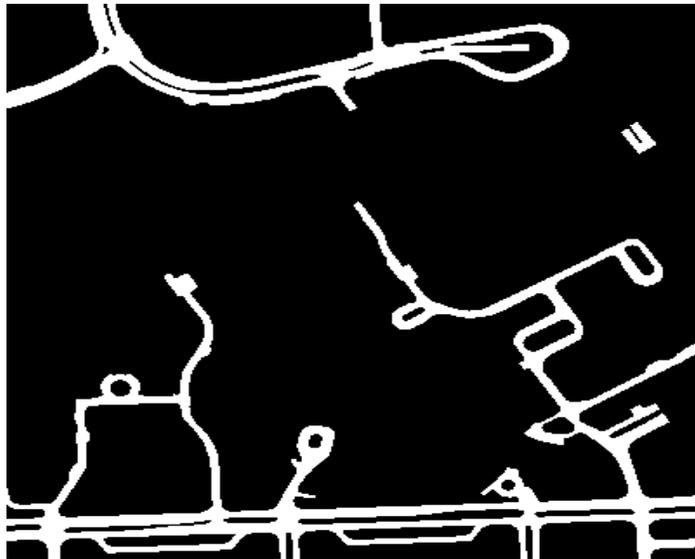


Figure 4-37: Road Candidates from GIS Data

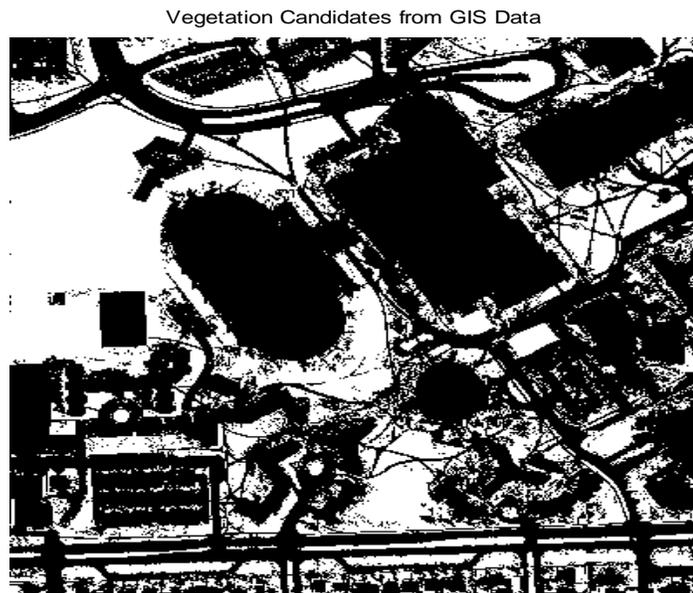


Figure 4-38: Vegetation Candidates from GIS data

In Figure 4-37, it is determined that the GIS data collected and built by the University of Calgary, focused on recording the main roads and neglected the secondary roads, while the classification technique was entirely successful in extracting all the road types by producing complete geometric properties, shape and width.

The classification results and GIS data were used to build a confusion matrix [Dance et al., 2004] for the sake of comparison. Table 4-1 shows the confusion matrix between the classification results and the GIS data (reference). This matrix represents the number of pixels that verify specific features against the results from the new extraction technique and the GIS reference data. The first row of the confusion matrix listed the available features in GIS data (reference data) while the first column lists the extracted features from the classification technique. The last row lists the total number of pixels in GIS data

corresponding to each feature of the three while the last column lists the total number of classified pixels for the Classification results.

Table 4-1: Confusion Matrix of the Classification Results Versus GIS Data

<i>Reference GIS Classification</i>				
Classification Results	<i>Buildings</i>	<i>Roads</i>	<i>Vegetation</i>	<i>Total</i>
Buildings	285990	535	9572	296097
Roads	27079	134476	82277	243832
Vegetation	32473	46168	431955	510596
Total	345542	181179	523804	1050525

The numbers of pixels that verify correct classification are listed in the confusion matrix in diagonal direction. The classification technique succeeded in extracting buildings with a success percentage equal to 83 % (285990 building pixels/345542 total GIS data building pixels) as shown in Figure 4-39 , while 74 % (134476 roads pixels/181179 total GIS data road pixels) for roads (neglecting the extracted secondary roads) as shown in Figure 4-40 and 82 % (431955 vegetation pixels/523804 total GIS vegetation pixels) for vegetation as shown in Figure 4-41. Figure 4-42 shows the overall classification assessment for the three classified features, building, road and vegetation with a total success percentage equal to 81 % (852421/1050525).

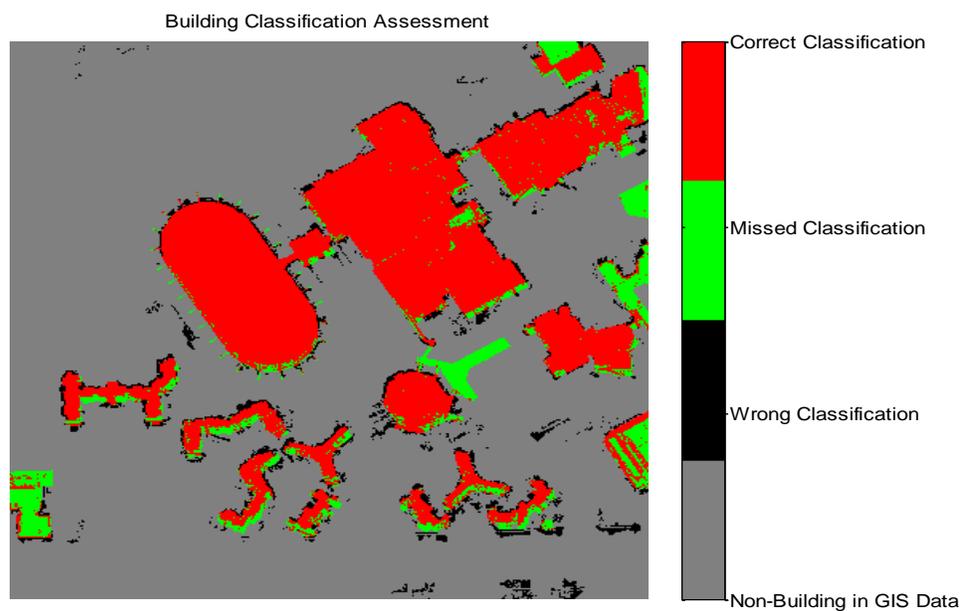


Figure 4-39: Building Classification Assessment versus GIS Data

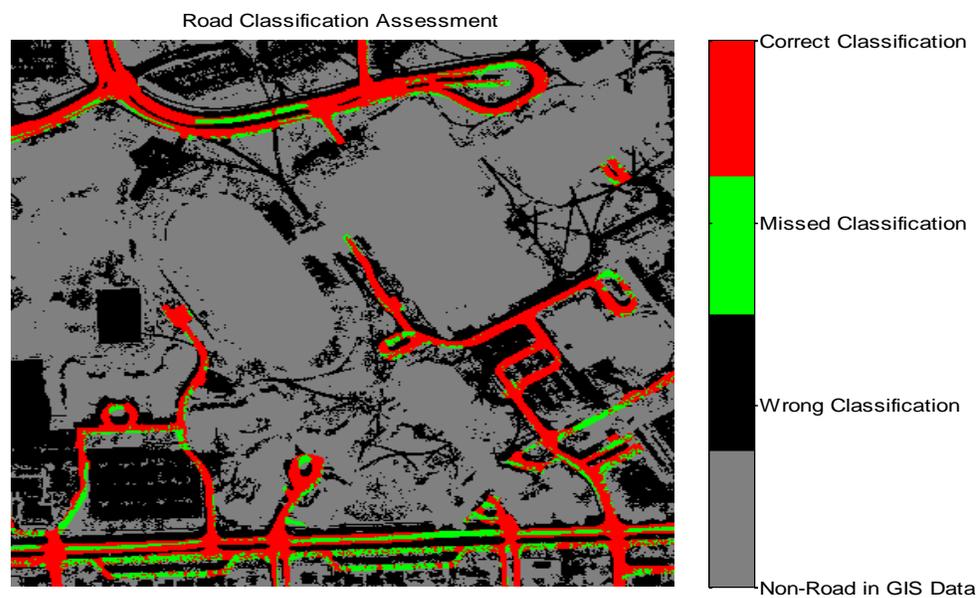


Figure 4-40: Road Classification Assessment versus GIS Data

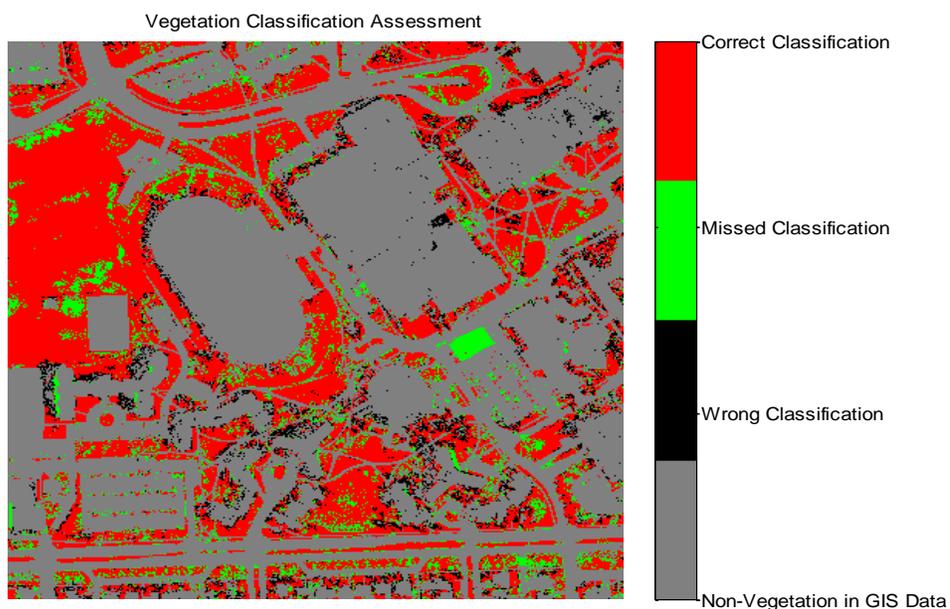


Figure 4-41: Vegetation Classification Assessment versus GIS Data

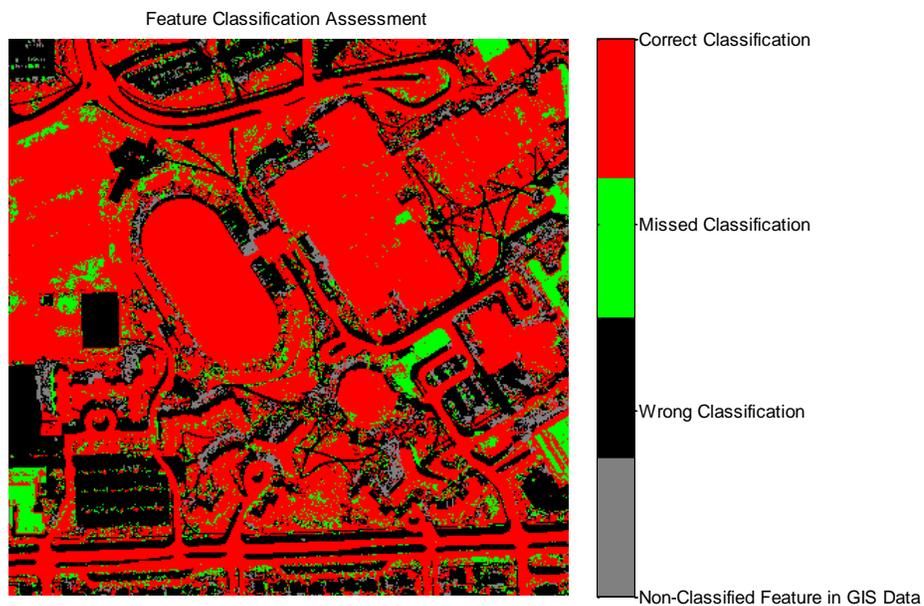


Figure 4-42: Classification Technique Assessment versus GIS Data for Building, Road and Vegetation Classification

The classification technique succeeded in classifying not only three features; building, road and vegetation, but also succeeded in detecting the shadow areas. The shadow areas might share pixels with any of the three main extracted features, buildings, road and vegetation. Table 4-2 shows the same confusion matrix as Table 4-1 however has an additional row of the shadow areas to represent the number of shadow area pixels and their classification related to the three main features in urban areas (building, road and vegetation) that affect pixel-based assessment using available GIS data.

Table 4-2: Confusion Matrix after Adding the Classification Results of Shadowy Areas

<i>Reference GIS Classification</i>	<i>Buildings</i>	<i>Roads</i>	<i>Vegetation</i>	<i>Total</i>
Classification Results				
Buildings	285990	535	9572	296097
Roads	27079	134476	82277	243832
Vegetation	32473	46168	431955	510596
Total	345542	181179	523804	1050525
Shadows	29382	10518	7493	

The visual, object-based assessment of the classification technique established on ground truth reference data, using the number of buildings, yields better results than the pixel-based assessment, using available GIS data since the available GIS data is inaccurate and

is produced from an RGB image without any ground truth data evaluation. The results from the previous two assessments achieved excellent results from both the visual object-based assessment,, and the statistical comparative pixel-based assessment. This indicates the high efficiency of the proposed automatic classification technique using a combination of RGB, semantic, LIDAR, and spatial data for introducing an automatic and complete urban classification technique (buildings, roads and vegetation).

Chapter Five: Monte Carlo Simulation Descriptor for Object Shape Recognition

The results of the feature extraction technique implementation are represented by object boundaries (Chapter 4) and textures (Chapters 3 and 4). These representations are used for creating a library of objects of interest (buildings) for an automatic object recognition application. Object recognition is one of these database approaches that have many applications especially in the military field, such as live monitoring change detection for targets or objects [Butkiewicz et al., 2008]. The thesis concentrates on object boundary representations to be utilized for an innovative and efficient object descriptor that can be employed in the automatic recognition of urban objects through different images (registered or not) captured at different times. Object descriptor is considered as a second step for an efficient urban features classification search algorithm, the first step is an object representation.

The automatic urban classification technique introduced in this thesis will help in building a library of the classified objects. Having an efficient search algorithm for automatic object recognition will be very important to efficiently utilize the classification output. The exact identification of an object can be determined with the aid of a knowledge database that contains the most probable objects existing in the application field [Ruel et al., 2004]. Object recognition is based on forming specific descriptors and having an efficient classification technique. An object is identified by the image as being a vehicle, a tree or a building and can be classified into many categories and/or classes, which can be efficiently identified from images. The classes can include usage, types, shapes or any other specific categories. The classification of objects to predefined classes is known as the object recognition process and is usually based on the fact that an object

consumes a region in the image and the boundaries of this region are used to generate object descriptors. This research thesis produces and assesses a new spatial domain object descriptor using the advantages of efficient urban classification investigated by an object boundary image based on the Monte Carlo Simulation (MCS) technique for object area determination application.

There are many object recognition techniques based on MCS such as traffic light recognition from color images produced by the Mobile Mapping System (MMS) [Tu and Li, 2002]. The Monte Carlo optimization technique was used for investigating a 3-D object recognition framework using *HSV* color space properties [Sungho et al., 2005]. Gronwall et al. [2006] used MCS for testing the performance of the rectangular estimation technique for the Automatic Target Recognition (ATR). Hundeishausen and Veloso [2007] introduced a new technique called Active Monte Carlo Recognition (AMCR) for object recognition with different scales and orientations. And Liang [2008] used MCS for the ATR for simulated objects in radar sensor networks.

In this thesis, the proposed techniques are based on the Monte Carlo Simulation (MCS) methods for estimating the probability of occurrence/existence for an object of interest required to be recognized from input imagery or data. But no previous research dealing with a new Monte Carlo Simulation Descriptor (MCSD) as will be illustrated later in this chapter. The efficiency of this new descriptor is tested and assessed in comparison to other varying descriptors for object shape recognition; they will be presented in this chapter in details.

5.1 Different Object Descriptors, Production and Assessment, Literature Review

The object descriptor can be represented in both spatial and frequency domains. Fourier and wavelet transform can be used for generating object descriptors [Tieng and Boles, 1997; Boulinguez and Quinquis, 1999; Osowski and Nghia, 2002; Kunttu et al., 2003; Mahmoud and Mahmoud, 2006; Ali et al., 2008; Smach et al., 2008]. These object descriptor parameters are then used in producing the objects library otherwise, the backbone of the object recognition process.

Many researches have shown the efficiency of wavelet and Fourier transform for generating object descriptors in object recognition applications. Osowski and Nghia [2002] presented a comparative study between Fourier descriptors and wavelet descriptors in object recognition applications. The object of interest is represented by its boundary and is used to extract the coordinates for boundary pixels. The boundary coordinates are used to form three 1-D functions. One of these functions is analysed by Fourier transform to investigate Fourier coefficients and the others are analysed by wavelet transform to investigate wavelet coefficients. 1-D function for Fourier transform is formed by complex numbers. Each function pin is formed from coordinates of a boundary pixel. Each pin is composed of an X-coordinate as a real component and a Y-coordinate multiplied by the square root of -1 as an imaginary component. Calculation of wavelet descriptors is based on dividing boundary coordinates into two 1-D functions. X-coordinates form a 1-D function and Y-coordinates form another 1-D function. The wavelet transform is applied to these two functions so as to investigate wavelet coefficients using the Daubechies (db4) wavelet function. The normalization process is required to make coefficients translation, orientation and scale invariant. Three neural

network classifiers are applied for the comparative study; Multilayer perceptron, self-organizing Kohonen-type network and Hybrid neural network. The best results are obtained when using self-organizing for both Fourier and wavelet descriptors. The comparative study is performed based on the results from self-organizing only. The wavelet descriptors provided better results than Fourier descriptors, but the entire procedure remained complicated and obscure.

Tiehua et al. [2004] tested the efficiency of wavelet coefficients in object recognition. The first step of object recognition was the feature extraction. The object of interest is represented by line segments. Each line segment has two corner points. The normalization process is applied to make each line segment translation, orientation and scaling invariant. The resultant line segment is divided into 2^n segments to be suitable for dyadic wavelet decomposition. The x components and y components for each line segment form a 1-D function. The wavelet transform is then applied to each function for each line segment to get wavelet coefficients for three levels of decomposition using B-spline 1.3 wavelet function. The last step of object recognition was the segment matching between segments of the object of interest and model object. The matching was applied for the approximation coefficients and for the detail coefficients. The matching process began by using approximation coefficients only and was repeated several times based on the number of details by adding one set of detail coefficients at a time. The segment matching is accepted if the dissimilarity function gives a value less than or equal to a predefined threshold. This step is followed by a hypothesis verification step. This step analyses the matching results to investigate the final recognition result of the object model related to the objects in the image. Wavelet descriptors gave adequate recognition

results, but this approach was complex using a combination of many techniques to end up with the classification result. This research thesis will produce a recognition technique using simple object description and classification methods.

Nabout and Tibken [2004] also introduced a wavelet object description method, where the object contours are extracted using the Object oriented Contour Extractor (OCE) method. The object boundary is represented by an outer contour line to form a polygon. This polygon is represented by polar representation and signal is represented by periodic angle function that contains the length of the polygon line in each pin. This function is analysed using the Daubechies (db4) wavelet function to obtain the object descriptors. This research is extended by introducing an object descriptor method based on the Mexican Hat wavelet function [Nabout and Tibken, 2007]. The object is represented by its outer contour line. The OCE method is used to extract the contour line and is described using a chain code. The Curvature Dependent Contour Approximation (CDCA) method is used to represent the contour line as a polygon. The angle differences between polygon vertices are calculated and form an angle function that contains the polygon lengths of its sides. As previously mentioned, this function is analyzed using the Mexican Hat wavelet function considering the first sixteen wavelet coefficients as the object descriptor. The recognition is mainly to differentiate between circular, triangle and square object shapes. The minimum distance method is used as a classification technique. The same research work is repeated with the Haar wavelet function [Nabout and Tibken, 2008]. Generally, as an output it has been found that the wavelet descriptors are efficient when using a minimum distance equation as a classification method.

The previous research work showed the efficiency of using the wavelet transform in object recognition in different applications. In this chapter the automatic urban classification technique will be completed by an efficient search algorithm for the identification of different objects from libraries created during the classification step. The proposed technique will be based on MCSD.

In addition, the efficiency of the new spatial domain descriptor, MCSD, will be introduced versus the wavelet descriptors for object recognition. The application and assessment will be introduced for both the boundary and texture of the objects (buildings).

5.2 Object Description and Classification Methodologies

As mentioned in the previous sections, a new spatial domain descriptor is used for object recognition from any data type based on the MCS technique. This section introduces an overview of MCS applications for image processing and the method of calculating its new spatial domain descriptor. Moreover, the mathematical background of object classification methodology used in the proposed recognition technique, Minimum Distance Classifier (MDC), will be introduced.

5.2.1 Monte Carlo Simulation for Object and Image Processing Applications

Monte Carlo Simulation (MCS) is a modelling process technique using a generation of random values [Huber, 1997]. MCS is used to generate different filters based on random variables with specific distributions. These filters have a great impact for image denoising applications [Xu and Pattanaik, 2005; Wong et al., 2008]. MCS succeeded in deriving object area by generating random numbers with specific probability density function inside the object boundary. This section introduces the method of generating a

new object spatial domain descriptor based on the MCS technique.

5.2.2 Monte Carlo Simulation Descriptor for Object Recognition

The extracted object (Chapters 3 and 4) is represented by a boundary image. This image contains one object boundary only, where the image borders are bounding and touching the object boundary. Uniformly distributed random values are generated inside this image, and the distribution is controlled by a Probability Density Function (PDF) shown in Equation 5.1 where a and b are the distribution parameters. These parameters are used for calculating the mean and variance as shown in Equation 5.2 [Gonzalez et al., 2004].

$$P_z(z) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq z \leq b \\ 0 & \text{otherwise} \end{cases} \quad 5.1$$

$$m = \frac{a+b}{2} \quad 5.2$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$

where: m = mean

σ^2 = variance

Random values will be segmented into two groups as shown in Figure 5-1. The first group includes the random values that are located inside the object boundary, shown as red points. The second group represents the total generated random values, shown as all points (red and green).

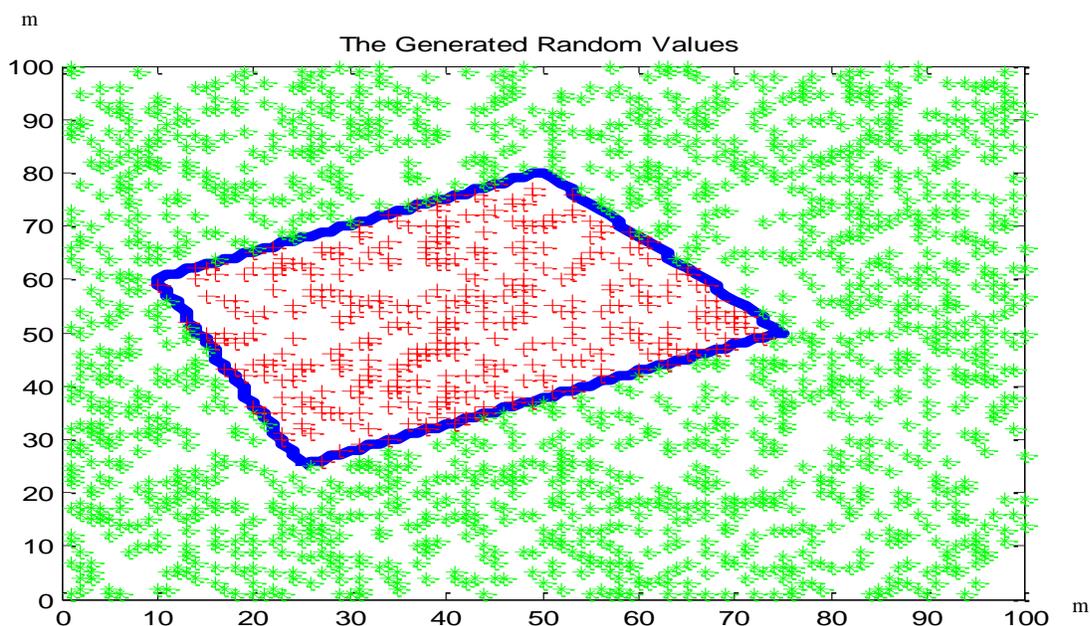


Figure 5-1: Generated Random Values Segmentation

The ratio between the number of random values inside the object boundary and the number of all generated random values represent the ratio of object area with respect to the total area of the image. This ratio is used as an object boundary or area descriptor and is known as the Monte Carlo Simulation Descriptor (MCSD) [Elhifnawy et al., 2010], which is not dependent upon object scale or orientation, thereby is considered as effective object descriptive information.

5.2.3 *Minimum Distance Classifier*

The last step of the object recognition process is classification. Object classification is based on calculating decision functions using the object of interest and library of classes descriptors as $df_1, df_2, df_3, \dots, df_{DF}$. Decision functions for assessing and testing the efficiency of the new investigated object descriptor are calculated using the Minimum Distance Classifier (MDC) method. This method tends to assign objects to the specific

class by the minimum distance classifier matching. This method is based on calculating the distance between vectors using Euclidean Distance (ED) (Equation 5.3). The object is classified to the class related to the vector that gives the minimum value [Gonzalez et al., 2004].

$$df = \|OD - LD_j\| \quad 5.3$$

where: df = decision function

OD = object of interest descriptor

LD = library of classes descriptors

$j = 1, 2, 3, \dots, DF$

DF = total number of classes

5.3 Library Preparation from Building Extracted Image

The most important step for object recognition is the generation of a library of different classes. The classes in this research are presented by objects with different shapes. The source of the library is the buildings extracted from the urban classification step introduced in the previous chapters. This library contains all geometric and semantic information and descriptors about the buildings extracted from LIDAR data such as area, centroid, boundary and intensity information. Figure 5-2 shows a library preparation algorithm for each object (class) using different descriptors, MCSD and wavelet descriptors.

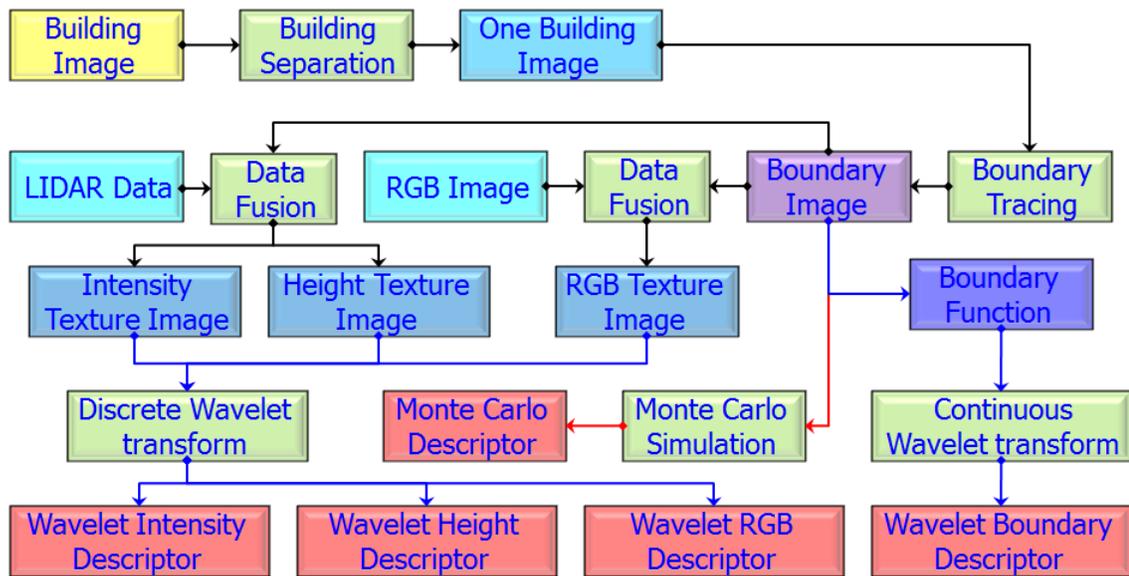


Figure 5-2: Library Preparation Algorithm

The algorithm begins with the buildings extracted from LIDAR data, as a result of the urban classification technique, shown in Figure 5-3. All extracted buildings are separated to get an image for each building. Building separation is performed in three steps. The first step involves the removal of small connectivity between buildings executed using morphological operators. The second step is labelling connected pixels based on the eight-connectivity behaviour. The last step is the production of an image for connected pixels (one building alone) after removing small objects based on predefined area thresholding. The number of images equals the number of extracted buildings from LIDAR data. Figure 5-4 shows buildings after the separation process. These images represent objects with different shapes and the library then contains all information for each object such as area, perimeter and centroid. All available information can be used to recognize the corresponding object from any image. The research thesis concentrates on the boundary and texture representatives for object recognition application. The next sub-

sections illustrate the methodology to represent and describe objects in the wavelet transform and MCS domains.

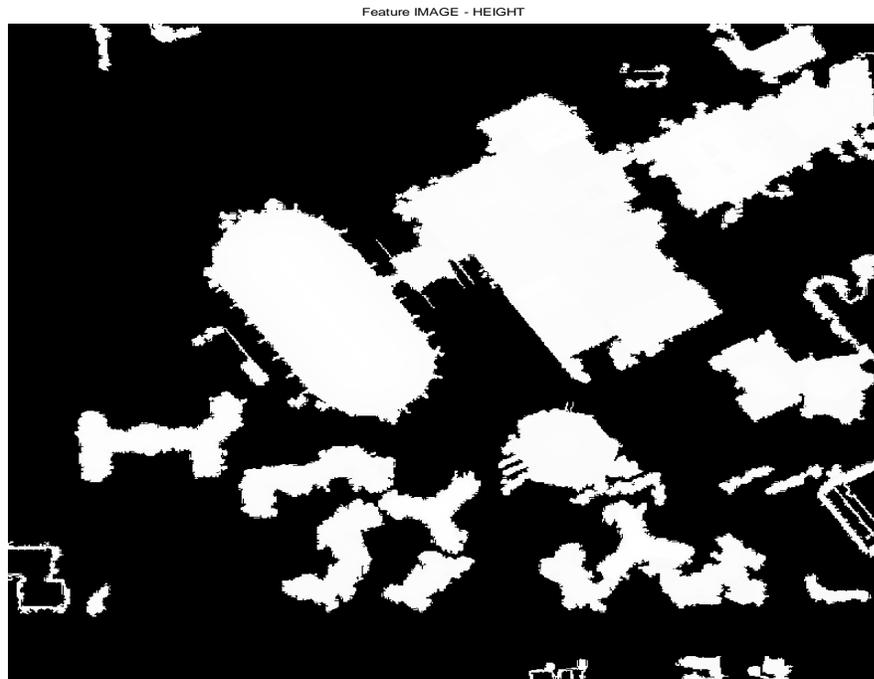


Figure 5-3: Building Image

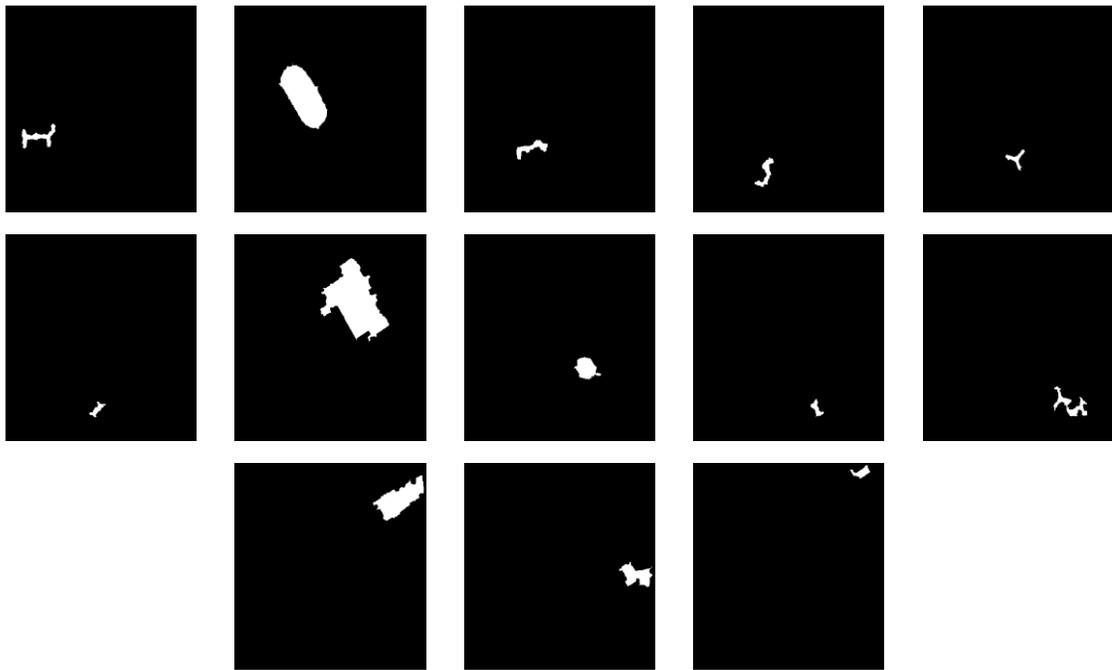


Figure 5-4: Building Separation

5.3.1 *Boundary and Texture Object Representation*

The boundary tracing technique is applied on each image to extract building boundary coordinates. The boundary coordinates of each building are used to generate building boundary image. Each object is represented by its boundary as shown in Figure 5-5. The boundaries in these images are in irregular shapes without any boundary enhancement or removing attached vegetation. This approach is useful for generating a library of objects with random shapes, irregular boundaries, random scales and orientation. The consequences and success of this approach can be generalized for almost all object shapes.

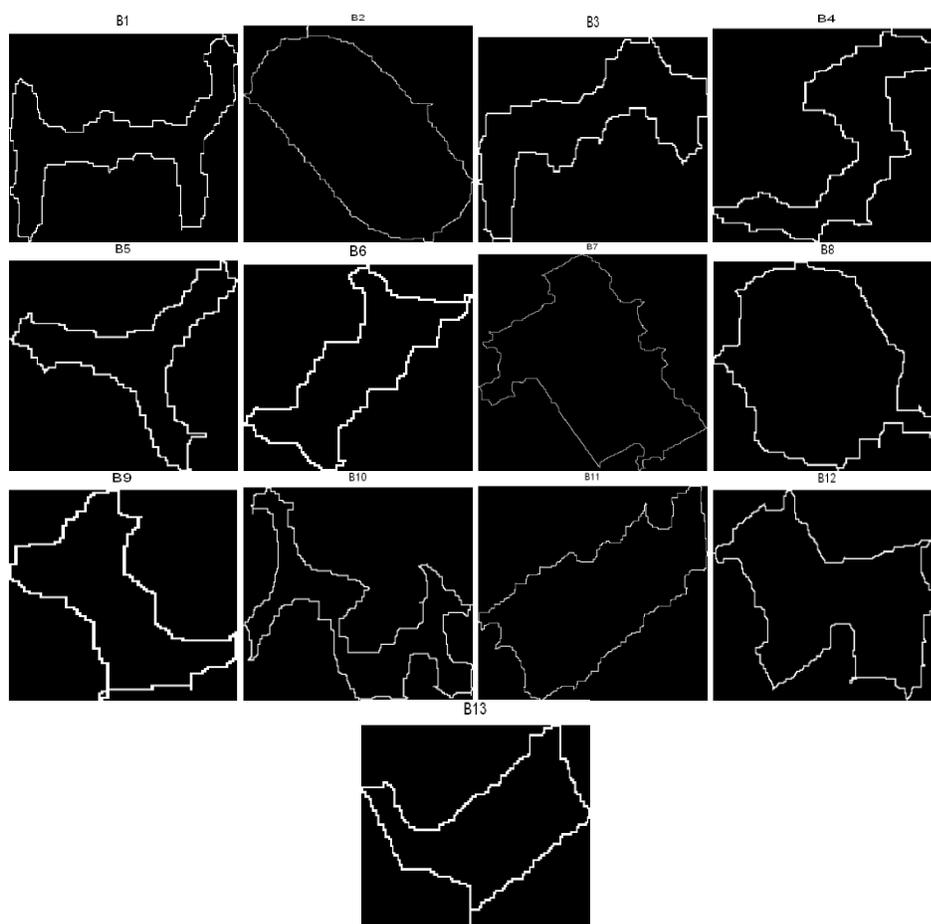


Figure 5-5: Object Boundary Images

Figure 5-6 shows texture object representation after filling object boundary by available data (LIDAR and RGB).

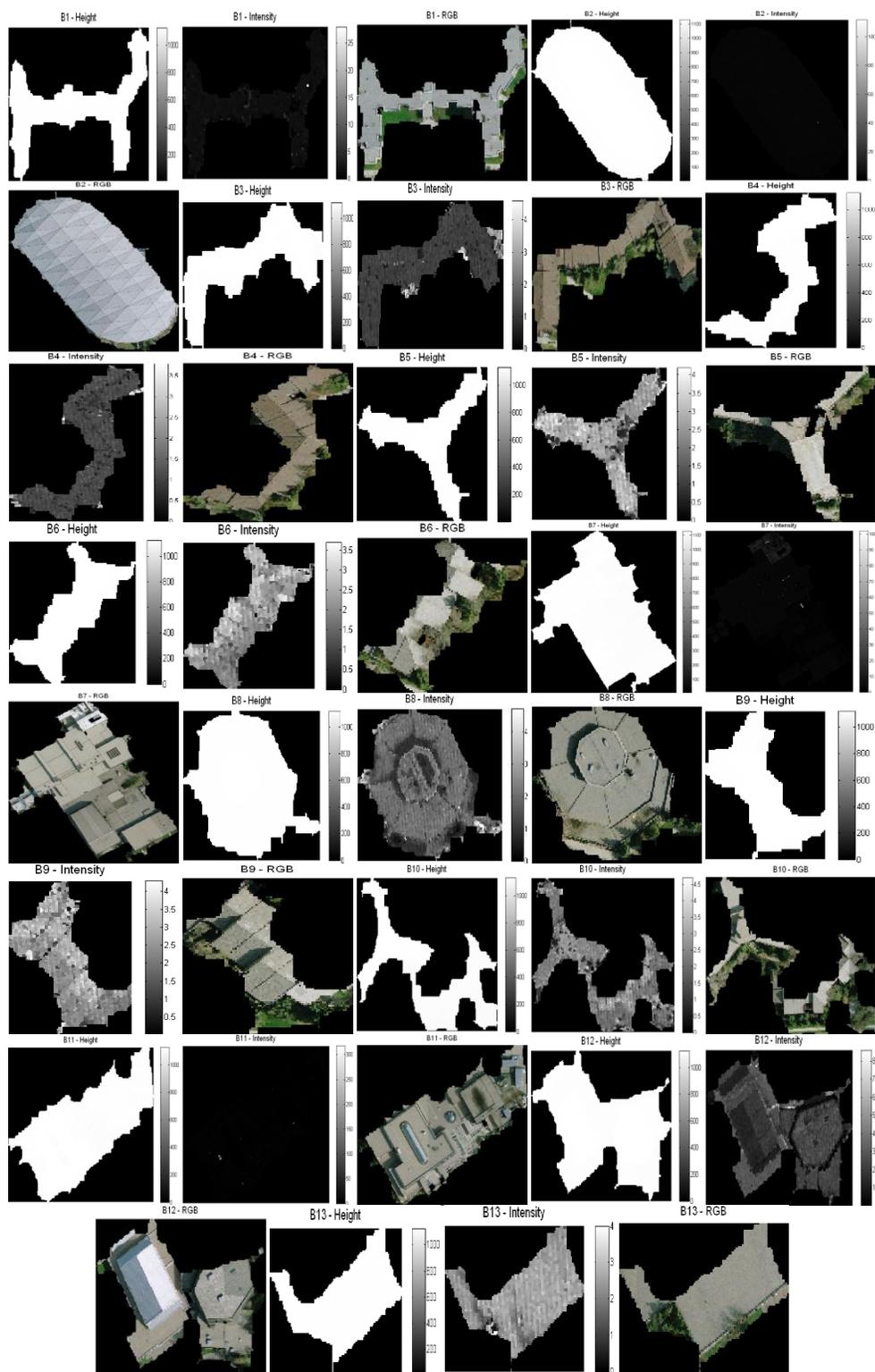


Figure 5-6: Object Texture Representation

5.3.2 Wavelet Object Description

Another approach is the representation of object boundary using continuous wavelet transform. The boundary representative is described using continuous wavelet transform. The first step is to transform the boundary coordinates to a one dimensional function using Equation 5.4 where x represents the x -coordinates and y represents the y -coordinates of the object boundary [Gonzalez et al., 2004].

$$BF(q) = x(q) + i y(q) \quad 5.4$$

where: $q = 1, 2, \dots, bq$

$bq =$ total number of boundary points

$i = \sqrt{-1}$

The produced 1-D function is a simple and requires no need for dyadic decomposition. And in order to test the efficiencies of complex wavelet descriptors, the continuous wavelet transform (Section 4.3.4) is applied on each object function to obtain object wavelet boundary descriptors.

The texture object representations are images, so wavelet texture descriptors are calculated by applying two dimensional discrete wavelet transform (Section 4.3.6) on each texture image (Height, Intensity and RGB).

Four real wavelet functions, Haar, Daubechies, Symlets and Discrete approximation of Meyer, and four complex wavelet functions, two Complex Gaussian and two Complex Morlet are tested for generating object wavelet boundary descriptors. Only the four real wavelet functions are used to generate object wavelet texture height, intensity and RGB

descriptors. Figure 5-7 shows the wavelet functions used for generating object wavelet boundary and texture descriptors.

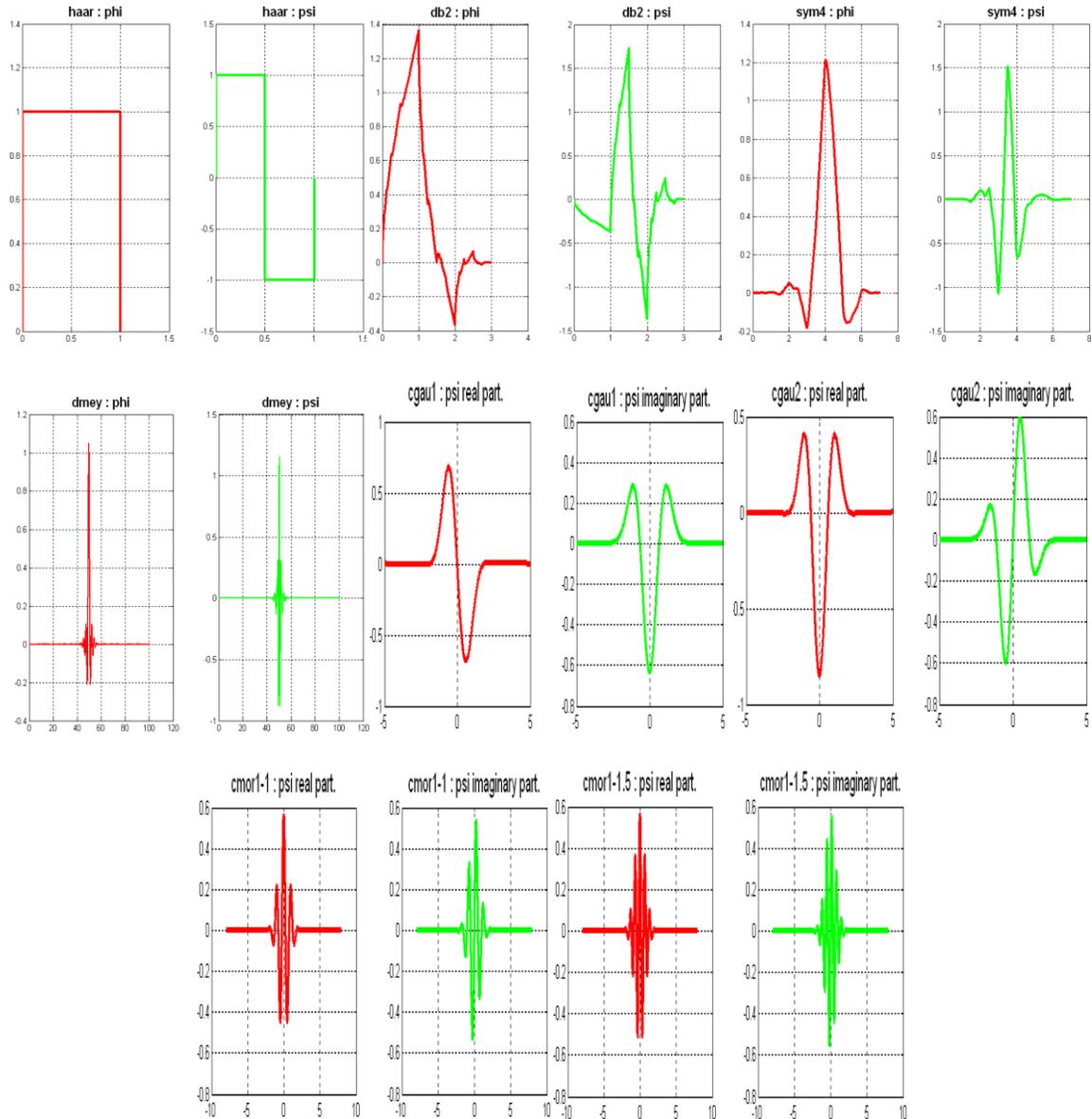


Figure 5-7: Wavelet Function for Object Descriptors

Object wavelet boundary descriptors are a vector of all wavelet coefficients after analysing 1-D boundary by a wavelet transform (CWT, Section 4.3.4). Object wavelet texture descriptors are a vector of all wavelet coefficients after analysing a texture image

by scaling and wavelet functions (2-D DWT, Section 4.3.6). Wavelet coefficients are arranged as four parts in the descriptors vector. The first part is corresponding to the approximation wavelet coefficients, the second part is the horizontal detail wavelet coefficients, the third part is the vertical detail wavelet coefficients and the fourth part is the diagonal detail wavelet coefficients. Each object in the library has four wavelet descriptor vectors, boundary and three textures, in addition to a vector of Monte Carlo Simulation Descriptors (MCSD).

5.3.3 *MCS Object Description*

The object boundary is represented by a 1-D function to calculate wavelet boundary descriptors. However, for calculating MCSD, the object is required to be represented by its boundary image. MCSD is a vector of five elements calculated and corresponding to five different numbers of random values generated and distributed all over the image using Equation 5.5.

$$NRANV(u) = row * col / (u + 1) \quad 5.5$$

where: $NRANV$ = total number of random values

row, col = image dimension

u = the order of the descriptor in MCSD vector

For each number of random values the vector element is calculated as shown in section 5.2.2. For assessing and gathering a trusted conclusion about the efficiency of this descriptor, it is generated eight times for all objects representing eight trials. Table 5-1 shows the MCSD of all library objects for first trial, as an example.

Table 5-1: MCS D for Library Objects – First Trial

Object	MCS D				
B1	0.244	0.269	0.271	0.278	0.266
B2	0.418	0.450	0.471	0.481	0.490
B3	0.274	0.303	0.316	0.302	0.319
B4	0.247	0.269	0.278	0.284	0.275
B5	0.211	0.229	0.232	0.252	0.250
B6	0.251	0.275	0.272	0.288	0.298
B7	0.370	0.402	0.416	0.426	0.433
B8	0.448	0.491	0.510	0.514	0.528
B9	0.307	0.320	0.348	0.339	0.349
B10	0.228	0.247	0.255	0.264	0.262
B11	0.345	0.368	0.384	0.392	0.409
B12	0.383	0.418	0.431	0.442	0.467
B13	0.301	0.328	0.333	0.359	0.345

5.4 Recognition Descriptors Assessment

This section assesses the MCS D versus wavelet descriptors for object shape recognition application. The object classification technique is applied using Minimum Distance

Classifier (MDC) (Section 5.2.3). The object of interest is associated with the class that gives the minimum value. Figure 5-8 shows the general methodology of applying the proposed object recognition technique. This algorithm is suitable for recognizing any object of interest from available data (LIDAR or RGB).

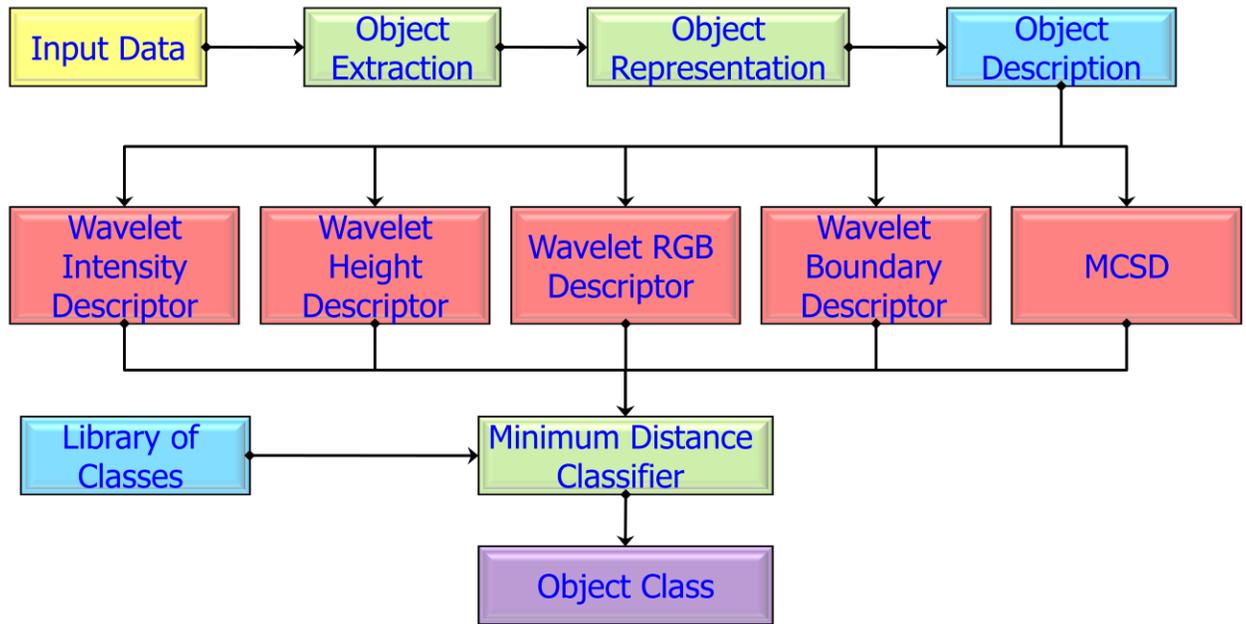


Figure 5-8: Object Shape Recognition Algorithm

Because the main target of this section is the assessment and testing of the efficiency of the descriptors (wavelet and MCSD) for object shape recognition application for buildings, this technique is applied for the recognition of the same objects in the library in two different cases. The first case is the object recognition from the original LIDAR height image, where both the object and the library have the same scale and orientation. The second case is the object recognition from the distorted LIDAR height image after changing scale and orientation. The object recognition algorithm is applied eight times

using eight different wavelet functions for both cases, as shown in Figure 5-7 and eight MCSD to assess the efficiency of all descriptors (wavelet and MCSD).

5.4.1 *Object Recognition from Original LIDAR Height Image*

The first case study is the object recognition from original LIDAR height image without changing scale or orientation. The recognition is applied using the minimum distance classifier technique. Figure 5-9 shows object recognition results for original input objects without change using eight different wavelet functions. Haar, Daubechies, Symlets and Discrete approximation of Meyer, real wavelet functions, are all used to obtain boundary and three different texture descriptors. Two Complex Gaussian and two Complex Morlet, complex wavelet functions, are used to obtain boundary descriptors only. The horizontal axis represents the objects of interest and the vertical axis represents the library of objects. The red square colour marker represents successful recognition objects, and the green circle colour marker represents failed recognition objects. All wavelet descriptors (boundary and texture) succeeded in the recognition process with one hundred percent for all objects of interest.

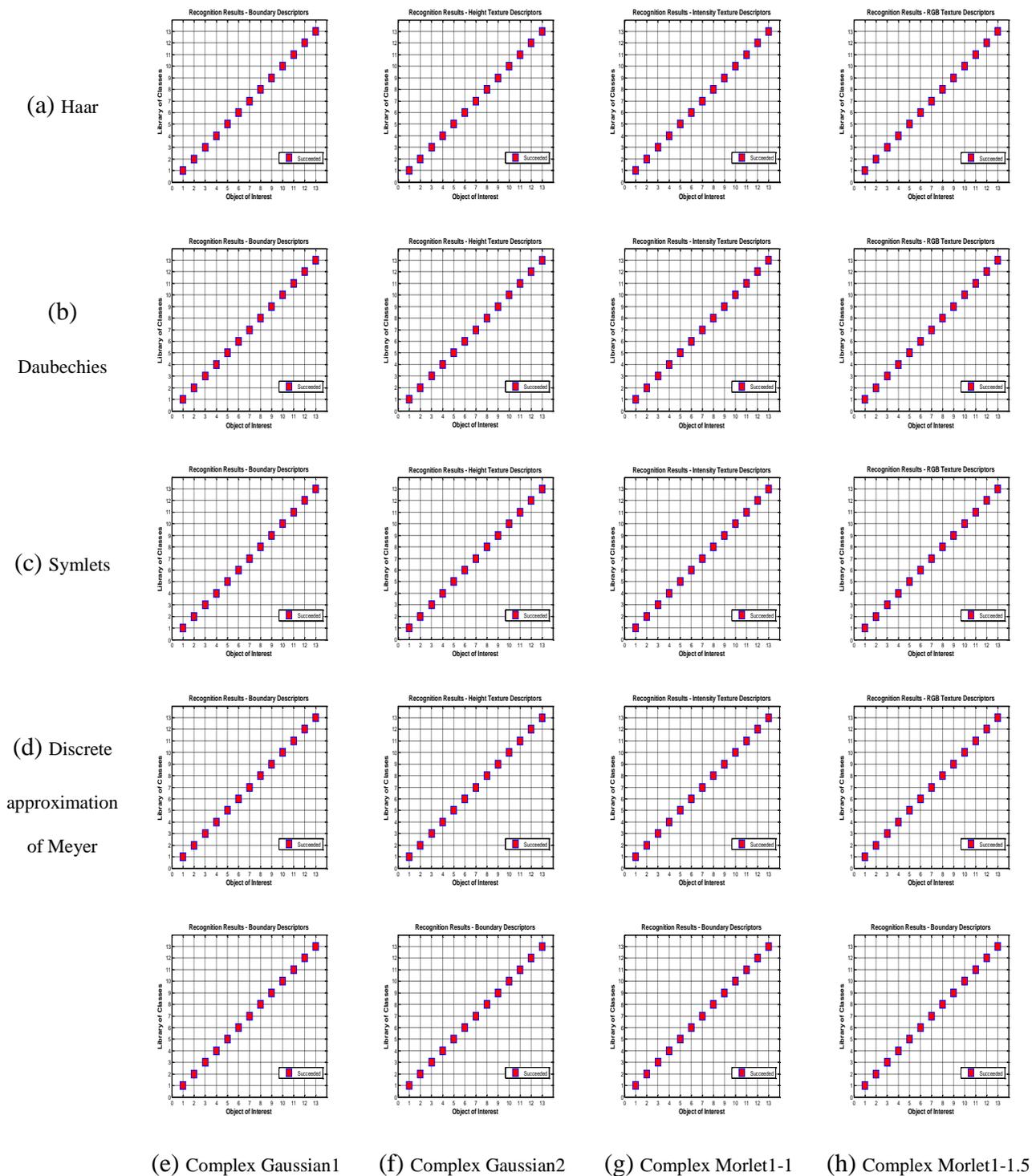


Figure 5-9: Object Recognition Results from Original Input Data using Wavelet Descriptors

As previously mentioned in Section 5.3.3, MCSD is calculated eight times for all objects of interest. Object recognition using MCSD is applied for all eight trials (Figure 5-10). The first trial MCSD succeeded in recognizing nine objects with their correct classes but failed to recognize four objects. The recognition process succeeded with sixty nine percent. The second, fifth and seventh trials MCSD succeeded in recognizing eleven objects with their correct classes but failed to recognize two objects. The recognition process succeeded with eighty four percent. The third trial MCSD succeeded in recognizing ten objects with their correct classes but failed to recognize three objects. The recognition process succeeded with seventy six percent. The fourth, sixth and eighth trials MCSD succeeded in recognizing twelve objects with their correct classes but failed to recognize one object. The recognition process succeeded with ninety two percent. Finally, MCSD succeeded in recognizing objects with general percentage of eighty four as shown in Figure 5-11.

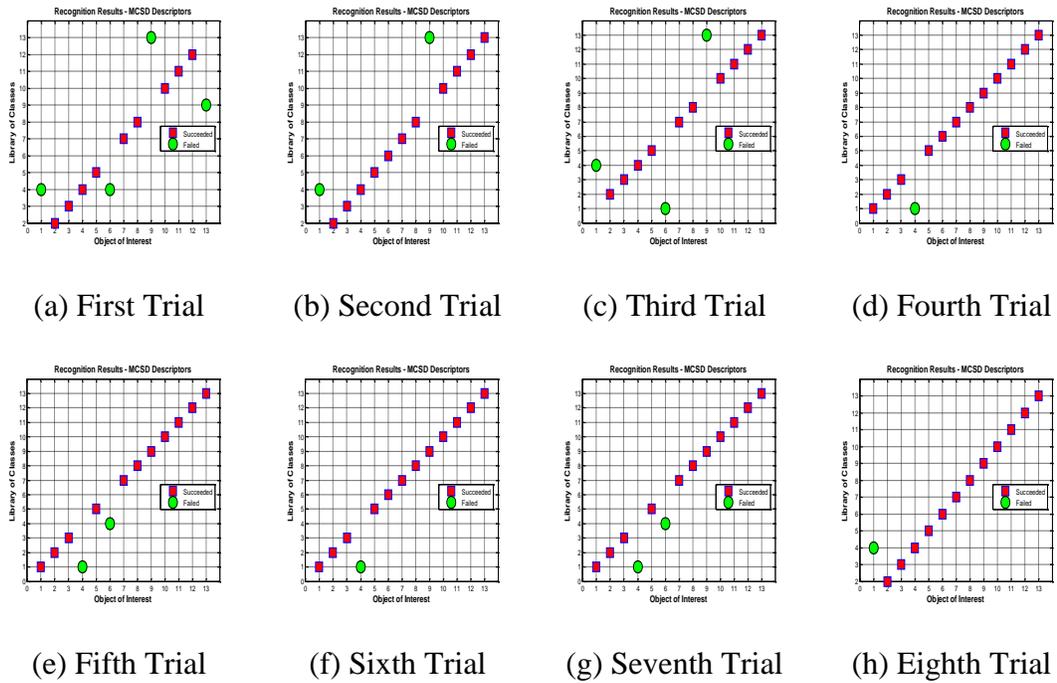


Figure 5-10: Object Recognition Results from Original Input Data using MCSD

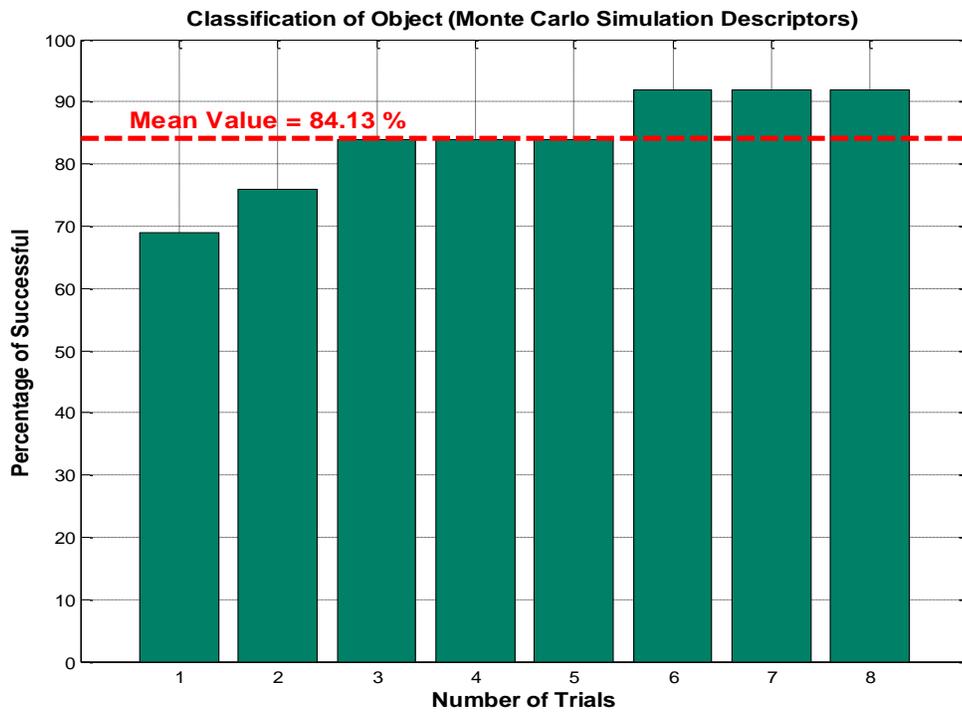


Figure 5-11: MCSD Recognition Successful Percentage form Original Input Data

5.4.2 Object Recognition from Scaled and Oriented Original LIDAR Height Image

The recognition process is repeated for wavelet boundary and texture descriptors after changing the scale and orientation of LIDAR and RGB images. The scaling process is applied to the original height, intensity and RGB images after distorting with scale factor $s_x = 1.5$ in X - direction and $s_y = 2$ in Y - direction. The scaled images are oriented with a rotation angle that equals 90° . Figure 5-12 and Figure 5-13 show the height and RGB images after scaling and orientation processes.

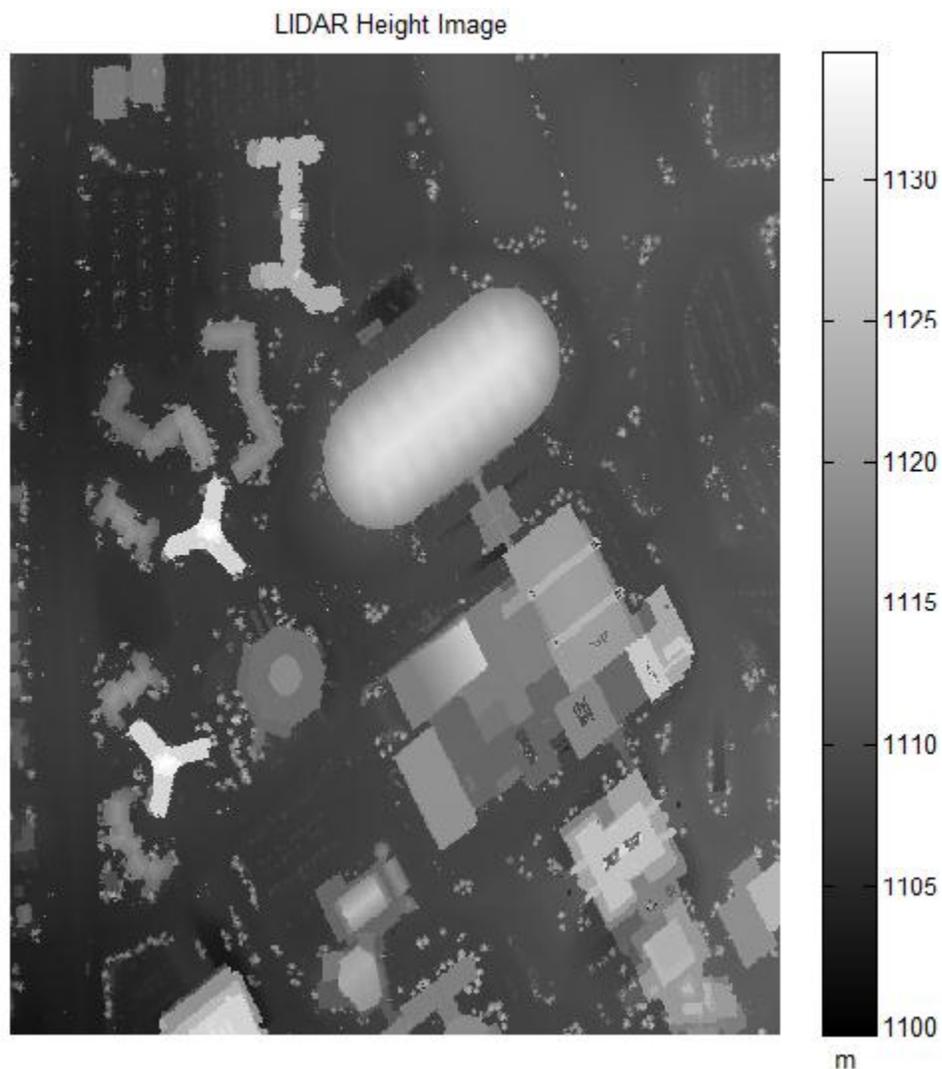


Figure 5-12: Scaled and Oriented LIDAR Height Image

Scaled and Oriented RGB Image**Figure 5-13: Scaled and Oriented RGB Image**

Figure 5-14 shows the recognition results using eight different wavelet functions. Wavelet boundary descriptors succeeded in recognizing one object in cases using the different proposed wavelet functions and failed to recognize the other twelve objects. Wavelet boundary descriptors succeeded in recognizing objects with seven percent. Wavelet texture descriptors failed to recognize all objects using four different wavelet functions in case of height and intensity properties. Wavelet RGB texture descriptors succeeded in recognizing one object and failed to recognize the other twelve objects in cases using three different wavelet functions (Haar, Daubechies, Symlets). Wavelet height and intensity texture descriptors succeeded in recognizing objects with seven percent. In cases using Discrete approximation of Meyer wavelet function, wavelet RGB texture descriptors succeeded in recognizing two objects but failed to recognize other

eleven objects. Discrete approximation of Meyer wavelet RGB texture descriptors succeeded in recognizing objects with fifteen percent. In general, object recognition process succeeds using wavelet descriptors with successful percentage equals to seven percent as shown in Figure 5-15.

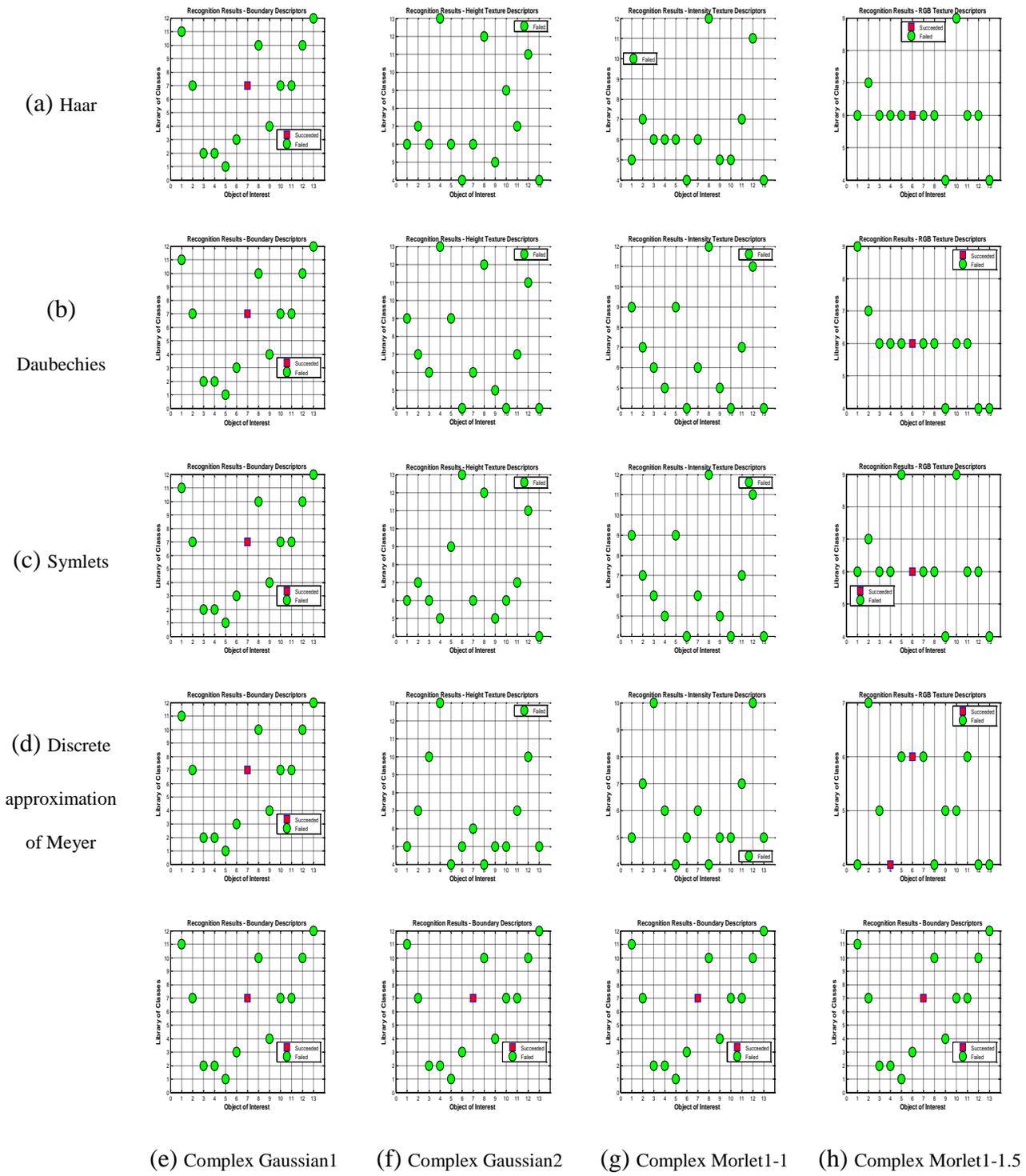


Figure 5-14: Recognition Results for Scaled and Oriented Objects using Wavelet Descriptors

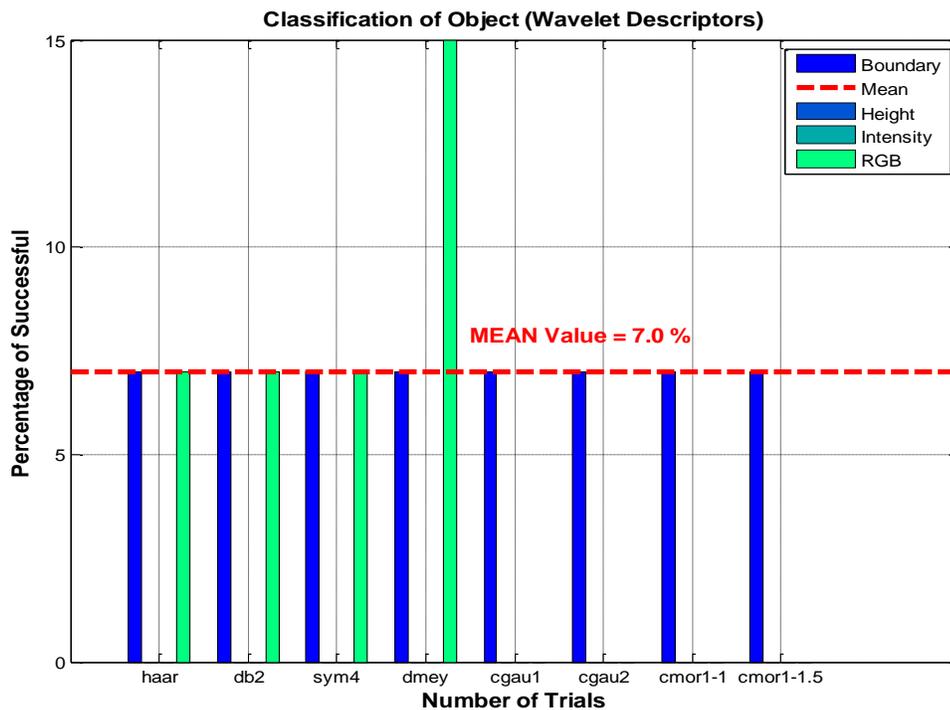


Figure 5-15: Wavelet Descriptors Recognition Successful Percentage for Scaled and Oriented Objects

In the case of MCS D, it succeeded in recognizing twelve objects but failed to recognize just one object in the case of seven trials (from trial two to trial eight) with a successful percentage equal to ninety two percent. In the case of the first trial, MCS D succeeded in recognizing eleven objects but failed to recognize two objects with a successful percentage equal to eighty four percent. Figure 5-16 demonstrates all object recognition results for all eight trials. In general, object recognition process succeeds in recognizing objects using MCS D with successful percentage equals to ninety one percent as shown in Figure 5-17.

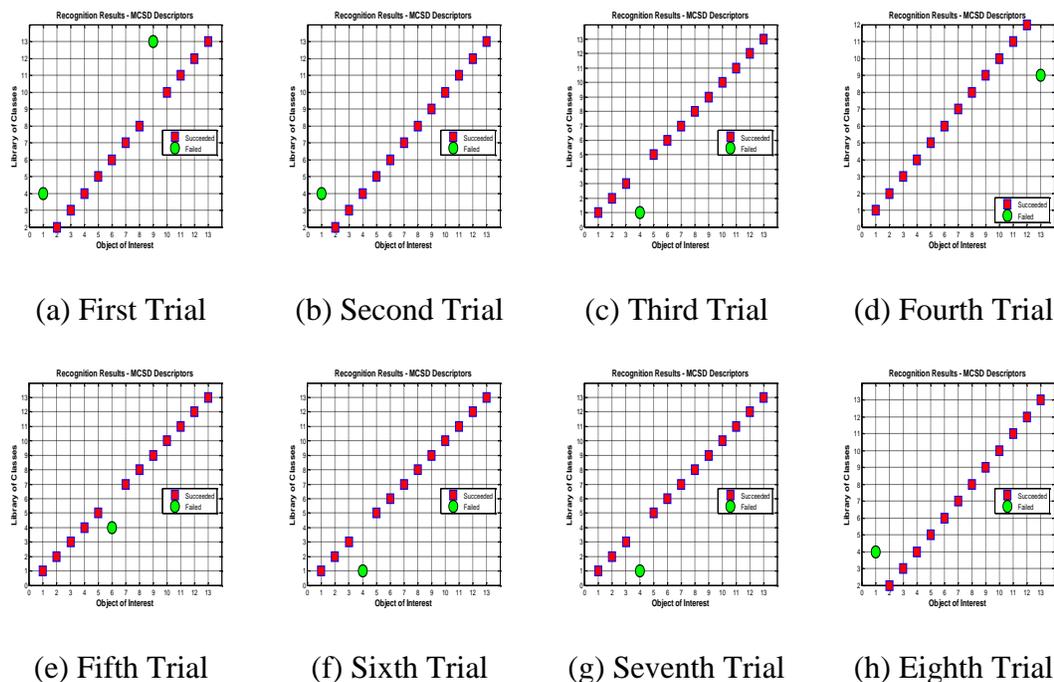


Figure 5-16 Recognition Results for Scaled and Oriented Objects using MCSD

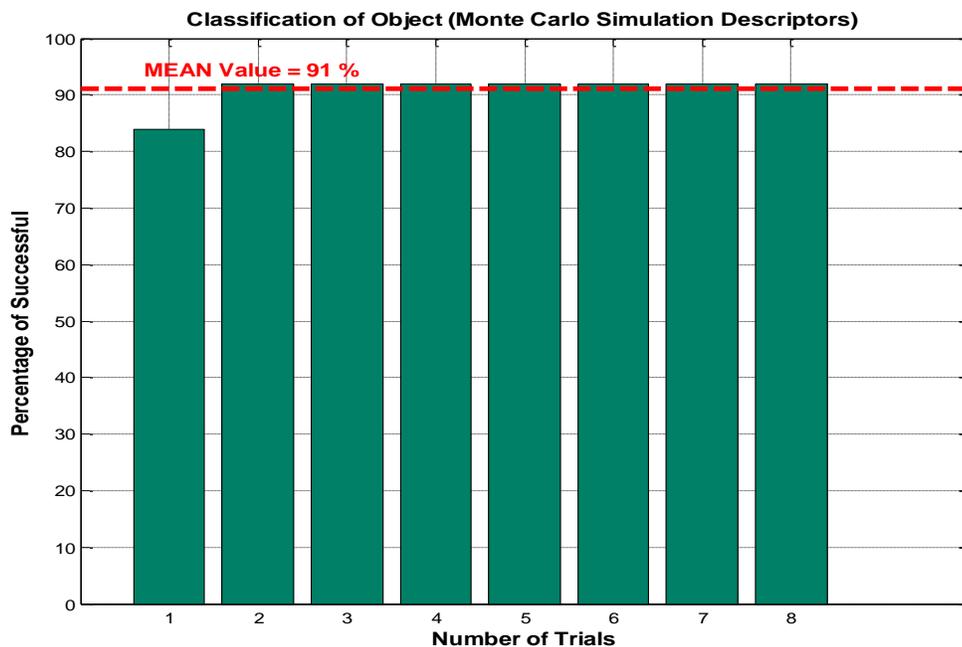


Figure 5-17: MCSD Recognition Successful Percentage for Scaled and oriented Objects

Based on the previous results, it is found that wavelet descriptors are efficient with one hundred percent of cases recognizing objects with the same scale and orientations as library of classes, but cannot be used for recognizing objects in random scales and orientations. In cases of MCSD, there is no need to register the input image referenced by the same scale and orientation as the library of classes. MCSD succeeded in recognizing all different objects in random scale and orientation with a general successful percentage equal to eighty seven percent. MCSD with minimum distance classifier method form an efficient and simple recognition technique regardless of the scale or orientation of the input image.

Chapter Six: Summary, Conclusions, Contributions and Recommendations

Efficient urban classification and recognition techniques are introduced. Feature extraction technique from combined semantic and spatial data overcomes the problems for extracting different features sharing same texture or height information. Proposed feature extraction techniques with recognition technique using Monte Carlo simulation descriptor form an efficient object extraction and recognition technique that is suitable for different features regardless of their shapes, scales or orientations.

6.1 Summary

This research thesis identified four different approaches for introducing complete object extraction, classification and recognition of urban areas. The first approach is for feature classification from an RGB image using a combination between supervised and unsupervised segmentation methods. Different features sharing texture properties are hardly separated from an RGB image.

The second approach is designed for building extraction from LIDAR data based on Multi Resolution Analysis (MRA) using wavelet transform. Trees located close to the buildings and have height information similar to some of the buildings are extracted as building candidates. The third approach used two different data, RGB and LIDAR, combined together to be efficiently extract all urban area features.

The fourth approach for object recognition using new spatial domain descriptor representing objects by boundary images and classifying the object of interest used the minimum distance classifier method. The new spatial domain descriptor is based on the Monte Carlo Simulation (MCS) technique. And the object is described by the ratio between the number of random values generated inside the object boundary and the total

number of random values that cover the object boundary image. The thesis assessed the efficiency of this descriptor versus different wavelet descriptors for object boundary and texture representations and concluded that the proposed descriptor gave better results than wavelet descriptors for object recognition.

6.2 Conclusions and Contributions

This thesis introduced the real automated urban classification technique from airborne LIDAR/Imaging systems. The proposed technique does not require any human interface in the classification process. The new classified features and especially the buildings are ready for commercial use and adaptive and efficient search, through the developed object recognition algorithmic that took into account both the boundary and texture proprieties. The thesis applied traditional segmentation methods with a new implementation approach to classify urban area features from an RGB image. The proposed classification technique introduces an efficient combination of supervised and unsupervised segmentation methods to classify vegetation, roads and buildings in urban areas. *RGB* Color space is suitable for identifying vegetation and shadows areas, but must transform to YC_bC_r and *HSV* domains to identify color ranges used in detecting and separating roads and buildings. The proposed technique succeeded in identifying color ranges that are corresponding to sandy and unhealthy vegetation areas (mixed together). The proposed technique is fully automated and suitable for all urban areas using high resolution RGB images. It has encounters difficulties when dealing with a similarity in texture properties of different features, buildings and roads, which was the motivation for aiding with additional information (spatial data) to extract, separate, and classify different features efficiently.

A standalone technique based on Multi-Resolution Analysis (MRA) using wavelet transform is used for the production of a 3-D building model from LIDAR data. Wavelet is a more sensitive MRA tool for analysing signals with different frequency/space resolutions. LIDAR height image is analysed by wavelet transform which efficiently helped in detecting building edges from LIDAR height image regardless of the building shape or image scale. It is found that trees located close to the buildings with the same height information are extracted as building candidates. This problem was the motivation for aiding the wavelet building extracting technique with semantic information to efficiently separate buildings from LIDAR data.

Although RGB images contain high descriptive information for ground and non-ground features, it is not enough to clearly classify all features. LIDAR data provides important spatial information that helps to distinguish between ground and non-ground objects sharing same texture properties (roads and buildings). The combination of these two types of data, along with the application of the proposed combined filtering statistical wavelet MRA analysis, introduced an effective algorithm for automatic urban feature classification.

In order to utilize the amount of information available from the proposed urban classification technique, a new search algorithm for object (buildings) recognition is introduced based on the Monte Carlo Simulation descriptors. It is also proven that the wavelet descriptors have greater accuracy for object recognition in cases of objects of interest with the same scale and orientations for both the objects used in the search and the developed library. The Monte Carlo Simulation Descriptors (MCSD) are used based on random number generators with a specific distribution function for object recognition

using distorted objects with libraries containing a different scale and orientation than the distorted object. MCSD is used to overcome this drawback and is used to describe the object regardless of the scale or orientation of the input image. It succeeded in recognizing different objects regardless of their shape, edge regularity, scale or orientation. MCSD was combined with minimum distance classifier method to form an effective object recognition technique.

The contributions of this thesis work can be summarized in the following points:

- A new combination of five traditional filtering techniques (Vegetation and Shadow indexes, Otsu segmentation, color transformation and global color thresholding) for urban classification from digital images has been efficiently implemented for features extraction with significant color channels diversity. Each of these techniques was efficiently implemented in a certain order and with specific statistical hypothesis between the three-color channels of RGB images and projected in different color spaces to introduce a complete and exact urban classified features without human intervention. Vegetation areas, buildings, roads, shadows and sandy areas were perfectly separated by the proposed new implementation for the cases of color diversity between these five features. The five features were efficiently classified in the area of study with more than 90% success based on visual assessment. The proposed new methodology and implementation succeeded in overcoming the main drawbacks and disadvantage of all the former digital imaging classification techniques. However, in the case of features with the same semantic properties, which fall in the same color channel, the digital RGB images were effectively aided with spatial information from

LIDAR data for the same area of interest. All the buildings were extracted from the height information (projected in the gray scale domain) with almost 90% success using first generation wavelet transform and the other 10% were mainly due to the interference from trees with almost the same height values. The proposed technique is very efficient to separate and classify buildings from the LIDAR height data using multi-resolution analysis wavelet transform.

- The two previously mentioned approaches, urban classification for RGB images and building separation from LIDAR height data, were combined together to overcome the two main disadvantages of having either same semantic information or same spatial values. The combination of both approaches led to almost over 95% success in separation of different features and in achieving an efficient automatic urban classification not only for areas with significant color diversity between different features, but also for complex images and LIDAR data that has a significant amount of features with common height values and/or color channels values.
- The success of the proposed classification technique led to the presence of a lot of objects that will introduce a huge amount of information that require storage and availability for easy and efficient reuse for some applications like change detection. The proposed classification technique is followed by a fast efficient object recognition technique that can be the main engine for monitoring any changes and developments in urban areas. The success of the new search algorithm was at least 85% with several complex structures (buildings) for both

texture and boundary changes using wavelets and Monte Carlo random numbers generation techniques.

- One of the main contributions of this thesis is the development of a complete software package that is able to introduce the automatic urban classification and recognition implementations efficiently and quickly.

6.3 Recommendations

The recommendation to extend this research thesis is mostly related to an increase in the number of features to be classified and to apply it to a more complex urban structure. Moreover, it is crucial to apply this proposed technique to different data images from different sources such as satellite images. The following list contains some of the interested research topics that are recommended for future work:

- Terrestrial mobile mapping has many important applications based on feature extraction such as traffic monitoring and city development, so it is recommended to test the investigated feature extraction and classification techniques using RGB/LIDAR data captured from terrestrial mobile mapping systems.
- Urban classification technique succeeded in identifying vegetation, shadows, sandy areas, roads and buildings for urban area. Water areas are considered as urban area features, so it is recommended to develop this urban area classification technique for detecting water areas such as lakes and swimming pools
- Wavelet transform succeeded in producing a 3-D building model from LIDAR data. It is recommended to investigate a standalone technique based on the wavelet transform to extract trees and produce Canopy Height Model (CHM) for different types of trees.

- Monte Carlo Simulation Descriptor (MCSD) succeeded in recognizing buildings from airborne LIDAR regardless the object shape, scale or orientation. It is recommended for assessing MCSD by testing its efficiency for recognizing different urban area objects such as cars and vehicles.

References

- Al-Durgham, M. (2007) *Alternative Methodologies for the Quality Control of Lidar Systems*, M.Sc., University of Calgary (Canada), Canada
- Ali, A., S. A. M. Gilani, and N. A. Memon (2008) "Affine Invariant Contour Descriptors Using Independent Component Analysis and Dyadic Wavelet Transform," *Journal of Computing and Information Technology - CIT*, vol 16, no Copyright 2009, The Institution of Engineering and Technology, pp. 169-181
- Bartels, M., and H. Wei (2006) "Towards Dtm Generation from Lidar Data in Hilly Terrain Using Wavelets," *Proceedings of 4th International Workshop on Pattern Recognition in Remote Sensing in conjunction with the 18th International Conference on Pattern Recognition*, Hong Kong, China, 33-36
- Bong, D. B. L., K. C. Lai, and A. Joseph (2009) "Automatic Road Network Recognition and Extraction for Urban Planning," *Proceedings of World Academy of Science: Engineering & Technology*, vol 53, pp. 209-215
- Boulinguez, D., and A. Quinquis (1999) "Underwater Buried Object Recognition Using Wavelet Packets and Fourier Descriptors," in *Image Analysis and Processing, 1999. Proceedings. International Conference on*, 1999, 478-483
- Boyer, K. L., and C. Unsalan (2005) "A System to Detect Houses and Residential Street Networks in Multispectral Satellite Images," *Computer Vision and Image Understanding*, vol 98, no 3, pp. 423-461
- Butkiewicz, T., R. Chang, Z. Wartell, and W. Ribarsky (2008) "Visual Analysis for Live Lidar Battlefield Change Detection," Orlando, FL, United states, pp. The International Society for Optical Engineering (SPIE), SPIE
- Cheng, H., J. W. Tian, J. Liu, and Q. Z. Yu (2004) "Wavelet Domain Image Denoising Via Support Vector Regression," *Electronics Letters*, vol 40, no 23, pp. 1479-1481
- Dance, C., J. Willamowski, L. Fan, C. Bray, and G. Csurka (2004) "Visual Categorization with Bags of Keypoints," in *ECCV International Workshop on Statistical Learning in Computer Vision*
- Dong, H., J. Li, and M. A. Chapman. (2007) "Semi-Automated Extraction of Urban Highway Intersections from Ikonos Imagery," in *Advances in Mobile Mapping Technology*, Taylor & Francis, London, 139-145
- El-Sheimy, N., C. Valeo, and A. Habib (2005) *Digital Terrain Modeling : Acquisition, Manipulation, and Applications*, Artech House, Boston

Elhabiby, M. (2007) *Wavelet Representation of Geodetic Operators*, Ph.D., University of Calgary (Canada), Canada

Falkowski, M. J., A. M. S. Smith, A. T. Hudak, P. E. Gessler, L. A. Vierling, and N. L. Crookston (2006) "Automated Estimation of Individual Conifer Tree Height and Crown Diameter Via Two-Dimensional Spatial Wavelet Analysis of Lidar Data," *Canadian Journal of Remote Sensing*, vol 32, no 2, April, pp. 153-161

Ghanma, M. (2006) *Integration of Photogrammetry and Lidar*, Ph.D., University of Calgary (Canada), Canada

Gitelson, A. A., Y. J. Kaufman, R. Stark, and D. Rundquist (2002) "Novel Algorithms for Remote Estimation of Vegetation Fraction," *Remote Sensing of Environment*, vol 80, no 1, pp. 76-87

Gonzalez, R. C., and R. E. Woods (2002) *Digital Image Processing*, Prentice Hall, Upper Saddle River (NJ)

Gonzalez, R. C., R. E. Woods, and S. L. Eddins (2004) *Digital Image Processing Using Matlab*, Pearson Prentice Hall, Upper Saddle River, N.J.

Graps, A. (1995) "An Introduction to Wavelets," *Computational Science & Engineering, IEEE*, vol 2, no 2, pp. 50-61

Gronwall, C., F. Gustafsson, and M. Millnert (2006) "Ground Target Recognition Using Rectangle Estimation," *Image Processing, IEEE Transactions on*, vol 15, no 11, pp. 3400-3408

Habib, A., K. I. Bang, A. P. Kersting, and D. C. Lee (2009) "Error Budget of Lidar Systems and Quality Control of the Derived Data," *Photogrammetric Engineering and Remote Sensing*, vol 75, no 9, Sep, pp. 1093-1108

Huber, T. (1997) "Introduction to Monte Carlo Simulation," Physics Department, Gustavus Adolphus College, <http://physics.gac.edu/~huber/envision/instruct/montecar.htm>

Hundeishausen, F. V., and M. Veloso (2007) "Active Monte Carlo Recognition," Bremen, Germany, pp. 229-243, Springer Verlag

Jansen, M., and A. Bultheel (1999) "Multiple Wavelet Threshold Estimation by Generalized Cross Validation for Images with Correlated Noise," *IEEE Transactions on Image Processing*, vol 8, no 7, pp. 947-953

Jensen, J. R. (2005) *Introductory Digital Image Processing : A Remote Sensing Perspective*, Prentice Hall, Upper Saddle River, N.J.

Kanungo, T., D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu (2002) "An Efficient K-Means Clustering Algorithm: Analysis and Implementation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol 24, no 7, pp. 881-892

Keller, W. (2004) *Wavelets in Geodesy and Geodynamics*, Walter de Gruyter, Berlin; w York

Kim, C., A. Habib, and Y.-C. Chang (2008) "Automatic Generation of Digital Building Models for Complex Structures from Lidar Data," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol XXXVII, no B4, July, pp. 456-462

Kim, C., A. Habib, and P. Mrsitik (2007) "New Approach for Planar Patch Segmentation Using Airborne Laser Data," *ASPRS—American Society for Photogrammetry and Remote Sensing, Annual Conference*, May 7-11, 2007, Tampa, Florida

Kunttu, I., L. Lepisto, J. Rauhamaa, and A. Visa (2003) "Multiscale Fourier Descriptor for Shape Classification," in *Image Analysis and Processing, 2003.Proceedings. 12th International Conference on*, 17-19 Sept. 2003 536-541

Laptev, I., H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner (2000) "Automatic Extraction of Roads from Aerial Images Based on Scale Space and Snakes," *Machine Vision and Applications*, vol 12, no 1, pp. 23-31

Li, J., Y. Lia, and M. A. Chapmanb (2007a) "Building Edge Extraction from Lidar Based on Jump Detection in Non-Parameter Regression Model," *The 5th International Symposium on Mobile Mapping Technology (MMT'07)*, 29-31 May, Padua, Italy

Li, Y., J. Li, and M. Chapman. (2007b) "A Fuzzy Relational Method for Image-Based Road Extraction for Traffic Emergency Services," in *Geomatics Solutions for Disaster Management*, Li, J., S. Zlatanova, and A. G. Fabbri, eds., Springer Berlin Heidelberg, 37-48

Liang, Q. (2008) "Automatic Target Recognition Using Waveform Diversity in Radar Sensor Networks," *Pattern Recognition Letters*, vol 29, no 3, pp. 377-381

Mahmoud, S. A., and A. S. Mahmoud (2006) "Arabic Character Recognition Using Modified Fourier Spectrum (Mfs)," in *Geometric Modeling and Imaging--New Trends, 2006*, 16-18 Aug. 1993 155-159

Mena, J. B., and J. A. Malpica (2005) "An Automatic Method for Road Extraction in Rural and Semi-Urban Areas Starting from High Resolution Satellite Imagery," *Pattern Recogn. Lett.*, vol 26, no 9, pp. 1201-1220

Mokhtarzade, M., and M. J. V. Zoej (2007) "Road Detection from High-Resolution Satellite Images Using Artificial Neural Networks," *International Journal of Applied Earth Observation and Geoinformation*, vol 9, no 1, pp. 32-40

Nabout, A., and B. Tibken (2004) "Object Recognition Using Contour Polygons and Wavelet Descriptors," in *Information and Communication Technologies: From Theory to Applications, 2004. Proceedings. 2004 International Conference on*, 19-23 April 2004, pp. 473-474

Nabout, A. A., and B. Tibken (2007) "Object Shape Recognition Using Mexican Hat Wavelet Descriptors," Piscataway, NJ, USA, pp. 1313-1318, IEEE

Nabout, A. A., and B. Tibken (2008) "Object Shape Description Using Haar- Wavelet Functions," Piscataway, NJ, USA, pp. 1085-1090, IEEE

Nardinocchi, C., M. Scaioni, and G. Forlani (2001) "Building Extraction from Lidar Data," Piscataway, NJ, USA, pp. 79-84, IEEE

Ooi, W. S., and C. P. Lim. (2006) "Hybrid Image Segmentation Based on Fuzzy Clustering Algorithm for Satellite Imagery Searching and Retrieval," in *Applied Soft Computing Technologies: The Challenge of Complexity*, Abraham, A., B. de Baets, M. Köppen, and B. Nickolay, eds., Springer Berlin / Heidelberg, 355-372

Osowski, S., and D. D. Nghia (2002) "Fourier and Wavelet Descriptors for Shape Recognition Using Neural Networks--a Comparative Study," *Pattern Recognition*, vol 35, no 9, pp. 1949-1957

Otsu, N. (1979) "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol 9, no 1, pp. 62-66

Pradhan, B., S. Mansor, A. R. Ramli, A. R. B. Mohamed Sharif, and K. Sandeep (2005) "Lidar Data Compression Using Wavelets," USA, pp. 598305-598301, SPIE - The International Society for Optical Engineering

Rottensteiner, F., and C. Briese (2002) "A New Method for Building Extraction in Urban Areas from High-Resolution Lidar Data," *International Archives of Photogrammetry and Remote Sensing (IAPRS)*, vol 34, no 3A, pp. 295-301

Rottensteiner, F., J. Trinder, S. Clode, and K. Kubik (2003) "Building Detection Using Lidar Data and Multispectral Images," CSIRO

Ruel, S., C. E. English, L. Melo, A. Berube, D. Aikman, A. M. Deslauriers, P. M. Church, and J. Maheux (2004) "Field Testing of a 3d Automatic Target Recognition and Pose Estimation Algorithm," USA, pp. 102-111, SPIE-Int. Soc. Opt. Eng

Sandau, R., ed. (2010) "Digital Airborne Camera, Introduction and Technology," Springer Netherlands

Schwalbe, E., H.-G. Maas, and F. Seidel (2005) "3d Building Model Generation from Airborne Laser Scanner Data Using 2d Gis Data and Orthogonal Point Cloud Projections," *ISPRS WG III/3, III/4, V/3 Workshop on Laser Scanning*, September 12-14, 2005, Enschede, the Netherlands, 12-14

Shan, J., and C. K. Toth (2009) *Topographic Laser Ranging and Scanning: Principles and Processing*, CRC Press/Taylor & Francis Group

Shorter, N., and T. Kasparis (2009) "Automatic Vegetation Identification and Building Detection from a Single Nadir Aerial Image," *Remote Sensing*, vol 1, no 4, pp. 731-757

Sirmacek, B., and C. Unsalan (2008) "Building Detection from Aerial Images Using Invariant Color Features and Shadow Information," in *23rd International Symposium on Computer and Information Sciences, 2008. ISCIS '08.*, 27-29 Oct. 2008, 1-5

Smach, F., C. Lemaître, J.-P. Gauthier, J. Miteran, and M. Atri (2008) "Generalized Fourier Descriptors with Applications to Objects Recognition In svm Context," *Journal of Mathematical Imaging and Vision*, vol 30, no 1, pp. 43-71

Sohn, G., and I. Dowman (2003) "Building Extraction Using Lidar Dems and Ikonos Images," *International Archives of Photogrammetry and Remote Sensing, WG III/3 Workshop on 3-D Reconstruction from Airborne Laserscanner and InSAR Data*, vol 34, no 3/W13, 08–10 October.,

Song, Y., and J. Shan (2008) "Buiding Extraction from High Resolution Color Imagery Based on Edge Flow Driven Active Contour and Jseg," *Remote Sensing and Spatial Information Sciences*, vol XXXVII, no Part B3a, Beijing, 2008, pp. 185 - 190

Sungho, K., J. Gijeong, and K. In So (2005) "An Effective 3d Target Recognition Model Imitating Robust Methods of the Human Visual System," *Pattern Analysis and Applications*, vol 8, no 3, pp. 211-226

Tate, N. J., C. Brunsdon, M. Charlton, A. S. Fotheringham, and C. H. Jarvis (2005) "Smoothing/Filtering Lidar Digital Surface Models. Experiments with Loess Regression and Discrete Wavelets," *Journal of Geographical Systems*, vol 7, no 3-4, pp. 273-290

Tiehua, D., L. Kah Bin, H. Geok Soon, Y. Wei Miao, and Z. Hao (2004) "2d Occluded Object Recognition Using Wavelets," in *Computer and Information Technology, 2004. CIT '04. The Fourth International Conference on*, 14-16 Sept. 2004, 227-232

Tieng, Q. M., and W. W. Boles (1997) "Recognition of 2d Object Contours Using the Wavelet Transform Zero-Crossing Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 19, no 8, pp. 910-916

Tu, Z. W., and R. X. Li (2002) "A Framework for Automatic Recognition of Spatial Features from Mobile Mapping Imagery," *Photogrammetric Engineering and Remote Sensing*, vol 68, no 3, Mar, pp. 267-276

Tuncer, O. (2007) "Fully Automatic Road Network Extraction from Satellite Images," Piscataway, NJ, USA, pp. 708-714, IEEE

Vain, A., Y. Xiaowei, S. Kaasalainen, and J. Hyypä (2010) "Correcting Airborne Laser Scanning Intensity Data for Automatic Gain Control Effect," *Geoscience and Remote Sensing Letters, IEEE*, vol 7, no 3, July, 2010, pp. 511-514

Vosselman, G., and D. Sander (2001) "3d Building Model Reconstruction from Point Clouds and Ground Plans," *International Archives of Photogrammetry and Remote Sensing*, vol XXXIV, no Part 3/W4, 22-24 Oct., pp. 37-43

Vu, T. T., and M. Tokunaga (2001) "Wavelet and Scale-Space Theory in Segmentation of Airborne Laser Scanner Data " *22nd Asian Conference on Remote Sensing*, 5-9 November,, Singapore

Vu, T. T., and M. Tokunaga (2002) "Wavelet-Based Clustering Method to Detect Building in Urban Area from Airborne Laser Scanner Data," In: *MapAsia 2002*, <http://www.gisdevelopment.net/technology/ip/techip009.htm>

Vu, T. T., and M. Tokunaga (2004) "Filtering Airborne Laser Scanner Data: A Wavelet-Based Clustering Method," *Photogrammetric Engineering and Remote Sensing*, vol 70, no 11, pp. 1267-1274

Wang, C.-K., and P.-H. Hsu (2006) "Building Extraction from Lidar Data Using Wavelet Analysis," *27th Asian Conference on Remote Sensing*, Ulaanbaatar, Mongolia

Wegner, J. D., U. Soergel, and A. Thiele (2009) "Building Extraction in Urban Scenes from Highresolution Insar Data and Optical Imagery," Shanghai, China, pp., IEEE Computer Society

Wei, H., and M. Bartels (2006) "Unsupervised Segmentation Using Gabor Wavelets and Statistical Features in Lidar Data Analysis," Hong Kong, China, pp. 667-670, Institute of Electrical and Electronics Engineers Inc.

Wong, A., A. Mishra, P. Fieguth, and D. Clausi (2008) "An Adaptive Monte Carlo Approach to Nonlinear Image Denoising," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 8-11 Dec. 2008:1-4

Xu, R., and S. N. Pattanaik (2005) "A Novel Monte Carlo Noise Reduction Operator," *Computer Graphics and Applications, IEEE*, vol 25, no 2, pp. 31-35

Yin, S., and W. Wang (2006) "Denoising Lidar Signal by Combining Wavelet Improved Threshold with Wavelet Domain Spatial Filtering," *Chinese Optics Letters*, vol 4, no 12, pp. 694-696

ZENG, Q. (2008) "Data Filtering and Feature Extraction of Urban Typical Objects from Airborne Lidar Point Cloud," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol XXXVII. Part B3b, no B3b, pp. 153-156

APPENDICES

Appendix A: RGB/LIDAR Data used for Implementation of Automatic Urban Area Classification Techniques

A-1 Images for Urban Area Classification from RGB Data Alone

This section shows the different RGB images used for in the assessment and development of the new urban area feature classification technique from RGB images and their classification results. These images are used for identifying the optimum methods to be combined inside the new classification technique for detecting vegetation and shadow areas. It is found that Otsu segmentation algorithm is the optimum method for detecting vegetation and shadow areas after applying on color invariant images produced from RGB color space. These images are also used for determining, through continuous testing with different values, the thresholding values for detecting road, sandy and unhealthy vegetation areas after transforming RGB image to YC_bC_r and HSV color spaces. The thresholding values are implemented after applying on Luminance, Hue and Saturation color channels.

Input Image



Figure A-1: Input RGB Image – im1

Segmentation Image

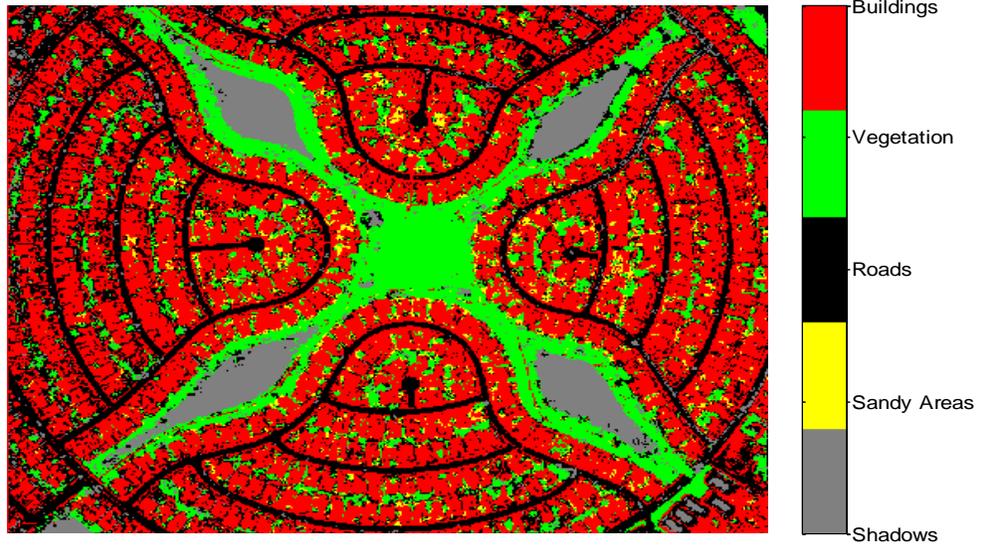


Figure A-2: Classification Results of Image – im1

Input Image



Figure A-3: Input RGB Image – im2

Segmentation Image

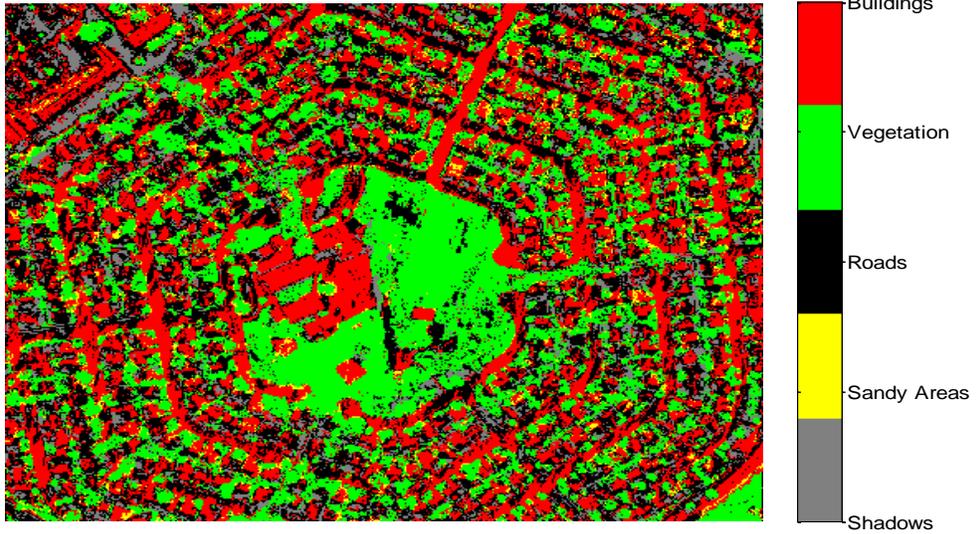


Figure A-4: Classification Results of Image – im2

Input Image



Figure A-5: Input RGB Image – im3

Segmentation Image

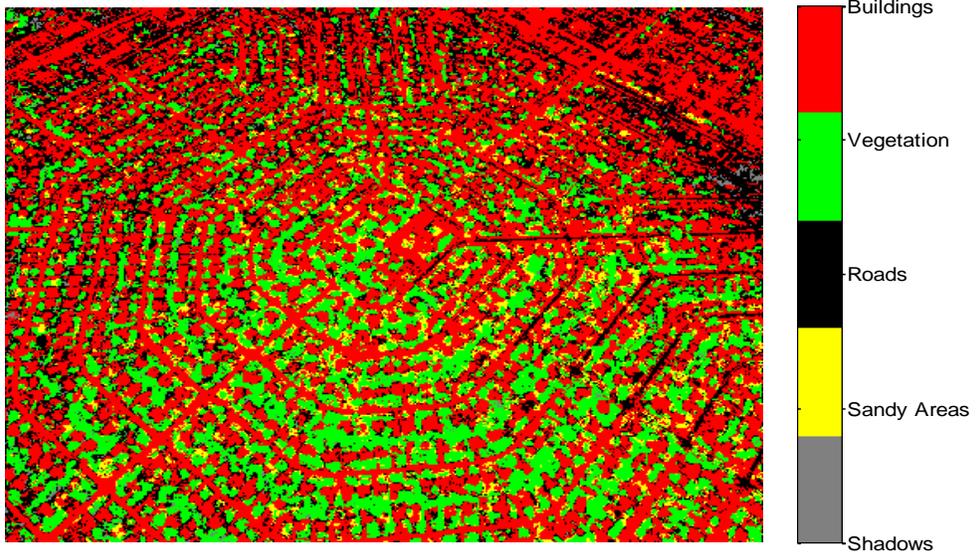


Figure A-6: Classification Results of Image – im3

Input Image



Figure A-7: Input RGB Image – im4

Segmentation Image

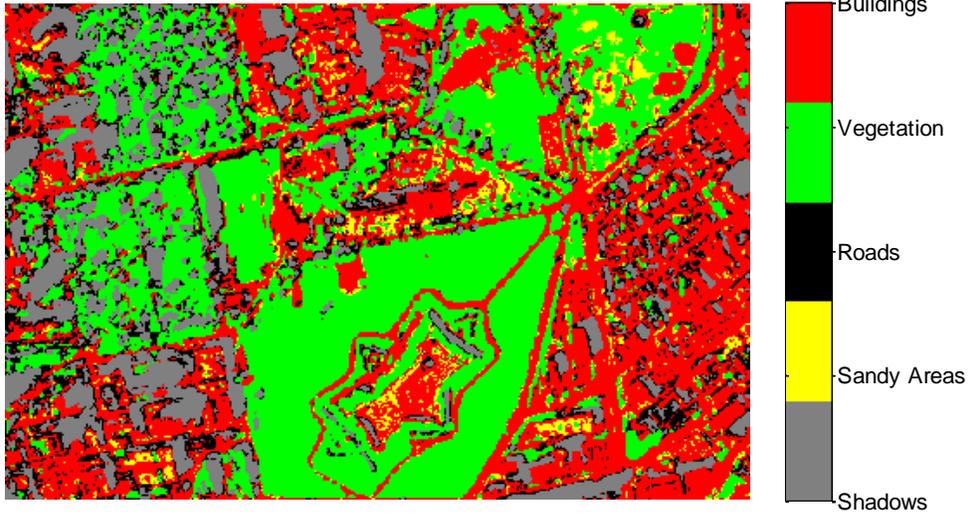


Figure A-8: Classification Results of Image – im4

Input Image



Figure A-9: Input RGB Image – im5

Segmentation Image

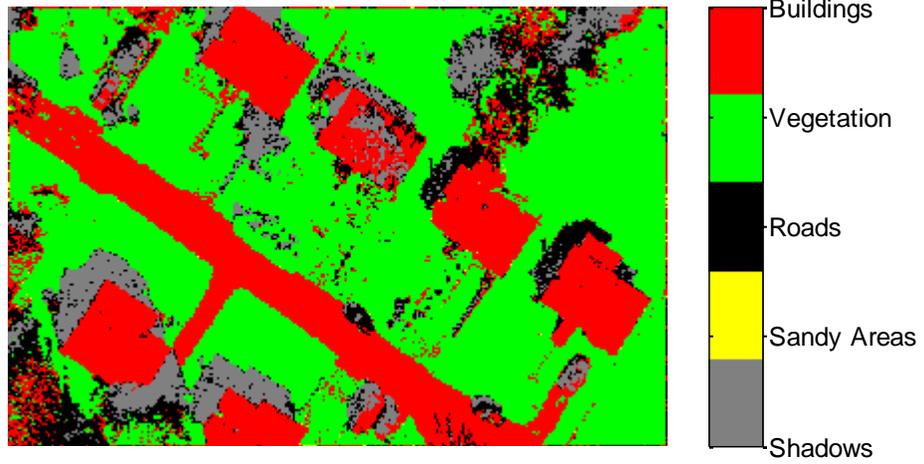


Figure A-10: Classification Results of Image – im5

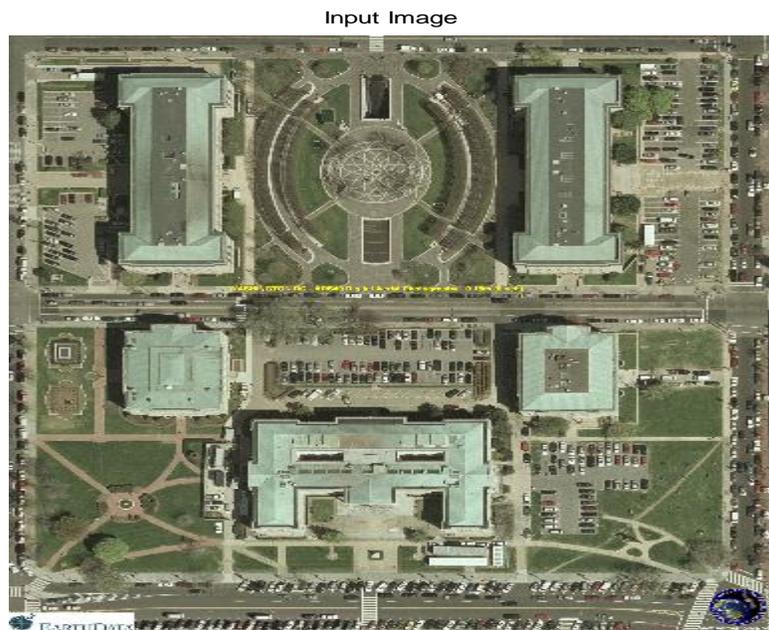


Figure A-11: Input RGB Image – im6

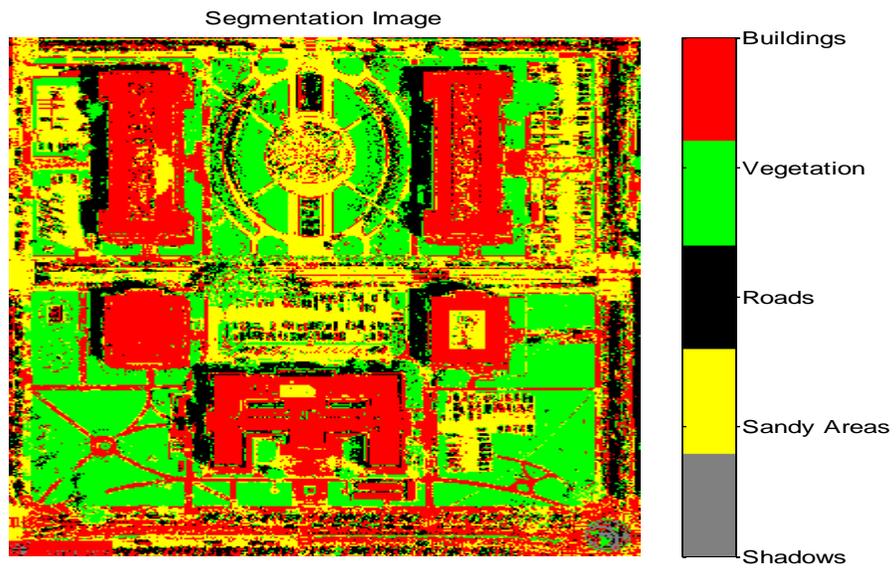


Figure A-12: Classification Results of Image – im6

Input Image



Figure A-13: Input RGB Image – im7

Segmentation Image

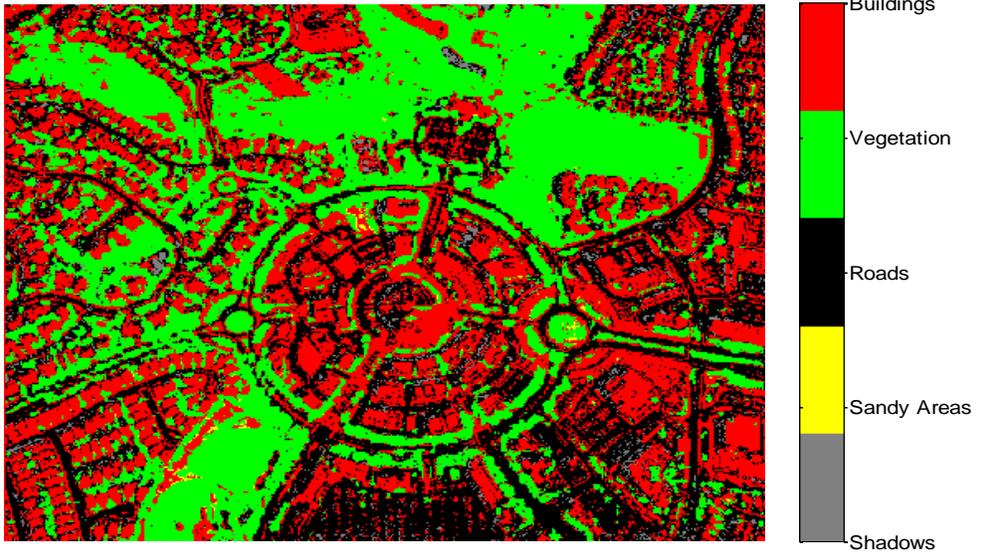


Figure A-14: Classification Results of Image – im7

Input Image



Figure A-15: Input RGB Image – im8

Segmentation Image

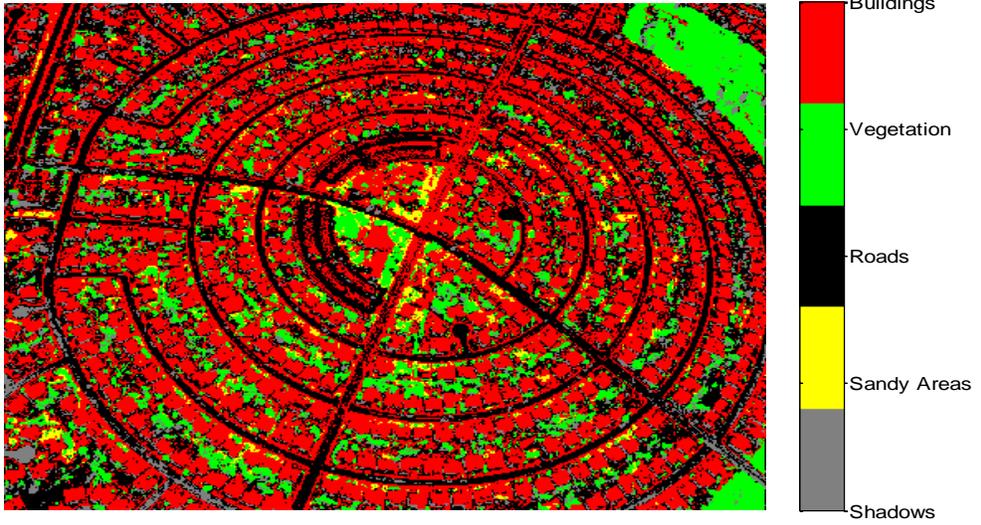


Figure A-16: Classification Results of Image – im8



Figure A-17: Input RGB Image – im9

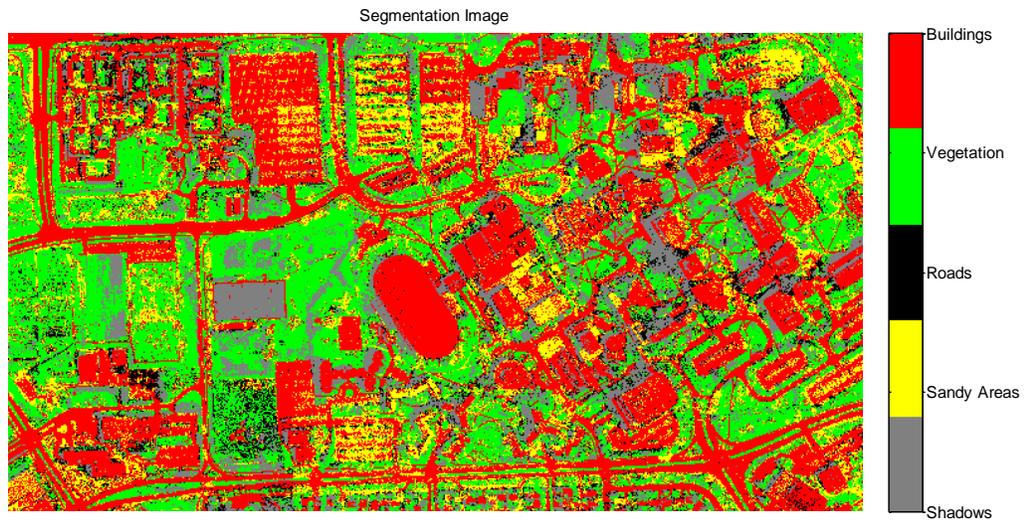


Figure A-18: Classification Results of Image – im9

The proposed urban area classification technique achieved excellent results in the case of having high differentiable texture properties, but it faces difficulties in separating features sharing the same texture properties, and this is the motivation for aiding RGB image with spatial information from LIDAR Data, which was developed in this research thesis.

A-2 RGB/LIDAR Data for the developed Urban Area Classification technique

Figure A-19 shows the LIDAR data for the same area of interest shown in Figure A-17.

The proposed algorithm succeeded in classifying urban area features effectively, especially the features sharing same texture properties (roads and buildings), by aiding with the height LIDAR information shown in Figure A-20.

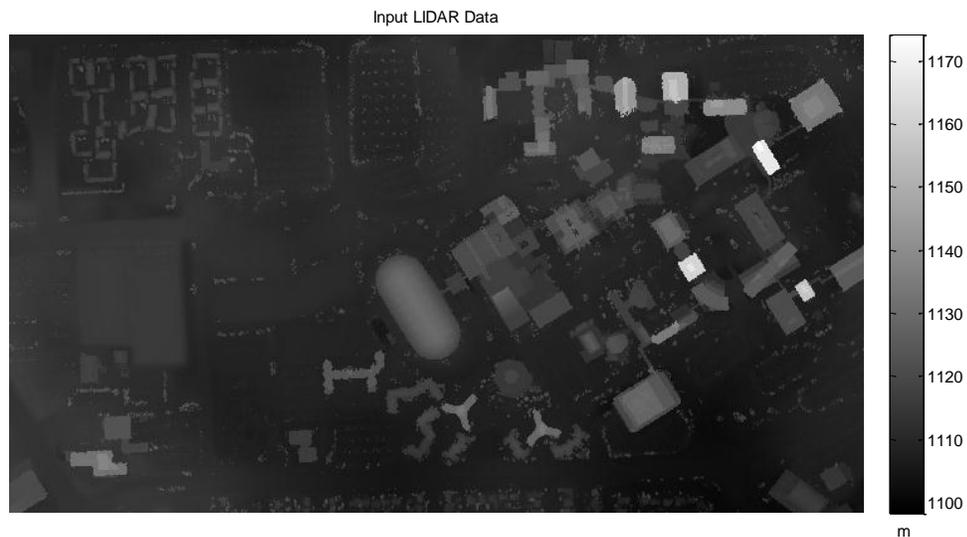


Figure A-19: Input LIDAR Data for RGB Image – im9

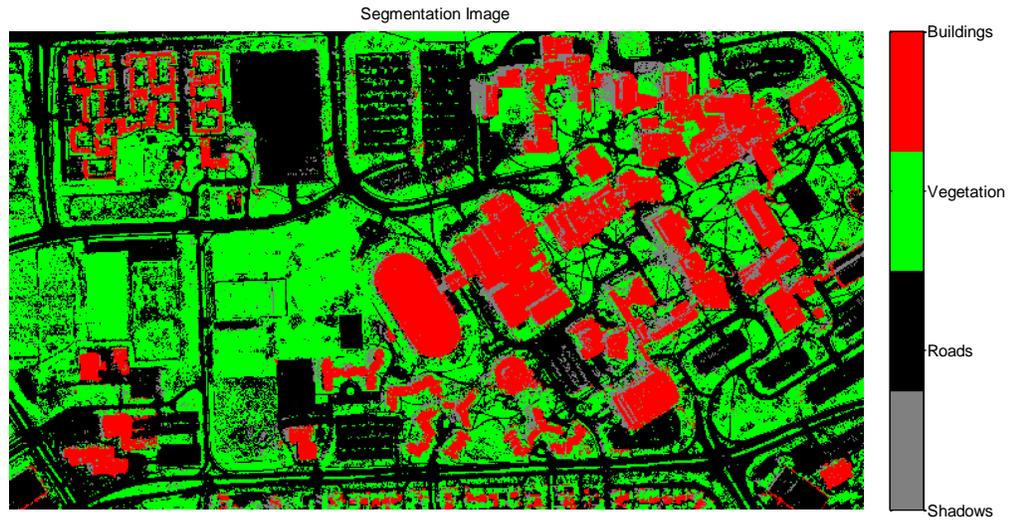


Figure A-20: Classification Results of Image – im9 and Corresponding LIDAR Data using Combined Classification Technique based on RGB/LIDAR Data

Appendix B: The Workflow of Complete Software Package for Urban Area

Classification and Object Recognition

The proposed complete software package for urban area classification and object recognition is based two main algorithms. The first algorithm for urban area classification and the second one for object recognition using extracted objects.

B-1 Urban Area Classification Workflow

Figure B-1 shows the schematic diagram of the software package workflow for urban area classification.

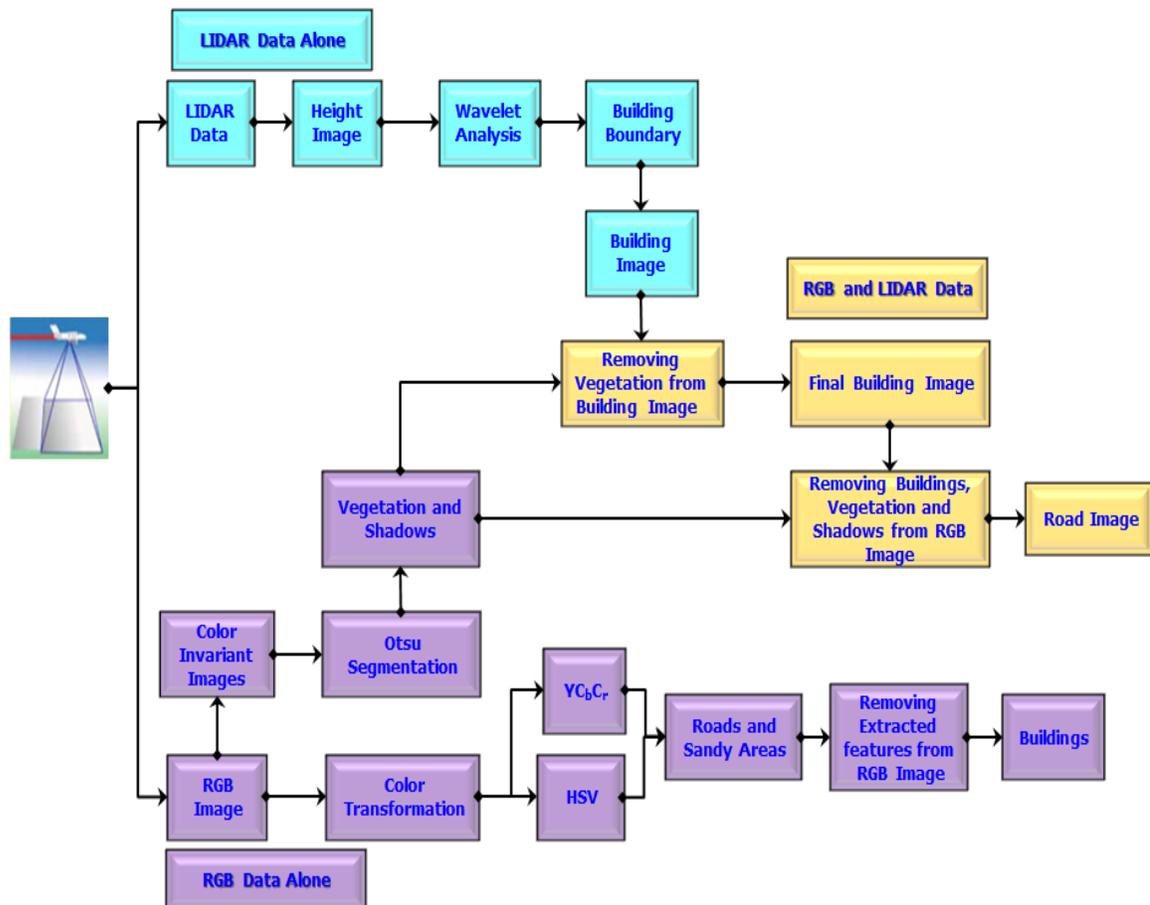


Figure B-1: Urban Area Classification Workflow through the Proposed Software Package

The software starts with the input data. Based on the input data, the software takes its workflow as follows:

- If the input data is LIDAR data alone the software concentrates on building extraction using wavelet analysis and it will end up with an image of the extracted buildings with LIDAR data representations, height and/or intensity.
- If the input data is an RGB image alone the software concentrates on classifying all urban area features. This step consists of three stages. The first stage is using Otsu segmentation method that is applied on the color invariant images that are produced from RGB color channels and it will end up with two images for shadows and vegetation areas. The second stage is applying global color thresholding on luminance, hue and saturation color channels that are produced after transforming an input RGB image from *RGB* color space to YC_bC_r color spaces and it will end up with two images for roads and sandy and unhealthy vegetation areas. The third step is producing an image of the buildings by removing all successive extracted features from an input RGB image.
- If the input data are LIDAR data and RGB image the software applies a combination between the previous mentioned classification technique to end up with an efficient urban features classification overcoming the problems in cases of features sharing semantic or height information. The final building image is produced after removing vegetation areas that is extracted from an RGB image alone from the image of extracted buildings that is produced from LIDAR data alone. The final road network and all paved areas image is produced after removing shadows, vegetation and buildings from an input RGB image.

B-2 Object Recognition Workflow

Figure B-2 shows the schematic diagram of the software package workflow for object recognition using extracted objects from urban area classification results.

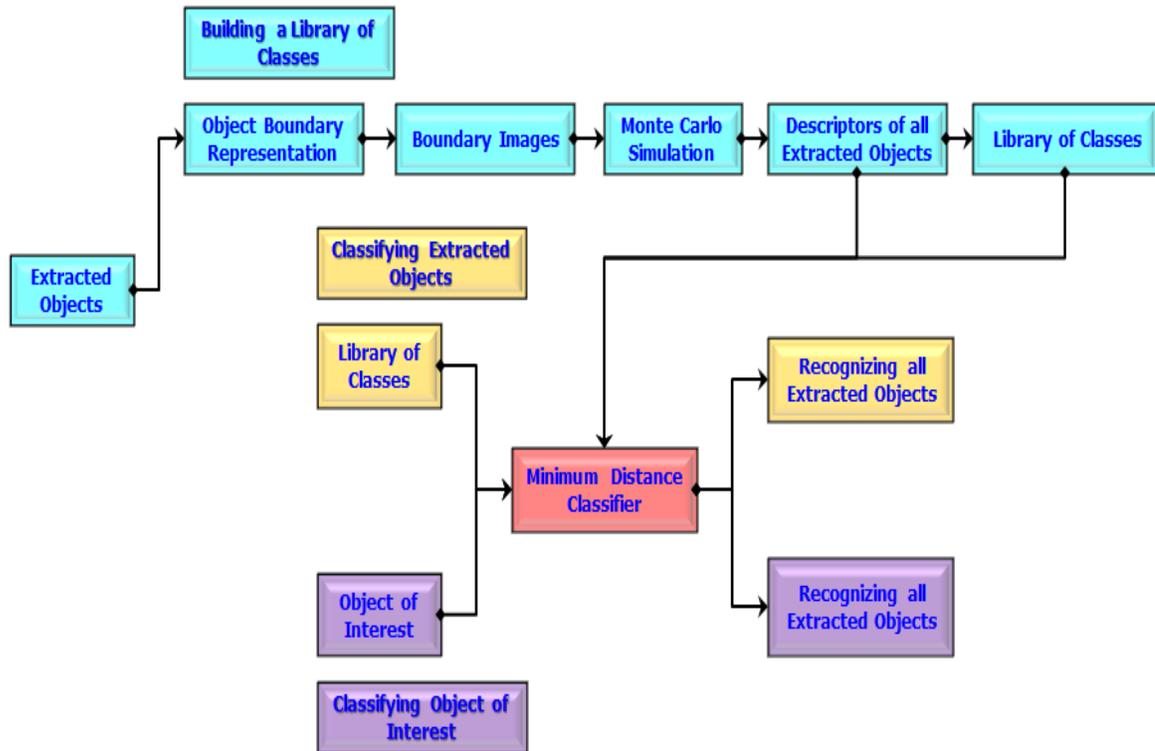


Figure B-2: Object Recognition Workflow through the Proposed Software Package

The software uses the extracted objects in three cases of applications as follows:

- If it is required to build a library of classes of the extracted features to be able to reuse for search applications the extracted features are represented by boundary images for each object. The Monte Carlo Simulation method is applied on these boundary images to calculate the object descriptors that are arranged with full available information, semantic and/or spatial, to build a library of classes.
- If a library of classes is available and it is required to recognize all or some of the extracted objects with respect to the existence library the minimum distance

classifier method uses all descriptors for both the available library of classes and the extracted objects to classify the required objects of interest.

- If there is information about a specific object of interest, image and/or spatial, and it is required to search about this object with respect to available data that is used in classification process to get more extensive information about this object the software uses the available information for the object of interest to produce the object descriptor. The minimum distance classifier uses the library of the extracted objects and the descriptor of the object of interest to recognize this object with respect to the extracted objects from the area of study. With minimum knowledge about any object, as dimension, shape and/or usage, and when the software succeeds in recognizing this object, it is possible to get full information about this object, semantic and/or spatial based on the available data; this is the target behind this software package.