Instantaneous Shoreline Extraction Utilizing Integrated Spectrum and Shadow Analysis From LiDAR Data and High-resolution Satellite Imagery

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

I-Chieh Lee, M.S., B.S.

Graduate Program in Geodetic Science

The Ohio State University

2012

Dissertation Committee:

Dr. Rongxing Li, Advisor

Dr. Alan Saalfeld

Dr. Alper Yilmaz
Abstract

Shoreline delineation and shoreline change detection are expensive processes in data source acquisition and manual shoreline delineation. These costs confine the frequency and interval of shoreline mapping periods. In this dissertation, a new shoreline delineation approach was developed targeting on lowering the data source cost and reducing human labor. To lower the cost of data sources, we used the public domain LiDAR data sets and satellite images to delineate shorelines without the requirement of data sets being acquired simultaneously, which is a new concept in this field. To reduce the labor cost, we made improvements in classifying LiDAR points and satellite images. Analyzing shadow relations with topography to improve the satellite image classification performance is also a brand-new concept. The extracted shoreline of the proposed approach could achieve an accuracy of 1.495 m RMSE, or 4.452m at the 95% confidence level. Consequently, the proposed approach could successfully lower the cost and shorten the processing time, in other words, to increase the shoreline mapping frequency with a reasonable accuracy. However, the extracted shoreline may not compete with the shoreline extracted by aerial photogrammetric procedures in the aspect of accuracy. Hence, this is a trade-off between cost and accuracy.

This approach consists of three phases, first, a shoreline extraction procedure based mainly on LiDAR point cloud data with multispectral information from satellite images. Second, an object oriented shoreline extraction procedure to delineate shoreline
solely from satellite images; in this case WorldView-2 images were used. Third, a shoreline integration procedure combining these two shorelines based on actual shoreline changes and physical terrain properties. The actual data source cost would only be from the acquisition of satellite images. On the other hand, only two processes needed human attention. First, the shoreline within harbor areas needed to be manually connected, for its length was less than 3% of the total shoreline length in our dataset. Secondly, the parameters for satellite image classification needed to be manually determined. The need for manpower was significantly less compared to the ground surveying or aerial photogrammetry.

The first phase of shoreline extraction was to utilize Normalized Difference Vegetation Index (NDVI), Mean-Shift segmentation on the coordinate (X, Y, Z), and attributes (multispectral bands from satellite images) of the LiDAR points to classify each LiDAR point into land or water surface. Boundary of the land points were then traced to create the shoreline. The second phase of shoreline extraction solely from satellite images utilized spectrum, NDVI, and shadow analysis to classify the satellite images into classes. These classes were then refined by mean-shift segmentation on the panchromatic band. By tracing the boundary of the water surface, the shoreline can be created. Since these two shorelines may represent different shoreline instances in time, evaluating the changes of shoreline was the first to be done. Then an independent scenario analysis and a procedure are performed for the shoreline of each of the three conditions: in the process of erosion, in the process of accession, and remaining the same. With these three conditions, we could analysis the actual terrain type and correct the classification errors to obtain a more accurate shoreline.
Meanwhile, methods of evaluating the quality of shorelines had also been discussed. The experiment showed that there were three indicators could best represent the quality of the shoreline. These indicators were: (1) shoreline accuracy, (2) land area difference between extracted shoreline and ground truth shoreline, and (3) bias factor from shoreline quality metrics.
Dedication

To the memories of my father, and paternal grandparents

And to my mother and my wife
Acknowledgements

Without the guidance from Dr. Rongxing Li, my advisor, this dissertation might have never been accomplished. I appreciate the huge amount of efforts he has put into this dissertation and making it better. I would like to express my gratitude to Dr. Kaichang Di, whose constructive comments have been invaluable to me in perfecting my research. I would like to thank Dr. Alan Saalfeld for his valuable comment to the dissertation and his understanding of my situation. I would also like to thank Dr. Alper Yilmaz for giving me inspiration and being supportive all the way. I would also like to thank Dr. Tatiana Oberyszyn for serving as a member in the dissertation committee.

I wish to thank the current and previous members at the OSU Mapping and GIS Laboratory: Dr. Jung-Kuan Liu, Dr. Shaoming Zhang, Dr. Qin Ye, Dr. Bo Wu, Dr. Xutong Niu, Dr. Pingbo Tang, Dr. Chun Liu, Dr. Shaojun He, Yunjae Choung, Dr. Yuan Zhou, Sagar Deshpande, Ding Li, Lin Yan, Dr. Juwon Hwangbo, and Boris Skopljak. Without their friendships it would have been impossible for me to complete this long journey.

My paternal grandfather passed away this morning, following my father and paternal grandmother who had departed for heaven in the past two years. I hope this dissertation is worth sacrificing the time I should have spent with them. I’m grateful for the support from my mother and my wife during these tough times, without them, I could have never gone this far.
Vita

2000................................B.S. Surveying Engineering, National Cheng Kung University, Taiwan
2003................................M.S. Surveying Engineering, National Cheng Kung University, Taiwan
2005 to 2006 .........................Firmware Engineer, Ability Enterprise Co., Ltd., Taiwan
2006 .................................Research Assistant, Geotechnical Lab, Department of Civil Engineering, National Taiwan University, Taiwan
2007 to 2011 .........................Graduate Research Assistant, Mapping and GIS Laboratory, Department of Civil, Environmental, and Geodetic Engineering, The Ohio State University

Publications


**Fields of Study**

**Major Field:** Geodetic Science  
**Studies in:** LiDAR  
Remote Sensing  
Photogrammetry  
GIS
Table of Contents

ABSTRACT ........................................................................................................................... II

ACKNOWLEDGEMENTS ........................................................................................................ VI

VITA ...................................................................................................................................... VII

TABLE OF CONTENTS ......................................................................................................... IX

LIST OF FIGURES ................................................................................................................ XII

LIST OF TABLES .................................................................................................................. XV

CHAPTER 1. INTRODUCTION ............................................................................................... 1

1.1 BACKGROUND ON SHORELINE DELINEATION .............................................................. 1

1.2 SHORELINE DEFINITION ............................................................................................... 5

   1.2.1 Shoreline and Bluffline ........................................................................................... 5

   1.2.2 Instantaneous and Tide-Coordinated Shorelines ..................................................... 7

   1.2.3 Being Free of Tidal Effects in Great Lakes Region .................................................. 8

1.3 REVIEW AND ISSUES .................................................................................................... 9

1.4 SOLUTIONS AND SIGNIFICANCE ............................................................................. 20

1.5 DISSERTATION OVERVIEW .......................................................................................... 24

CHAPTER 2. DESCRIPTIONS OF THE DATA SOURCES AND RELATED ALGORITHMS ........... 27

2.1 DATA SOURCES ............................................................................................................ 27

   2.1.1 LiDAR Systems ....................................................................................................... 28

   2.1.2 World View 2 Satellite Imagery ............................................................................. 35

2.2 FUNDAMENTAL ALGORITHMS USED IN THIS RESEARCH ....................................... 36

   2.2.1 The Mean-Shift Segmentation Algorithm............................................................... 37

   2.2.2 Modified Convex Hull ............................................................................................ 41

   2.2.3 NDVI ....................................................................................................................... 45

2.3 TYPES OF COASTAL TERRAINS AND DETERMINATION OF GROUND TRUTH ............. 46
CHAPTER 3. SHORELINE EXTRACTION FROM LiDAR DATA AND SATELLITE IMAGERY ......................... 52
  3.1 METHODOLOGY OF THE PROPOSED SHORELINE EXTRACTION APPROACH ................................. 52
  3.2 PHASE ONE: SHORELINE EXTRACTION BASED ON LiDAR DATA WITH MULTISPECTRAL INFORMATION ....... 54
    3.2.1 LiDAR Point Cloud Preprocessing ................................................................................. 57
    3.2.2 LiDAR Point Cloud and Satellite Imagery Registration ...................................................... 59
    3.2.3 LiDAR Point Cloud Classification ................................................................................. 61
    3.2.4 Shoreline Extraction from Classified LiDAR Point Cloud ................................................... 70
  3.3 PHASE TWO: SHORELINE EXTRACTION BASED ON SATELLITE IMAGERY ................................ 71
    3.3.1 Shadow Area Determination ........................................................................................... 78
    3.3.2 Water Surface and Vegetation Area Determination ............................................................ 85
    3.3.3 Sediment Bank/Manmade Structure Area Determination by Spectrum Matching .............. 89
    3.3.4 Classification Assignment on panchromatic image ............................................................ 95
    3.3.5 Classification Adjustments and Refinements before Shoreline Extraction .......................... 103
  3.4 SHORELINE COMPARISON METHOD ....................................................................................... 113
    3.4.1 Indicator of Shoreline Accuracy ...................................................................................... 113
    3.4.2 Indicator of Area Difference on Land .............................................................................. 115
    3.4.3 Shoreline Quality Metrics .............................................................................................. 117

CHAPTER 4. SHORELINE INTEGRATION UTILIZING SCENARIO ANALYSIS ........................................ 121
  4.1 SHORELINE INTEGRATION FOR MANMADE STRUCTURE AREAS ............................................. 123
  4.2 SHORELINE INTEGRATION FOR AREAS WITH EROSION ....................................................... 128
  4.3 SHORELINE INTEGRATION FOR AREAS WITH ACCESSION .................................................... 129

CHAPTER 5. EXPERIMENT AND ACCURACY ANALYSES OF THE EXTRACTED SHORELINES .......... 133
  5.1 DESCRIPTIONS OF STUDY AREA, DATA SOURCES, AND DATA SETS ..................................... 133
  5.2 STRATEGY AND PROCEDURE TO PERFORM SHORELINE EXTRACTION .................................. 142
  5.3 ACCURACY ESTIMATION OF THE EXTRACTED SHORELINE IN PHASE ONE ......................... 145
  5.4 PARAMETER DETERMINATION AND ACCURACY ESTIMATION OF THE SHORELINE EXTRACTED IN PHASE TWO ................................................................. 157
    5.4.1 Parameter Determination .................................................................................................. 157
    5.4.2 Accuracy Estimation ........................................................................................................ 169
  5.5 ACCURACY ESTIMATION OF THE SHORELINE INTEGRATED IN PHASE THREE .................. 172
  5.6 EVALUATION OF THE SHORELINE QUALITY METRICS ......................................................... 187
  5.7 SUMMARY AND DISCUSSION .................................................................................................. 192

CHAPTER 6. CONCLUSION AND FUTURE RESEARCH ................................................................. 196
6.1 CONCLUSIONS ........................................................................................................... 196
6.2 FUTURE WORKS ......................................................................................................... 197
REFERENCES .................................................................................................................... 201
List of Figures

Figure 1.1 The coastal erosion situation in Painesville, OH. ................................................................. 5
Figure 1.2 Definitions of a shoreline and a bluffline.............................................................................. 6
Figure 1.3 Jurisdiction boundaries are depending on tide-coordinated shorelines.................................. 8
Figure 1.4 These legal shoreline candidates are all linear features along the shore.............................. 12
Figure 1.5 The four stages of a breaking wave ...................................................................................... 12
Figure 1.6 The water run-up process forming the water line.................................................................... 13
Figure 1.7 The linear features along the coastal region .......................................................................... 13
Figure 1.8 It is difficult to determine the boundary between the water surface and the land................... 14
Figure 1.9 Spectral reflectance of the common materials within the coastal region............................... 14
Figure 1.10 Extracted shoreline comparsion in 16m and 2m resolution NIR band satellite image........... 17
Figure 1.11 Strip discrepancies among four LiDAR strips ...................................................................... 19
Figure 2.1 A terrestrial LiDAR system manufactured by RIEGL.............................................................. 30
Figure 2.2 The concept of an air-borne LiDAR system.......................................................................... 32
Figure 2.3 Spectral bands of different satellite images provide by DigitalGlobe................................. 36
Figure 2.4 One-dimensional visualization of calculating a mean shift.................................................. 39
Figure 2.5 The convex hull algorithm .................................................................................................. 42
Figure 2.6 The procedure of modifying the convex hull algorithm ...................................................... 44
Figure 2.7 The types of shoreline that are available in study area.......................................................... 48
Figure 3.1 The conceptual flow chart of the proposed shoreline extraction approach............................ 53
Figure 3.2 The flow chart of the shoreline extraction procedure in Phase One....................................... 56
Figure 3.3 Spectra reflectance measured from spectrographic aerial mapping system with on-board
navigation system .......................................................................................................................... 63
Figure 3.4 LiDAR intensity map near Sandusky, Ohio .......................................................................... 67
Figure 3.5 The four intensity profiles corresponding to the four profiles in Figure 3.4......................... 68
Figure 3.6 Elevation profile of the flight path ....................................................................................... 69
Figure 3.7 Elevation profile of Profile 1 in Figure 3.4 ........................................................................... 69
Figure 3.8 The distribution of the LiDAR points on water surface ....................................................... 70
FIGURE 3.9 FLOW CHART OF PHASE 2: SHORELINE EXTRACTION BASED ON SATELLITE IMAGERY ............................................. 72
FIGURE 3.10 CLASSIFICATION EXAMPLE USING SUPERVISED CLASSIFICATION ALGORITHM ............................................. 76
FIGURE 3.11 SHADOWS CREATED BY TREES ......................................................................................................................... 80
FIGURE 3.12 SHADOW CREATED BY VERTICAL STRUCTURE ................................................................................................. 80
FIGURE 3.13 SHADOWS OF A BLUFF TOP ............................................................................................................................. 81
FIGURE 3.14 MULTISPECTRAL IMAGE PIXELS ARE CLASSIFIED AS SHADOW AREAS ...................................................... 83
FIGURE 3.15 MULTISPECTRAL IMAGE PIXELS ARE CLASSIFIED AS WATER SURFACES AND VEGETATION ...................... 87
FIGURE 3.16 SPECTRAL PLOT OF THE FOUR MAJOR OBJECTS ON THE SURFACE ......................................................... 90
FIGURE 3.17 SPECTRAL PLOT OF THE MANUALLY SELECTED SAND PIXELS .......................................................... 91
FIGURE 3.18 AFTER THE NORMALIZATION OF FIGURE 3.17 ............................................................................................... 92
FIGURE 3.19 MULTISPECTRAL IMAGE PIXELS ARE CLASSIFIED AS SEDIMENT BANKS OR MANMADE OBJECTS .......... 93
FIGURE 3.20 FLOW CHART OF THE PANCHROMATIC IMAGERY CLASSIFICATION .............................................................. 98
FIGURE 3.21 THE CONCEPT OF STANDARD DEVIATION IMAGE STRETCHING ................................................................. 100
FIGURE 3.22 THE STACKING SEQUENCE GENERATING THE FINAL CLASSIFICATION FROM INDIVIDUAL CLASSIFICATIONS .... 102
FIGURE 3.23 IN A SANDY BEACH AREA, WATER SURFACE AND SEDIMENT BANK CLASSES ARE RIGHT NEXT TO EACH OTHER .... .. 104
FIGURE 3.24 FLOW CHART FOR THE CLASSIFICATION ADJUSTMENT PROCESS .............................................................. 105
FIGURE 3.25 THE NORTHERN SIDE OF UNCLASSIFIED AREA IS SHADOW CLASS ................................................................... 106
FIGURE 3.26 SHADOWS CREATED BY MANMADE OBJECTS SUCH AS VERTICAL STRUCTURES ........................................ 108
FIGURE 3.27 SHADOWS CREATED BY TREES AND BLUFF TOPS .......................................................................................... 109
FIGURE 3.28 POSSIBLE CONDITIONS FOR A SHADOW-SHORELINE RELATION ............................................................ 109
FIGURE 3.29 AN EXAMPLE OF A TREE CLUSTER .................................................................................................................. 110
FIGURE 3.30 A SHADOW AREA SPLIT INTO THE WATER SURFACE AND THE LAND CLASS BASED BY THE SURROUNDING CLASSES. 110
FIGURE 3.31 AN UNCLASSIFIED AREA IN-BETWEEN WATER SURFACE CLASS AND SEDIMENT BANK ..................... 112
FIGURE 3.32 UNCLASSIFIED AREA IN-BETWEEN SEDIMENT BANKS, VEGETATIONS, AND SHADOW AREAS ...................... 112
FIGURE 3.33 THE TRADITIONAL METHOD OF CALCULATING DISTANCES BETWEEN SHORELINES ........................................ 113
FIGURE 3.34 THE METHOD USED WHEN COMPARING EXTRACTED SHORELINE WITH THE GROUND TRUTH SHORELINE ....... 115
FIGURE 3.35 AN EXTRACTED SHORELINE COMPARED WITH THE GROUND TRUTH SHORELINE ........................................ 117
FIGURE 3.36 THE CONCEPT OF GENERALIZATION FACTOR .............................................................................................. 118
FIGURE 3.37 THE CONCEPT OF DISTORTION FACTOR (DF) .............................................................................................. 119
FIGURE 3.38 THE CONCEPT OF THE BIAS FACTOR (BF) .......................................................................................................... 119
FIGURE 3.39 THE CONCEPT OF THE FUZZINESS FACTOR (FF) ............................................................................................ 120
FIGURE 4.1 FLOW CHART OF THE SHORELINE INTEGRATION PROCEDURE ............................................................... 123
FIGURE 4.2 FLOW CHART FOR PERFORMING THE SHORELINE INTEGRATION WITHIN SLOPED STRUCTURE AREAS ........ 126
FIGURE 4.3 FLOW CHART OF SLOPED STRUCTURE AREA DETECTION ........................................................................... 128
List of Tables

TABLE 3.1 THE OBJECT PROPERTIES ON DIFFERENT DATA SOURCES. .................................................................62
TABLE 5.1 DATA SOURCES OF THE PROPOSED APPROACH ................................................................................142
TABLE 5.2 SHORELINE ACCURACY BEFORE MANUAL INSPECTION .................................................................147
TABLE 5.3 SHORELINE ACCURACY AFTER EXCLUDING SHORELINE CHANGES BETWEEN DATASETS ..............156
TABLE 5.4 THE RMSE OF SHORELINE ACCURACY WITH RESPECT TO THE CB AND SOP PARAMETERS ..........161
TABLE 5.5 SHORELINE ACCURACY WITH RESPECT TO THE SWS, CB, AND SOP PARAMETERS .........................165
TABLE 5.6 SHORELINE ACCURACY WITH RESPECT TO THE SOP (0.1~0.6), CB, AND SB PARAMETERS .................166
TABLE 5.7 THE PARAMETER SET CHOSEN ACCORDING TO THE PARAMETER DETERMINATION GUIDELINES ......171
TABLE 5.8 THE QUALITY ESTIMATES OF THE EXTRACTED SHORELINE ..........................................................171
TABLE 5.9 THE ACCURACIES OF THE EXTRACTED SHORELINES WITH INDIVIDUAL TERRAIN TYPES ...............171
TABLE 5.10 THE QUALITY ESTIMATES OF THE INTEGRATED SHORELINE .........................................................174
TABLE 5.11 THE ACCURACIES OF INTEGRATED SHORELINE WITH INDIVIDUAL TERRAIN TYPES .....................175
TABLE 5.12 SHORELINE QUALITY INDICATORS WITH RESPECT TO THE SHORELINES EXTRACTED IN PHASE ONE, PHASE TWO AND PHASE THREE .........................................................................................191
TABLE 5.13 THE ACCURACIES OF INTEGRATED SHORELINE WITH INDIVIDUAL TERRAIN TYPES INCLUDING ACCURACY AT 95% CONFIDENCE LEVEL .........................................................................................193
Chapter 1. Introduction

1.1 Background on Shoreline Delineation

Along the shore of Lake Erie, severe erosion has impacted all shorelines significantly (Ali, 2003). Bluffline erosion in the region of Painesville, Ohio can reach as serious as six feet (1.8m) per year (Srivastava, 2005). The Ohio Department of Natural Resources (ODNR) has been monitoring coastal erosion since 1988. The Coastal Erosion Area (CEA) designated by the ODNR in 1988 and 2010 confirmed that the erosion process has been present for a few decades. Figure 1.1 shows the evidence of the constant eroding process in Painesville. According to the ODNR CEA maps, the CEA is designated based on the amount of bluffline changes during 1990 and 2004 showing on the aerial photographs, since these years are the years the aerial photographs were acquired. However, two sets of delineated bluffline at an interval of 25 years might be enough for designating the CEA, they are not sufficient to determine the cause of shoreline erosion or to plot solutions to prevent the shoreline from being further eroded. Two major reasons to prevent higher frequency of shoreline mapping are the higher expense and the longer duration of the aerial photogrammetry project.

“A number of possible shoreline proxies have been suggested, including the beach scarp, high water line (HWL), berm crest, vegetation line, dune toe, dune crest, and the bluff edge (Leatherman, 2003)”. Thus, how to define and where to locate the shoreline
are based on applications and audiences. Since CEA is indicating where the real estate on
the bluff top might be lost due to the bluff erosion, ODNR uses the bluffline changes to
indicate the CEA. However, “most coastal researchers and government agencies in the
United States use the High Water Line (HWL) because this shoreline indicator is visible
in the field and can be interpreted on aerial photographs by the gray scale or color tone
(Leatherman, 2003)”. The United States Federal Geographic Data Committee (FGDC)
recognizes the definition of shoreline by the National Oceanic and Atmospheric
Administration (NOAA) (FGDC, 2010) as “the intersection of the land with the water
surface. The shoreline shown on charts represents the line of contact between the land
and a selected water elevation. In areas affected by tidal fluctuations, this line of contact
is the mean high water line. In confined coastal waters of diminished tidal influence, the
mean water level line may be used (Hicks, 2000)”.

“The intersection of the oceans and land is a margin of continuous change. Shoreline positions vary in response to several factors, and the movement of the shoreline can vary dramatically or negligibly in no uniform pattern at temporal scales of daily, seasonally, and decadal” (White, 2007). Both water level and shape of terrain determine the location of a shoreline. The reason that the location of a shoreline is constantly changing is the changes of water level and/or shapes of terrains. For example, in the Great Lakes region, the water level is related to the inflow/outflow, precipitation/evaporation, and shore-term fluctuations (Velissariou, 2009). Shapes of terrains are related to erosion, sedimentation process, and human activities, such as harbor construction and/or dredging. As for tidal coastlines, water level is mainly affected by tidal fluctuations. Water level is constantly changing, and the delineated shoreline
needs to correspond to a certain water level. In tidal areas, tide-coordinated water level is used to define the legal shoreline (Hicks, 2000). For example, Mean High Water (MHW) that we have mentioned in the previous paragraph is one of the tide-coordinated water levels. A shoreline which is delineated based on a tide-coordinated water level is called a tide-coordinated shoreline (Li et al., 2002a). In this case, a shoreline delineated while water level is at the MHW is called a MHW tide-coordinated shoreline. On the contrary, a shoreline that is delineated not at a certain tide-coordinated water level but at any instant in time is called an instantaneous shoreline (Boak and Turner, 2005). The detailed definition of shorelines is included in section 1.2.

Shorelines in the U.S. are officially managed by the National Geodetic Survey (NGS), NOAA. In the NGS, shoreline is delineated mainly from the stereoscopic environment of aerial photogrammetric procedures (White, 2007). Recently, the NGS has been developing procedures to extract shorelines from a water penetrating LiDAR system to delineate shorelines as well (White et al., 2010). Although the aerial photogrammetric shoreline mapping procedure could provide the most accurate shoreline among other semi-automatic or autonomous approaches, it is a time-consuming process and the placement of shoreline is subject to human interpretation (White, 2007). As a result, it is a cost intensive procedure involving not only the amount of manpower but also having experienced operators. Consequently, the high cost limits the frequency of periodic shoreline mapping. On the other hand, the shoreline changes could be caused by a long-term process or a single event. For example, a huge amount of precipitation within a short time period may cause land slide and result in shoreline changes; earthquakes may lead to tsunamis and reform a coastal area overnight. Shoreline delineation procedures
such as aerial photogrammetry and aerial LiDAR need to be planned in advance, and it is challenging to acquire these datasets during an event while shoreline changes rapidly.

In this dissertation, we proposed a new approach to delineate the instantaneous shoreline from the satellite imagery. With satellite imagery we could not only monitors long-term shoreline changes, but also capture shoreline changes during an event if the satellite imagery is obtainable. Its cost is also relatively low compared with aerial photogrammetry projects. However, the highest resolution of commercial satellite imagery currently available is around 0.5m. It is about several times lower than the resolution of an aerial photograph. Instead of delineating the tide-coordinated shoreline, we targeted on the instantaneous shoreline. Delineating the tide-coordinated shoreline is still our long-term goal; however, there is no direct method of delineating tide-coordinated shoreline from satellite imagery. Hence we simplified the problem as delineating the instantaneous shoreline from satellite imagery first, and figuring out how to acquire satellite imagery with a certain tide-coordinated water level later. We think that dividing this process into two stages is reasonable because the concept of shoreline delineation process from aerial photograph is similar as well.

Painesville Township, Ohio is the chosen study area for this research. According to the ODNR CEA map updated in 2010, 51% of the shoreline within our study area was designated as CEA (information derived from ODNR, 2010). The maximum recess rate within the study area was 10.5 ft/year (3.2 m/year) between 1990 and 2004 (ODNR, 2010). Considering the high erosion rate in this region, finding a cost-effective way to map shorelines is an important issue to increase the mapping interval. Moreover, because of the related research in this region over the past years, we have accumulated extensive
data sources, data sets, and knowledge regarding this region.

Figure 1.1 The coastal erosion situation in Painesville, OH. Image (a) shows the building from the bird’s eye view (Image Credit: Bing Map), (b) shows the same building from the side view. There is a pipe stretching out from the bluff demonstrate the amount of land has been eroded after the pipe was installed.

1.2 Shoreline Definition

In this research, we target on delineating the instantaneous shoreline from the satellite imagery. In this section, we describe the differences between shoreline and bluffline, and instantaneous shorelines and tide-coordinated shorelines, and the reason that we use a non-tide coast area as our study site.

1.2.1 Shoreline and Bluffline

A shoreline is roughly defined by the U.S. Geological Survey (USGS) as a line of contact, naturally occurred, between a land and a body of water (Graham, Sault, and
Bailey, 2003). The FGDC defined the US legal shoreline as the mean high water line (MHW). A bluffline, on the other hand, is defined by the US Army Corps of Engineers (USACE) is “a cliff or headland with an almost perpendicular face (USACE, 2002; International Hydrographic Bureau, 1990)”. There are two distinct features of a bluffline, bluff top and bluff toe. Liu et al. (2009) and Choung (2009) from the Mapping and GIS Laboratory, Ohio State University have been working on bluffline extraction from LiDAR point cloud and have obtained promising results. The focus of this dissertation is on shoreline delineation. By combining shoreline and bluffline delineation altogether, the topology of this coastal region can be shown thoroughly. Figure 1.2 depicts the definition of a shoreline and a bluffline.

Figure 1.2 Definitions of a shoreline and a bluffline (Liu et al., 2009; California Coastal Commission, 2004)
1.2.2 Instantaneous and Tide-Coordinated Shorelines

"The instantaneous shoreline is the position of the land-water interface at one instant in time (Boak and Turner, 2005)”. That instant is when data is acquired. For example, shorelines showed in satellite images, aerial photographs, or ground surveys are instantaneous shorelines (Sukcharoenpong, 2010).

Tide-coordinated shorelines are shorelines while tide is at a specific water level (Li et al., 2002a; Lipakis, 2008). This water level is based on a tidal datum and aggregated from water level observations over a period of time. For example, if an instantaneous shoreline is extracted while the water level is at mean lower-low water (MLLW) level, this shoreline can be used as a MLLW tide-coordinated shoreline. Tide-Coordinated shorelines have been used in various fields, for example, jurisdiction boundaries such as national borders, state borders, etc (Figure 1.3). In nautical charts tide-coordinated shorelines are also used to provide information for safe passage through waters. Li et al. (2002a) categorized tide-coordinated shorelines into two types: physical tide-coordinated shorelines and digital tide-coordinated shorelines. This categorization is based on the methodology used to delineate shorelines. Physical tide-coordinated shorelines are obtained by extracting instantaneous shorelines from aerial photographs (White, 2007; Woolard et al., 2003) acquired at certain tide-coordinated water levels. Digital tide-coordinated shorelines, on the other hand, are obtained from a Coastal Terrain Model (CTM) by combining a Digital Elevation Model (DEM) with bathymetry, then intersecting this CTM with a certain tide-coordinate water surface (Li, 1997; Li et al., 2002a; Robertson et al., 2004; Stockdon et al., 2002). Since deriving tide-coordinated shorelines from instantaneous shorelines is one of the major ways to extract
tide-coordinated shorelines, we decided to focus on instantaneous shoreline delineation in this research.

![Jurisdiction boundaries diagram](image)

Figure 1.3 Jurisdiction boundaries are depending on tide-coordinated shorelines (NGS, 2012)

1.2.3 Being Free of Tidal Effects in Great Lakes Region

Our test site is located in the Great Lakes region in Painesville, Ohio. Since this research does not target at tide-coordinated shorelines, having tide or not is not an issue when choosing experiment locations. Nevertheless, choosing a non-tide region may also simplify the shoreline delineation problem during the methodology development phase.
The problem may become a purely processing and classification problem, so that the physical constant changing water level can be put aside. Although there are no tidal effects in Lake Erie, there are still water level fluctuations, including seasonal water level fluctuations caused by water inflow/outflow and precipitation/evaporation, and shore-term fluctuations such as storm surges and seiches caused by wind (Velissariou, 2009). Nevertheless, tide-coordinated water level can still be calculated within the Great Lake region through the traditional way of calculating tide-coordinated water level in regions with tidal effect. Consequently, Great Lake tide-coordinated shorelines may still be used by the researchers, but that does not mean there are tides in the Great Lakes.

1.3 Review and Issues

Shoreline mapping in the United States dates back to 1807. Since then, ground surveys with plane table and theodolite have been the primary way of mapping shorelines (Shalowitz, 1964). Needless to say, this is a labor intensive and high cost method. Moreover, shoreline accuracy highly depends on experiences of technicians. In the last century, aerial photogrammetry was introduced to the field, and became the main method of shoreline mapping ever since. This is also the major method the NOAA currently utilizes in mapping the shorelines in the U.S., and is recognized by the federal government (Li et al., 2002b; Woolard et al., 2003). Shorelines delineated from aerial photogrammetry are the most accurate compared with all other surveying techniques, due to its high image resolution and high accuracy of horizontal control (Li et al., 2001). The reason why researchers keep developing new methods to map shorelines is the high cost of the aerial photogrammetry project. The high cost comes from two of the phases of the
project: the aerial photo acquisition phase and the manual processing phase. Before acquiring aerial photographs, ground control survey needs to be done, which is a labor intensive task. Moreover, acquiring aerial photos highly depends on weather conditions. Weather conditions have to meet the safety requirements of the aviation regulation. On top of that, in order to delineate tide-coordinated shorelines, aerial photo acquisition time has to be within the window of desired water levels. For example, to delineate a MLLW tide-coordinated shoreline, aerial photographs need to be taken when the water level is at the MLLW level. Needless to say, observation of water level changes for a period of time is definitely needed in order to predict a specific water level and establish a schedule for the flight plan. If the weather condition during the planned period is not satisfying the requirements of picture quality and flight safety, the acquisition must be rescheduled and all the man power and equipment costs would be wasted. The other reason that keeps the cost high is the manual processing of the aerial photogrammetry procedures: aerial triangulation and manual shoreline delineation. Aerial triangulation is a sophisticated and delicate procedure. Although there is computer software to assist engineers in solving this problem, they still have to be well trained and experienced to get the best accuracy. On the other hand, after aerial triangulation is done, shoreline could then be delineated from either a stereo image or an orthophoto. Shoreline delineation from stereoscopic environment is a time-consuming process and the placement of shoreline is subject to human interpretation (White, 2007). Shoreline delineation from orthophoto was a popular topic in this field. Researchers (Di et al., 2003; Woolard et al., 2003) had developed several algorithms to delineate shoreline semi-automatically. There is still no robust and recognizable algorithm to delineate shorelines from an aerial orthophoto autonomously.
There are several reasons that make the autonomous shoreline delineation process from orthophoto almost impossible. First, multiple linear features exist along a shoreline. Figure 1.4 depicts the linear features (vegetation line, high water line, berm crest, base of scrap (bluff), and manmade structures) along the shore on the land. These features have been mentioned in section 1.2.1 that they are also legal shoreline candidates. Figure 1.5 depicts the different stages of a breaking wave on the water surface and Figure 1.6 shows the physical process of water run-up. All these features can be seen from the aerial images and are linear and close to the actual shoreline (Figure 1.7). It is extremely difficult to use a computer algorithm to determine which line is the desired shoreline. Secondly, in some regions, it is difficult to separate water and lands using only red, green, and blue color images (Figure 1.8). In order to achieve full autonomy, some external information is definitely needed, for example, elevation information or Near Infrared (NIR) images. Figure 1.9 depicts the spectral reflectance of the common materials within a coastal region, and the R, G, B, and NIR spectral bands a remote sensing camera received. As seen from the image, NIR band is useful in distinguish vegetations and water surfaces. However, the NIR image is not a necessity data for the aerial photogrammetry process. Hence the acquisition of NIR band has to be added to the project requirements if needed.
Figure 1.4 These legal shoreline candidates are all linear features along the shore (Morton and Speed, 1998)

Figure 1.5 The four stages of a breaking wave (Sorensen, 2005)
Figure 1.6 The water run-up process forming the water line (Sorensen, 2005)

Figure 1.7 The linear features along the coastal region. The high water line is the shoreline defined by the NOAA. (Image Credit: Bing Map)
Figure 1.8 It is difficult to determine the boundary between the water surface and the land because of the moisture content of the sand is similar, and the water is clear, shallow and calm. (Image Credit: Google Map)

Figure 1.9 Spectral reflectance of the common materials within the coastal region. The reflectance data is provided by the USGS Digital Spectral Library 06.
NIR images are often simultaneously acquired with aerial photographs in recent years, but of course with an extra cost. There are three types of digital NIR cameras available: the medium format digital aerial photograph frame camera with filter for NIR imaging such as Applanix DSS (Applanix, 2008), Rollei AIC (Optech, 2012), and DiMAC ULTRALiGHT+ (DiMAC, 2012), the multiple frame cameras housed in one single unit such as Intergraph DMC (Intergraph, 2012), Microsoft UltraCam (Microsoft, 2012), and DiMAC WiDE+ (DiMAC, 2012), and line scanning sensors such as Leica ADS (Leica, 2012). For acquiring Color-Infrared (CIR) imagery, only R, G, and NIR bands are used, a frame camera with a CIR filter could satisfy the requirement (Applanix, 2008; Optech, 2012). In case of using the frame camera to acquire R, G, B, and NIR images, two aerial cameras are needed (Applanix, 2008; DiMAC, 2012). However, more sensor devices onboard the aircraft usually means a larger aircraft is needed and higher expenses. On the other hand, multiple frame cameras or multiple linear array cameras can acquire NIR simultaneously with red, green, and blue bands (Intergraph, 2012; Microsoft, 2012; Leica, 2012). However, they are not traditional frame cameras which the photogrammetry is designed upon; image processing procedures may be different. The choice of cameras is to weight pros and cons between the application, requirements, and costs.

Satellite images have been a popular data source of delineating shorelines in the recent decades. In the past, multispectral satellite image resolution is over 10 meters (Landsat 1~7 and SPOT 1~3). Shoreline in an image of this level of resolution would be like a clear-cut line especially in NIR band (Figure 1.10 a). This kind of sharp-edged linear feature can be easily extracted by binary segmentation used in image processing. In
the past decade, resolution of satellite images reached a sub-meter range and spectrum resolution was improved as well. One of the latest multi-spectral remote sensing satellites operated by DigitalGlobe was named WorldView 2 (WV2). It provides 8 bands of multi-spectral images in two-meter resolution and one panchromatic band in half-meter resolution. In a sub-meter satellite image, the shoreline is not a clear-cut line anymore (Figure 1.10 b). The coastal region in the sub-meter satellite image is more similar to that in an aerial image. As a result, the shoreline is not the only linear feature in the coastal region. Besides, you can also identify the ripple on the water surface, the wave fronts near the shore, the submerged sediments, and the wet sand caused by run-up waves. These objects create the gray intensity areas in the NIR band between the true water surface and the true land region and may not be classified correctly. Having no distinct separation, the shoreline extracted from this NIR band may not help to increase the robustness along with the increasing resolution of the satellite imagery.

For aerial photographs and satellite images, existence of shadows is another inevitable issue. Shadows may exist at any places with an elevation difference, and directions of shadows vary from time to time. This is not only troublesome for computers to delineate a shoreline, but also confusing for determining shoreline locations even for humans. When shorelines are within a shadowed area, it is either unable to determine locations or determined location is unreliable. It would be a great improvement if the errors created by shadows can be partially eliminated.
Figure 1.10 In a 16m resolution NIR band satellite image (a), the shoreline is a clear-cut line. In the image of 2m resolution (b), it is not that easy to identify the shoreline. Both images are the NIR band from WV2, and image (a) is resampled to 16m and equalization is applied to simulate the images from older satellite.

Light Detection And Ranging (LiDAR) systems are one of the newer technologies introduced for shoreline delineation. Aerial LiDAR point cloud contains a vast wealth of data with horizontal coordinates, elevation, and intensity for each data point. This was a revolutionary concept for the field of surveying. There are two major types of aerial LiDAR systems: Topographic LiDAR system and Bathymetric LiDAR system. A Topographic LiDAR system provides point clouds on lands and some of points on water surfaces. The laser pulse of a bathymetric LiDAR system can penetrate water. In an ideal condition, the LiDAR system can measure points in depths up to 50 m (Vosselman and
Maas, 2010). It can provide point clouds on lands, under water, and on water surfaces. In the majority of methods used to delineate shorelines from aerial LiDAR is to first generate a DEM from either topographic or bathymetric LiDAR, then intersect DEM with a water level elevation from tidal datum (Stockdon et al., 2002, Robertson et al., 2004, White et al., 2007, 2010.), e.g. MLLW. In addition, a topographic LiDAR point cloud has to be acquired during low tide in order to cover entire tidal zone and create any possible tide-coordinated shoreline. In this kind of shoreline delineating method, several key issues need to be taken care of precisely. The first one is LiDAR systematic errors. LiDAR strip discrepancies are the end result of the systematic errors (Figure 1.11). Although through the strip adjustment process most of these discrepancies may be removed, a small amount of them would still remain in the LiDAR points (Kilian et al., 1996; Crombaghs et al., 2000; Vosselman and Maas, 2001; Schenk, 2001; Maas, 2002). Shoreline extracted using the DEMs created by the LiDAR point cloud with these discrepancies may lead to irregular shapes of shorelines, especially in the overlapping areas of two LiDAR strips. Secondly, tidal observation is a necessary data source. The closer it is to a tidal station, the higher accuracy of tidal water level would be. For regions without a tidal station, tidal water levels are interpolated, and resulting shoreline accuracies will be degraded by this inaccurateness.
Figure 1.11 Strip discrepancies among four LiDAR strips (Toth and Grejner-Brzezinska, 2009). A same building having four set of LiDAR measurements.

Data integration is a popular way of trying to solve unsolvable problems while using single data sources. For example, a LiDAR point cloud is commonly integrated with an aerial orthophoto for terrain mapping (Liu et al., 2007). The major issue for this kind of integration is time shift between data sources. Water level is a constantly changing surface, and timing is extremely important for tide-coordinated shoreline delineation. It is a difficult task to get all data sources acquired simultaneously, especially satellite images. If data sources are not acquired simultaneously, a time series model is needed to estimate changes through time. As a result, putting multiple data sources altogether may not always simplify the process, it might create more problems than it solves. Multiple data sources such as LiDAR and aerial photographs can possibly be acquired simultaneously. This combination of data sources could be synchronized with multiple instruments put on the aircraft side by side under a specially designed flight plan (Vosselman and Maas, 2010). For data sources not acquired at the same time, it is better to work on applications where changes are not as dramatic as tidal effect.
The issues described previously are the goals we try to conquer with our proposed approach. These goals are summarized as below:

a) Lowering the expenses of data sources

b) Autonomous shoreline delineating algorithm

c) The target accuracy of the extracted shoreline meeting the accuracy requirements for NGS shoreline mapping projects

d) Minimizing the impact of shadow areas in satellite images.

e) Maintaining the robustness and accuracy while LiDAR systematic errors after strip adjustments still exist

f) Using data sources that are acquired at different instance in time.

1.4 Solutions and Significance

Erosion, sedimentation, and human activities are the major forces that influence shapes of coastal regions. All these forces are interrelated with each other and can reshape topographies of coasts. Shoreline and bluffline changes are results of these active forces. In other words, the more frequently shoreline changes are monitored, the more we can understand the corresponding coastal environment. Increased frequency of shoreline mapping can be done by shortening mapping time-table and reducing mapping costs. Under these criteria, there are two guidelines while developing this shoreline delineation method: using data sources that are less labor intensive and cost-less, and increasing the degree of autonomy of the shoreline delineating algorithms.

Although with aerial photogrammetry the accuracy of shoreline delineation is the highest, it is a labor intensive task on the ground, in the air, and in the office. The best
substitute for aerial photogrammetry currently is satellite imagery, although its resolution and accuracy are lower than those of an aerial orthophoto, and there is no way to generate DEM based on a single satellite image. However, multi-spectral satellite imagery contains extra bands, especially NIR bands for better water/land separation. Needless to say, satellite imagery can be acquired more rapidly with a significantly lower cost.

Shoreline in an aerial photograph or satellite imagery can be described as a linear feature. However, there are multiple linear features along a coastal region in a high-resolution aerial or satellite imagery. It is extremely difficult to distinguish shorelines from other kinds of lines. In a region of clear and calm water, the boundary between water and land is also difficult to define. In order to achieve autonomous shoreline delineation, additional data sources are needed.

Due to the lack of the elevation information for satellite imagery, it is obvious that a LiDAR point cloud would be a good candidate to be integrated with satellite imagery. Similar to aerial photogrammetry, LiDAR also requires intensive human labor interactions. However, manpower spent in ground control and post-processing procedures are much less than those required for aerial photogrammetry. Nevertheless, LiDAR also has its drawbacks. The most significant drawback in shoreline delineation is the systematic error removal between LiDAR strips. A LiDAR systematic error consists of errors from each individual device and mounting of the device. Some of these systematic errors can be calibrated in the lab, but some of them can only be calibrated on-the-job (Burman, 2000). According to the LiDAR model depicted by Schenk (2001), there were still systematic error parameters that could not be solved. The amount of these remaining errors may vary depending on the factors of the LIDAR project. In the dataset used in this
research, there is a 0.2~0.5m error in the elevation between the strips. As a result, these LiDAR point clouds and the DEM created by the LiDAR data cannot be treated as true value. On the other hand, LiDAR points are not uniformly distributed. They may be very dense in forest and bushes, but extremely sparse on water bodies such as lakes, ponds, or oceans. The density of the LiDAR points depends on the scanning parameters from hardware of a LiDAR system, flight elevation, and terrain type. Nominal Point Spacing is a term representing point density. Moreover, there is no way to determine where a laser beam hit the ground exactly, and no way to control a laser beam to hit an exact point that you need to measure.

A classification algorithm is required in finding the separation of water surface and land in either LiDAR data or satellite images. There are two major kinds of classification algorithms: supervised ones and unsupervised ones. Supervised classification algorithms usually require a small portion of data source to be manually classified, and this portion is used as a training set to train the classification models with adjusting the parameters. An algorithm with this kind of process is usually called a semi-automatic algorithm. In order to achieve a fully autonomous algorithm, unsupervised classification is usually used. It does not mean that classification parameters do not need to be determined, for these algorithms, parameters are usually obtained based on empirical values and physical meanings of data source. In addition, given this classification, adjusting parameters is also a solution of tolerating systematic errors and random errors from data source. Preventing these errors lead to inaccurate results of delineated shorelines.

Lastly, there is the timing problem between data sources and datasets. It is almost
impossible to obtain a LiDAR dataset and a satellite imagery simultaneously. We can possibly minimize the time difference to several days, but that is about as close as we can get. The shoreline position depends on tide elevation, and several days of data source timing difference already assure that the water elevations are not going to be consistent between the data sources. Consequently, to create a shoreline delineation algorithm, the assumption that data sources are acquired simultaneously has to be removed. Hence, a model which estimates changes of tides and terrains between data sources is needed in order to make delineated shorelines robust and reliable. Since we detach the acquisition time between data sources, we can set our standard higher by extending the time difference to the degree of years. In other words, we can use a 2006 LiDAR dataset, a 2010 satellite imagery, and a 2011 satellite imagery, to delineate the 2010 and 2011 instantaneous shorelines and detect the changes. We may also use a 2006 LiDAR dataset with several 2011 satellite images taken within a short time period and delineate several instantaneous shorelines and then interpolate these shorelines into the 2011 tide-coordinated shoreline with the help from tide observation dataset. Under these circumstances, the LiDAR data, which is the most expensive data source in this shoreline delineation method, can be replaced by free public data sources acquired years ago. The only expense would be from the latest satellite imagery acquisition.

When a human sees an aerial image of a coastal region, he can instantly envision the scene on the ground in 3-dimension, and then understand what each linear feature on the aerial image represents. Thus he can select the precise line needed to delineate the shoreline. This is done with the knowledge of the terrain type, structures, and materials on the ground in order to envision the scene. In order to achieve the same accuracy of a
manually delineated shoreline, a computer program has to do the same as a man does in one’s mind. First, materials along the shoreline must be determined in order to identify the terrain type. Secondly, the type of terrain can then be used to analysis the change patterns with time. For example, we first determine that the area is covered in sand using the multispectral satellite images. However, both the manmade structure and the sandy beach (sediment bank) are made of sand. Hence elevation information is incorporated to determine the exact terrain type. Then possible physical changes can also be estimated, such as erosion or sedimentation pattern and rate for a sandy beach area. Shadows have always been treated as an obstacle in remote sensing classification (Yamazaki et al., 2009), in this research, we used them as a form of providing elevation difference information and tried to minimize the impact caused by shadow areas.

In sum, the highlights of the proposed shoreline delineating approach are that: a) the shoreline delineation is based on understanding the terrain types and physical characteristic of the terrain while in the process of erosion/accession, which is new in shoreline extraction; b) removing the requirement of the data sources having to be acquired simultaneously leads to a new opportunity of obtaining shorelines under a cost-effective principle with rapid production; and c) utilizing the shadow analysis to benefit from it while minimizing the impact is also new in this field.

1.5 Dissertation Overview

The content arrangement of the dissertation is as follows: Chapter 2 describes the background knowledge and the algorithms used in this proposed approach. This includes the principle of LiDAR system, the properties of World View 2 satellite imagery, the
concept and usage of Mean-Shift segmentation, Normalized Difference Vegetation Index (NDVI) and Modified Convex Hull algorithm, and description of the terrain types that is existed while delineating ground truth shoreline.

The proposed shoreline delineation approach consists of three phases. The first phase is to extract a shoreline based on LiDAR point cloud data. The second phase is to extract shoreline based on World View 2 satellite imagery. The third phase is the integration of these two shorelines together by compensating each other’s weakness and maintaining the best of their results.

Chapter 3 describes the method that we have developed for the first and second phase. In first phase, a classification process is adopted to separate the LiDAR points into land points and water surface points. Then the shoreline is extracted by tracing the boundary of the land points. In the second phase, the satellite imagery is used to determine the materials and terrain types. These processes include spectrum matching procedures for sediment bank and concrete structure classification, shadow area classification and processing, and NDVI indicator to classify water surface areas and vegetation areas.

Chapter 4 describes the integration of these extracted shorelines. The shorelines are integrated based on the analysis of the water level difference between the LiDAR and satellite imagery data, the erosion and sedimentation between data sets, and shadow analysis in satellite imagery.

Chapter 5 describes the data sets used in this study and the accuracy estimation of the extracted shoreline from LiDAR point cloud, satellite imagery and the integrated shoreline. The accuracy estimation uses a process developed by this research and

Chapter 6 concludes the dissertation with discussions and future research topics.
Chapter 2. Descriptions of the Data Sources and Related Algorithms

In this chapter, all of the fundamental knowledge is described, including the data sources and the related algorithms. The data sources include topographic aerial LiDAR point cloud and multispectral satellite imagery. The data sets used by the research are aerial topographic LiDAR data, and World View 2 (WV2) satellite imagery. In this proposed shoreline extraction approach, there are several algorithms being used, including well known algorithms such as Mean-Shift segmentation, Normalized Difference Vegetation Index (NDVI), and the modified convex hull algorithm. In addition, when delineating ground truth shoreline, the terrain type is also been determined as well. In the final section, terrain types used in this research is described.

2.1 Data Sources

In order to significantly lower the cost of data acquisition, public domain data sources were used. Agencies such as United States Geological Survey (USGS), NOAA, US Army Corps of Engineers (USACE), National Aeronautics and Space Administration (NASA), or State government usually provide their acquired data to public for free. LiDAR survey is often done by government for geological, hydrological and forestry management purposes. It would be the best free data source that meets our needs for this
research. In the state of Ohio, state-owned public surveying data sets are managed by Ohio Geographically Referenced Information Program (OGRIP), a department under the Ohio office of Information Technology (OIT). LiDAR data is a part of the Ohio Statewide Imagery Program (OSIP). OSIP consists of 6-inch pixel resolution and 4-band (RGB NIR) color imagery, 1 foot or higher pixel resolution Oblique Imagery, 1M LiDAR, 5-foot contours (2-foot in some regions), and land use / land cover development (OGRIP, 2011).

In this research, the extracted shoreline is corresponding to the shoreline on the imagery used. Thus, deciding which image to use totally depends on which shoreline on image you would like to extract. For instance, when one needs a historical shoreline of 2002, a 2002 satellite imagery would be used in order to extract the 2002 shoreline. World View 2 is currently the only sub meter resolution satellite equipped with 8-band multispectral imagery. This satellite imagery data set would be the only data source used in our proposed approach with a price tag.

2.1.1 LiDAR Systems

LiDAR stands for Light Detection and Ranging, which is a remote sensing technology. It could provide the distance between the target and the scanner itself. The scanner emits a light beam (usually laser) toward the target and receives the beam reflected back from the target. Then the distance is calculated by measuring the time that light travels from the scanner to the target and gets back. The scanner also keeps track of the direction of the light been emitted. With the direction and the distance, the relative coordinates of the target could be determined (Vosselman and Maas, 2010). Nowadays, a
LiDAR system is sometime informally called Laser Scanner or Laser Scanning Device. This reveals the scanning function of the LiDAR system. Thus, a land-base (terrestrial) or airborne LiDAR system usually can scan a wider area by adjusting the laser emitting angle instead of measuring one single target point. For terrestrial LiDAR systems there is usually a vertical scanning scanner mounted on a horizontal rotating motorized platform (Figure 2.1). It can measure 360 degree for surrounding objects and create a point cloud under the local coordinate system. One could measure the coordinates of the LiDAR scanner in the object space and for every single point the coordinate can be translated into the object space. As a result, terrestrial LiDAR could easily produce a point cloud of surrounding regions. The distance of the region is confined by the limited power of the laser pulse (Vosselman and Maas, 2010) and line of sight. In order to measure a larger region, the scanner can be placed in multiple locations and the point clouds can be combined together into the same coordinate system. This terrestrial LiDAR can be very useful for building modeling, city modeling, or building interior modeling applications.
Air-borne LiDAR is also called Airborne Laser Scanning (ALS) System (Schenk, 2001). The theory behind the air-borne LiDAR is exactly the same as terrestrial LiDAR. The difference would be the direction of the laser beam and the constant moving location of the scanner. Air-borne LiDAR emits light from the airplane towards the ground in an angle perpendicular to the flight line (Figure 2.2). When the airplane moves forward, the scanning points form a zigzag pattern on the ground. The location of the scanner where every laser pulse emitted from is estimated by the onboard GPS / INS system. The light emitted from the scanner is usually laser with a small divergence angle. In an air-borne application, after the emitted light travels a few hundred meters in the air, the footprint on the ground is about the size of half a meter to one meter, depending on the distance between LiDAR and the target. If there are several objects within the footprint, especially objects with various elevations, the pulse would be divided and reflected back to the
scanner at different heights, resulting in multiple LiDAR points returned. Modern LiDAR scanners have the ability to distinguish multiple returns within a pulse. However, due to the Omni-directional property of the LiDAR receiver, we can only know there are several different elevations of objects reside in this footprint area, no way to tell the exact part of the footprint. As a result, the error of LiDAR point cloud is highly related to the size of footprint. Most of the LiDAR multiple returns were caused by vegetations, and some of them occurred on the edge of the buildings (or manmade objects). In these areas, part of the laser footprint reflected from the top of the trees, some might be reflected by the leaves or tree trunks, or eventually the ground. The older LiDAR system can detect only one return, but newer systems can record two returns (the first and the last), four returns (the first, the 2nd, the 3rd, and last), or more. The latest LiDAR system can directly record entire laser reflection waveform. The ability of receiving multiple return of the LiDAR system also leads to opportunities for new applications. Researchers are now using LiDAR point cloud data for canopy analysis related to forestry researches. The analysis of the full waveform can tell multiple distinct returns, and also determine other properties of surface, such as ground slope, roughness, materials, etc.(Mallet and Bretar, 2009). However, it still has weakness such as lack of laser points on vertical surfaces such as building facades.
Figure 2.2 The concept of an air-borne LiDAR system (Image Credit: Optech)

Each surveying technique has systematic errors, and LiDAR is no exception. Distance measurement error for LiDAR system is totally relying on the time measurement device of the receiver. One nanosecond of measuring error may cause 15
cm of ranging error. With the modern LiDAR system, a few centimeters of accuracy can be reached. For a terrestrial LiDAR system, the scanning angle and the rotary platform may also lead to scan angle measuring errors. On the other hand, with air-borne LiDAR there may be errors caused by scan angles and distance measurement as well. In addition, the GPS/INS causes location and attitude systematic errors. To remove systematic errors in traditional surveying techniques, redundant observations from a different device or aspect are often used. However, LiDAR is not able to measure the same point twice or more times because the exact location on the ground where the laser beam is going to hit is uncertain. Even though strip adjustment systematic error estimation process is widely used in the industry, it still cannot eliminate systematic errors completely (Schenk, 2001). As a result, systematic errors still reside in the produced data set and we cannot treat LiDAR point cloud data as true values. These are the advantages and disadvantages we need to understand before using a LiDAR point cloud as data source.

A LiDAR system utilizes laser as a tool to measure distance. Light is actually a form of electromagnetic waves. Properties differ with frequencies of electromagnetic waves when passing through substances. There are two major frequencies of electromagnetic waves that have been used in LiDAR systems: 1064nm (NIR light) and 532nm (green light). NIR light is the most popular type of laser used by LiDAR systems. These two types of light show no significant difference on land surveys, but huge differences where water is present. When light hits the water surface, part of it is reflected by the water surface and part of it is transmitted into the water. Since green light cannot be absorbed by water, the light transmitted into water will keep traveling until it hits a solid object, for example, a sea floor, and then is reflected back to the surface and
eventually to the LiDAR scanner. But in the case with NIR light, the light transmitted into water will be absorbed by water and nothing will be reflected back to the scanner except for the light reflected by the water surface.

The green-light LiDAR system is formally called the “Bathymetric LiDAR” and the NIR light LiDAR system is called the “Topographic LiDAR”.

The Scanning Hydrographic Operational Airborne LiDAR Survey (SHOALS) system is the first Bathymetric LiDAR system. This system was developed by the US Army Corp of Engineers (USACE) (Irish et al., 2000) in 1994 and manufactured by the Optech Inc. (Optech, 2012). SHOALS is the most commonly available commercial Bathymetric LiDAR system. NASA also developed their own version of Bathymetric LiDAR system called Experimental Advanced Airborne Research LiDAR (EAARL) system. The major improvements of EAARL are "(1) a relatively short (1.3 ns) laser pulse, (2) a radically narrowed receiver field-of-view (FOV)(1.5-2 mrad), (3) digitized signal temporal backscatter amplitude waveforms, and (4) software as opposed to hardware implementation of real-time signal-processing elements" (USGS, 2012). From our experience with these two data sources, EAARL system has a more successful rate of penetrating water surfaces than SHOALS system.

NGS uses EAARL system to develop their Bathymetric LiDAR shoreline delineation procedure. They first transform vertical datum from ellipsoidal height into Orthometric height (NAVD for instance). Then a DEM is created by the LiDAR point cloud in orthometric height datum. Finally, a tide-coordinated water level is used (MHW water level for example) to draw a contour on this DEM. This contour would be the shoreline at this water level (MHW shoreline) (White, 2007).
A Bathymetric LiDAR system is currently much more expensive than topographic LiDAR and as well as the data sets created. Moreover, government provided LiDAR dataset for the public is usually a topographic LiDAR dataset, since the major application for LiDAR data is usually not Hydrology related. Bathymetric LiDAR is only used when a dataset is used specifically for Hydrology related applications. For the purpose of lowering the cost of data acquisition, topographic LiDAR is a reasonable choice of data source.

Researchers at Mapping and GIS Levarotary, Ohio State University have developed bluffline delineation procedures based on LiDAR and aerial orthophotos (Liu et al., 2009). First, create profiles along the shoreline and detect the bluff top and bluff toe location on a 2D profile. Then, project these bluff top and toe points back to 3D space and connect these bluff top points and bluff toe points separately. Then, Iterative Closest Point (ICP) algorithm is used to match these bluffline with the linear features that is extracted from the aerial orthophoto. Choung (2009) further improved this procedure to detect 3D break lines solely from LiDAR point cloud. First, 2D Triangular Irregular Network (TIN) is established using Delaunay triangulation. Next, averaging the slope within the neighboring triangle and measure the slope change between triangles. Finally, detect the break line by the slope difference between triangles. The different of folding angle aspect could be used to distinguish the bluff top and bluff toe.

2.1.2 World View 2 Satellite Imagery

World View 2 is the latest remote sensing satellite provided by DigitalGlobe. The difference between this latest satellite and the previous Quickbird satellite is not just the
resolution improvement. The spectrum resolution of this satellite is also increased to cover the new additional color bands. There are one panchromatic band and 8 multispectral bands. These multispectral bands are Coastal Blue, Blue, Green, Yellow, Red, Red Edge, N-IR 1, and N-IR 2 (Figure 2.3). Moreover, the spatial resolution is also improved from 0.7m (Quickbird) to 0.5 m (World View 2) resolution (Panchromatic band, 2.0m in multispectral band). Image accuracy is also improved as well, from 23m (Quickbird) to 6.5m (World View 2) (DigitalGlobe, 2010).

![Image](image.png)

Figure 2.3 Spectral bands of different satellite images provide by DigitalGlobe

### 2.2 Fundamental Algorithms Used in this Research

The new approach developed by this research is based on several well known or modified reliable algorithms. These algorithms include the Mean-Shift segmentation algorithm, the Modified Convex Hull algorithm, the Shoreline comparison method used through out this research, and the Normalized Difference Vegetation Index (NDVI). The
Mean-Shift segmentation algorithm used in both Phase One and Phase Two for segmenting LiDAR data points and multispectral imagery. Modified Convex Hull is used in Phase One for tracing the boundary of the LiDAR points classified as land. The NDVI is used both in Phase One and Phase Two for separating the land and water surface.

2.2.1 The Mean-Shift Segmentation Algorithm

Mean Shift is a non-parametric iterative algorithm which can be used for density estimation, mode finding, and segmentation. It was first introduced by Fukunaga and Hostetler (1975), and then Comaniciu and Meer (2002) introduced Mean Shift to the computer vision field as a segmentation algorithm. The basic concept for the Mean Shift algorithm is finding the local data point maximum density location by shifting the searching window toward a denser region. There are several procedures involved in this process including Kernel density estimation, density gradient estimation, and a condition for convergence. The kernel density estimator with kernel K(x) when n data point x_i, i=1,…,n in the d-dimensional space is given as:

\[
\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K \left( \frac{x-x_i}{h} \right)
\]

(2.1)

where h is the d-dimensional bandwidth of this kernel density estimator. The kernel could be uniform (rectangular), Gaussian, Epanechnikov, or of any other shape depending on your application and preference. In practice, the Epanechnikov kernel is widely used in dataset or image segmentation (Comaniciu and Meer, 2002). The equation
of the Epanechnikov kernel in one dimension can be represented as:

\[ k_E(x) = \begin{cases} 
\frac{3}{4}(1-x^2) & 0 \leq x \leq 1 \\
0 & x > 1 
\end{cases} \]  

The density gradient can be represented as below

\[ \bar{x} = \frac{\sum_{i=1}^{n} K'(\frac{x-x_i}{h})x_i}{\sum_{i=1}^{n} K'(\frac{x-x_i}{h})} \]  

(2.3)

Assuming \( g(x) = -K'(x) \), the mean shift can be represented as

\[ m(x) = \frac{\sum_{i=1}^{n} g(\frac{x-x_i}{h})x_i}{\sum_{i=1}^{n} g(\frac{x-x_i}{h})} - x \]  

(2.4)

Figure 2.4 shows an example of how the segmentation is done in concept. The black bars along the axis represent the input samples. \( \hat{f}(x) \) is the estimate of kernel density, \( K(x) \) represents the kernel, \( g(x) \) is the derivative of the kernel, and \( m(x) \) represents the vector of the mean shift. The whole idea of the mean-shift algorithm is to find the direction toward the local maximum using discrete samples. How an algorithm like this can help us perform data segmentation? The procedure for Mean-shift
Segmentation is summarized below.

Figure 2.4 One-dimensional visualization of calculating a mean shift (Szeliski, 2011)

1. Define a type of kernel to be used for density estimation, usually a uniform, Gaussian, or Epanechnikov with a predefined radius (the bandwidth). This is the parameter of this algorithm, with one for each dimension;

2. Randomly select a data point within the samples as the initial location of the mean-shift algorithm;

3. Calculate the mean shift vector at this data point;

4. Apply the mean shift vector and move the kernel center to the new location, and repeat steps 3 to 4 until the mean shift vector becomes insignificant;

5. Defines the mode of a segment with this location. Each point within the radius of the kernel along the trajectories converges to the same segment; and

6. Repeat steps 2 to 4 until every point in the dataset is assigned to a segment.
This Mean Shift Algorithm had been used in our method for image and LiDAR point cloud segmentation. In LiDAR point cloud segmentation, with the multispectral data retrieved from the satellite imagery, the segmentation is a multi dimensional segmentation with X, Y, and multispectral and panchromatic bands form the satellite images. For each of these bands, the bandwidth for the kernel is the same. The method to assign different radiiuses to different dimensions is to normalize the data in different dimensions. This part of the coding is done with MATLAB by modifying a set of Mean-Shift Clustering codes developed by Bart Finkston (2006) released on MATLAB File Exchange website. The major modification of this set of codes done by this research is reducing the usage of memory and improving the computation performance (utilize parallel computing) of the codes in order to make it possible to apply the Mean-Shift Segmentation to a huge dataset such as LiDAR point cloud. As for image segmentation, the Edge Detection and Image SegmentatiON (EDISON) implementation were applied. The EDISON was developed by Robust Image Understanding Laboratory (RIUL) at the Rutgers University. This implementation was developed based on Comaniciu and Meer’s (2002) research, utilizing Mean-Shift segmentation on image only. This set of codes was developed in C++ language and wrapped into a MATLAB function developed by Shawn Lankton (2007) in order to use in MATLAB. The kernel of the Mean-Shift used in EDISON is of an ellipsoid shape with different radiiuses in image space and color space, and both of the radiiuses could be defined separately. The radius of kernel in image space is called Spatial Bandwidth, while in color space is called Color Bandwidth. These two are the primary parameters for EDISON. In later chapters when this EDISON algorithm is applied, only these two parameters would be mentioned.
2.2.2 Modified Convex Hull

The Modified Convex Hull algorithm (Sampath et al., 2007) is used to trace the boundary of the shoreline. This algorithm is modified based on the convex hull algorithm proposed by Jarvis (1977). The difference between this modified and the original convex hull algorithms is the extracted boundary. Instead of extracting the convex hull (Figure 2.6 second image in row #4), the boundary line traced from the modified convex hull could closely follow the boundary created with a set of points (Figure 2.6 third image in row #4). To achieve this goal, a searching window is set during the edge tracing procedure to make sure the line linking process is linked to neighboring points only (Sampath et al., 2007).

Before getting into the modified convex hull algorithm, first we need to understand the convex hull algorithm (Figure 2.5).

1. Find the left-most point that belongs to the boundary and assign this point as starting point P.
2. Create a line segment from point P to all the other points in this dataset. Calculate the clockwise angles between the vertical axis and all the segments, and select the segment with the smallest angle as part of the convex hull and the point connected to this segment (A) would be the next point on the convex hull.
3. Create a line segment from point A to all the other points except for the points belong to the convex hull in this dataset. Calculate the clockwise angles of the previous convex hull segment (AP) and all the other segments, and select the segment with the smallest angle as part of the convex hull and the point connected to this segment would be the next point on the convex hull.
4. Repeat step 3 until the point selected in step 3 is the very first starting point, then stop the process.

Figure 2.5 The convex hull algorithm (Sampath et al., 2007)
The steps of the modified convex hull algorithm are as follows (Figure 2.6):

1. Find the left-most point that belongs to the boundary and assign this point as starting point P.

2. Find the subset of the point cloud which is the neighboring point of point P.

3. Keep searching the boundary point based on the convex hull algorithm, with the constraint of the new segment created with the new point not intersecting the already identified convex hull segments.

4. After a new boundary point is selected, repeat step 2, 3, and 4.

5. Continue the process until the point selected in step 4 is the very first starting point.
Figure 2.6 The procedure of modifying the Convex Hull algorithm and the results compared with the traditional convex hull algorithm (Sampath et al., 2007)

The outcome of modified convex hull algorithm is significantly different with the traditional convex hull algorithm. The traditional convex hull algorithm leads to a polygon that surrounding all of the points, the modified convex hull algorithm on the other hand, leads to a polygon that closely contacts with the boundary point of this point.
cloud (Figure 2.5 row #4). In the case of tracing shoreline from a subset of LiDAR points, the result of modified convex hull algorithm could provide a better representation of a boundary line since the line is closely contacted with the boundary points.

Shorelines are usually polylines instead of polygons. This modified convex hull algorithm is used to trace the boundary of the LiDAR points on land. Since the boundary of the land is a polygon and the shoreline is just a part of it, a trimming process is needed to trim out the polygon which is at the edge of the LiDAR data set. After trimming, the resulting polyline would be the part of the polygon that would be considered as the shoreline.

2.2.3 NDVI

NDVI stand for Normalized Difference Vegetation Index. Like the name of this algorithm suggests, it is mainly used for vegetation detection. The vegetation tends to absorb visible light and reflect NIR light (Figure 1.9). Hence, an area found with high intensity in an NIR band and low intensity in a visible band would be vegetation. On the other hand, water absorbs N-IR light, the property in NIR band is the opposite of vegetation. As a result, vegetation and water surfaces are on the opposite end of the NDVI value.

This is the formula of calculating NDVI (Tucker and Sellers, 1986):

\[
NDVI = \frac{I_{NIR} - I_{RED}}{I_{NIR} + I_{RED}}
\]

(2.5)
where I represent the intensity either in NIR band or RED band. The NDVI value varies between -1.0 and +1.0 by design. In general, the same set of threshold can be applied to datasets acquired by the same satellite with the same hardware and similar processing procedure. Consequently, the threshold suggested in this research, is based on the World View 2 satellite imagery. The Classification of vegetation and water surface area is based on the NDVI indicator in this research. This algorithm would be primary method use to classify water surface and vegetation from satellite imagery. For the purpose of shoreline delineation, water surface class and land class are the most important classes. Vegetation class is one of the major subset of Land class. Thus, NDVI could provide two of these major classes that we needed for shoreline extraction.

2.3 Types of Coastal Terrains and Determination of Ground Truth

With this new proposed approach, different types of terrains along the shoreline need to be identified and processed separately during the accuracy estimation and shoreline integration. As a result, shoreline terrain types need to be identified when determining ground truth for shoreline accuracy estimation. The ground truth of the shoreline is determined by manually digitizing the shoreline from the panchromatic band of the satellite imagery, the same satellite image used in our shoreline line extraction procedure. When manually digitizing shoreline, for some regions, exact shoreline locations may be quite difficult to determine. In these cases, multispectral bands, orthophoto and / or LiDAR point cloud are used to help with manual digitization. Since the satellite imagery we used is not tide-coordinated, we could only partially follow the principle of shoreline delineation guidelines defined by the NGS, NOAA (NGS, 2011).
The part of the guideline that we are following is the determination of the shoreline location. The principle defined by the NGS is using the aerial imagery acquired at the tide of mean high water level and delineating the mean high water line on the stereoscope environment. The identification of mean high water line on aerial photograph is the closely correlated physical evidence, such as debris lines or the wet-dry sand abutment (Floyd, 1995) and subject to human interpretation (White, 2007; NGS, 2011). Since satellite imagery could not provide elevation information like the stereo aerial images did, the wet-dry line is used as the ground truth shoreline. For areas where the wet-dry line cannot be observed from satellite image, most of them are manmade vertical structures. Under this circumstance, the edge of manmade vertical structures is considered as the shoreline, as suggested by the North Carolina Department of the Environment and Natural Resource (NC DENR) Division of Coastal Management (DCM) shoreline delineation guidelines (Geis and Bendell, 2010). We also assigned terrain types to the delineated ground truth shoreline for latter analysis of the shoreline accuracy and terrain type relations. Seven (7) terrain types have been chosen to represent the terrain types within our study area. Six of them are corresponding with the terrain types that defined by NC DCM. These terrain types are: sediment bank (Figure 2.7 a), vertical structure (Figure 2.7 b), sloped structure (Figure 2.7 c), groin (Figure 2.7 d), breakwater (Figure 2.7 e), and pier (Figure 2.7 f). One terrain type in addition to NC DCM is the bluff (Figure 2.7 g).
Figure 2.7 The types of shoreline that are available in study area. (a) Sandy beach, a typical sediment bank. The shoreline should be the wet/dry line. (b) Bulkhead is a typical form of vertical structure (Image credit: Geis and Bendell, 2010). (c) Riprap revetment is of a typical sloped structure (Image credit: Geis and Bendell, 2010). (d) Groin is designed to trap sand and create sediment bank. (e) Sand has build up between these two breakwaters (Image credit: Geis and Bendell, 2010). (f) A typical pier structure (Credit: Geis and Bendell, 2010). (g) A typical bluff along the south shore of Lake Erie
Figure 2.7 continued

(b)

Continued
Figure 2.7 continued

(d)

(e) Continued
Figure 2.7 continued

(f)

(g)
Chapter 3. Shoreline Extraction from LiDAR Data and Satellite Imagery

In this chapter, the phases of shoreline extractions from LiDAR data and satellite imagery are presented. Section 3.1 described the overall concept of the proposed approach; section 3.2 and 3.3 described the shoreline extraction procedures and parameters for each phase in detail. Shoreline accuracy estimation procedure is also described in section 3.4.

3.1 Methodology of the Proposed Shoreline Extraction Approach

The data sources used in this research are the LiDAR data, World View 2 satellite imagery panchromatic and multispectral bands. The three data sources used in this research are of different resolutions and accuracies. The advantage of the LiDAR data is the high accuracy in vertical coordinate (5cm - 30cm). However, the horizontal accuracy is 2-5 times lower than the vertical one (Ackermann, 1999). On the other hand, the 0.5-meter-resolution WorldView2 satellite imagery can potentially improve the accuracy in horizontal coordinates if both of these data sources are integrated. This dissertation proposed an approach to achieve the integration.

The proposed shoreline extraction approach is divided into three phases. In Phase One, the shoreline is extracted based on the LiDAR points that are classified as land with additional multispectral information from the satellite imagery used as features of
classification. In Phase Two, spectrum matching, shadow analysis and NDVI are applied to classify the pixels in multispectral satellite imagery. Then shoreline is delineated by tracing the boundary of the water surface class after the satellite image classification. The major difference between Phase One and Phase Two are the data that we are classifying. Phase One is classifying LiDAR points, Phase Two is classifying image pixels. Phase Three is integration of these shorelines utilizing shoreline types and scenario analysis of the changed shorelines. The methodology to incorporate the three phases is showed in the flow chart in Figure 3.1.

Figure 3.1 The conceptual flow chart of the proposed shoreline extraction approach
3.2 Phase One: Shoreline Extraction Based on LiDAR Data with Multispectral Information

In Phase One, the shoreline extraction process is based on the LiDAR point cloud. The process is to classify LiDAR points into land and water surface points through NDVI and Mean-Shift segmentation. The entire process consists of several sub-processes. These sub-processes include LiDAR point cloud preprocessing, LiDAR point cloud and satellite imagery registration, classification of LiDAR points, shoreline extraction from the classified LiDAR points, and the accuracy estimation of the extracted shoreline. The flow chart of the entire process is shown in Figure 3.2. Before discussing the advantages of our solution, we need to know other available solutions. Other shoreline extraction methods using LiDAR data mostly utilize LiDAR point cloud to generate DEM, in either raster or TIN. Then the shoreline is extracted using a predefined water level to intersect the DEM and create a contour line representing the shoreline. The major issues for these methods are: 1) while creating DEM from LiDAR point cloud, LiDAR points are treated as true values without errors; 2) before creating DEM, filtering or even manually editing of the LiDAR point cloud is required in order to create a reasonable DEM (Vosselman and Maas, 2001); and 3) creating a contour line from DEM may result in an irregular shape (jagged shape) of lines, so additional processes are needed to remove or smooth out these irregularities and generate unpredictable inaccurateness. The distinguish feature of our proposed procedure in Phase One is extracting the instantaneous shoreline in the LiDAR data by classifying LiDAR points into land and water surface subsets. Under our framework, there are still systematic and random errors in the LiDAR point cloud during
classification. Hence, a median filter to remove the random errors is sufficient. Furthermore, no additional smoothing process is needed for the extracted shoreline. This solution not only saves time and labor required for data preprocessing, but also leads to no additional errors with the extracted shoreline.
Figure 3.2 The flow chart of the shoreline extraction procedure in Phase One
3.2.1 LiDAR Point Cloud Preprocessing

There are basically two issues needed to be handled in this process. The first one is to identify the LiDAR points with multiple returns. The second one is to remove the random error points. Usually, the work of assigning LiDAR multiple-return information to LiDAR points is done by the data provider. However, the data we used in this research did not provide any LiDAR multiple-return information. Hence, we had to handle it by ourselves.

In this research, the terrain surface was what we needed and it is usually represented by the last LiDAR return. Most of the LiDAR datasets do provide the information of the Number of Total Returns for each given laser pulse (NTR) and the Current Return of the laser pulse (CR). In this case, when NTR was equal to CR, it means that was the last return. However, the LiDAR dataset that we used in this research, there is no information of NTR or CR. Under this circumstance, we have to figure out which are the last return LiDAR points by our-self. These are the proposed solutions:

a. Instead of finding a set of points are multiple returns in a numerical way, we could obtain it geometrically. Slope, roughness of the surface, and vertical structures were used to estimate this set of points on terrain surface, a building, or a tree. However, the robustness of the result might not be good. It highly depends on the LiDAR point spacing, distribution of the LiDAR points, and in-situ environment.

b. Since multiple returns were generated by the same pulse, all of the multiple return of LiDAR points belong to the same pulse shared the same location of origin. In other words, the location of the laser origin was invariant, as well
as the scan angle. The only difference among these multiple return points was the distances between the object and the laser scanner. Since the LiDAR points were recorded sequentially by the scanning order, the multiple return points should be in neighboring records. Two neighboring records of the LiDAR points forms a line. If this line intersected with the flight path, then these two LiDAR points were multiple returns of a single LiDAR pulse. After processing the entire LiDAR point cloud, all multiple returns within a pulse for the whole point cloud were identified. Within each pulse, the LiDAR point which was the closest to the laser scanner was the first return, and all the rest multiple returns were assigned accordingly. Whether this method worked depended on the quality of the LiDAR data. If the decimal places of the LiDAR coordinate were enough and the calculation of the coordinate was accurate enough, this method would have worked. However, the LiDAR dataset that we used for this research could not provide the precision we needed to calculate if the LiDAR points are originated from the same pulse.

c. A simpler version of the previous method is to find the closest neighboring LiDAR points within a predefined distance, and calculate height difference between them. If the height difference is larger than the predefined value, then this set of points are multiple returns. Since the last return was what we needed, we kept the point with the lowest elevation and removed all the other returns.

Removing LiDAR points with random error is another major preprocessing
process. There are several types of situations that may cause LiDAR points with random error, such as LiDAR laser beam hitting a flying bird or an object, sun glint influence on the LiDAR receiver, and breaking waves on water body which might interfere the laser pulse. The solution for this is commonly using median filter filtering through the LiDAR point cloud (Liu et al., 2009). First, a radius for the searching window must be defined. Secondly, for each LiDAR point, all points within this window size are obtained then the elevation is considered as the median elevation within this set of LiDAR points. After this process, the LiDAR point cloud should be free of random error. Although, some of the small features such as a rock on the shore or a chimney on the rooftop might be smoothed as well, structural features of the terrain were our main concern.

3.2.2 LIDAR Point Cloud and Satellite Imagery Registration

Datasets are usually acquired independently, may uses different sensors and using different coordinate systems. Registration between datasets is to create transformations for datasets and put them into a same coordinate system. Although coordinated systems can be transformed, they could still do not match with one and other perfectly after transformation. In order to integrate multiple datasets, the registration between datasets is crucial. This discrepancy may be caused by systematic errors, acquired sensor differences, acquired platform differences, and datum differences. In case of our research, LiDAR and satellite imagery were totally different kinds of sensors, one was an air-borne platform while the other was space-borne. In addition, both LiDAR point cloud data and satellite imagery were georeferenced by the data providers, but they were under different datum and coordinate systems. There are several solutions to register between satellite imagery
and LiDAR point cloud. The most common way is to manually find corresponding points between these data sets and establish a transformation equation. There are several typical transformation models such as translation, similarity transform, affine transformation, projection transformation, and non-linear transformation. To determine which transformation model to use depends on the physical transformation model between data sets. In this research, the LiDAR data set that we used was in NAD83 and the WV2 imagery is in UTM. Both of these coordinate systems are based on GRS 80 ellipsoid. Within a small area, the transformation was linear and it is very likely that an affine transformation could satisfy the discrepancy. In this research, the coordinate system transformation was done using ArcGIS. We did not deal with the coordinate system transformation directly. However, after the coordinate system transformation was done, there were still discrepancies between the datasets. A set of five manually determined corresponding points were used to obtain the transformation parameters. After estimation, there was a 1~2-meter shift (translation) between these datasets. The amount of the offset was determined and the LiDAR point cloud was shifted during the registration process.

After the process of LiDAR and WV2 imagery registration were completed, the LiDAR points were projected onto the WV2 satellite imagery and then the multispectral intensity was assigned to each LiDAR point. In the end, every LiDAR points was equipped with X, Y, Z and R, G, B, NIR, PAN retrieved from the satellite imagery. These values retrieved from satellite imagery multispectral band would be essential for the LiDAR point cloud classification.
3.2.3 LiDAR Point Cloud Classification

The concept of this binary classification of separating LiDAR points into water surface and land is to “divide and conquer”. Lands are composed of several major objects such as vegetations, sand beaches, manmade structures, and other miscellaneous objects. On the other hand, water surfaces are composed of submerged vegetations, wet sand (the NOAA defined a shoreline as the high water line, where the wet sand areas are a part of the water surfaces), breaking waves, and ripples. Consequently, it is almost impossible to find an algorithm to achieve this binary classification all at once.

Table 3.1 shows the major objects that are available in a coastal region. These objects show various properties in data from different sources. Within the same data source, the objects are easily mixed with one another if they have the same description in Table 3.1. For example, land vegetation, wet sand, water surface, and submerged vegetation can be mixed up in a panchromatic band. The methodology of this classification is to establish a decision tree: First, NDVI is used to separate the land vegetations. Secondly, LiDAR elevation is applied to classify objects of the lowest elevation, which are the water surface and submerged vegetations. After these two processes are completed, wet sand objects are still not classified as water surface. Hence, panchromatic and NIR bands of the satellite imagery are used to separate the wet sand class from all other objects (Figure 3.2).
The NDVI indicator can be used to identify sea grass (Hill et al., 2007). The spectral curve of sea grass is different from vegetation on land (Figure 3.3). The intensity of sea grass in the red band is stronger than water surface, but weaker than sand. In the NIR band, the intensity of sea grass is about the same as water surface, but relatively weaker. However, the intensity of vegetation on land is very strong. Hence, the NDVI value of land vegetation would be close to 1, that of water surface would be close to zero, that of sand would be close to -1, and that of sea grass would in-between water surface and sand. The empirical threshold for land vegetation would be higher than 0.5, we tested this threshold using our dataset and the performance was very good in separating water surface and vegetation. Hence the threshold was set to 0.5 for classifying land vegetations.
In the next process the Mean-Shift algorithm is applied to segmenting the LiDAR points based on elevation. Mean-shift segmentation is based on density function estimation. The bandwidth will influence the density function and it is also the major parameter for mean-shift segmentation. To determine the bandwidth for mean-shift algorithm is something of an art (Szeliski, 2011). In the early stage of this research, we focused on determining the bandwidths which may lead to the best segmentation (Lee et al., 2009; Lee et al., 2010). However, it was so complicated and sometimes subjective
based on the operator’s judgments. Hence we decide to use mean-shift segmentation to pick out one segment at a time. To segment the entire LiDAR dataset into preferred segments requires a bandwidth that can generate a density function that all segments are separated. If only one segment is needed, while the number of samples is huge and the distribution is narrow, the density function will approximately match the true distribution and the bandwidth would not matter according to the Parzen-Window density estimation theory (Duda et al., 2001). LiDAR points on water surfaces satisfy all these above-mentioned requirements. As a result, the selection of bandwidth is not an issue in this process. However, this concept does not apply to every class. In Phase Two, determining the bandwidth is still one of the major processes. The goal of this process is to find all the LiDAR points that belong to water surface, which are the LiDAR points of the lowest elevations. This is a one-dimensional segmentation, only the elevations of the LiDAR points are used. Figure 2.4 in section 2.2.1 shows exactly how this one-dimensional segmentation is performed. The water surface segment was the largest group and at the lowest elevation within this dataset. This segment was directly assigned to water surface and the remaining LiDAR points were mainly the sand and manmade objects. For sloped and vertical structures areas, the elevation difference provided a very good separation of the land and water surface. The major issue for the remaining LiDAR points was how to separate the shoreline boundary on sandy beach.

Since the shoreline that we wanted to delineate is the high water line, which is the wet/dry line on the image, the last process of the classification was to separate the wet and dry materials, mainly for shoreline in the sandy beach area. Within all these multispectral bands provided based on the WV2 imagery, NIR and panchromatic bands
were the best ones to differentiate the wet part and the dry part on the same material. Since the wet/dry separation can be seen clearly with the naked eye in these bands, the segmentation algorithm should be able to achieve the same result with well-evaluated parameters. Hence, NIR, and panchromatic attributes in the LIDAR data were used in the Mean-Shift segmentation for segmenting the remaining LiDAR points. This segmentation process was a two-dimensional segmentation featuring NIR and panchromatic information. The LiDAR points on the wet sand are classified in one class, and there are several other segments such as dry sand and buildings. From the segmentation point of view, this process is very similar to the previous process for segmenting water surface class. The distribution of the intensity in NIR and panchromatic band is also narrow, but the number of samples is not that large compared with water surface points. Hence the bandwidth needs to be determined. Under the condition of the WV2 image, 5, 10, and 15 could be used as initial bandwidths. If these bandwidths create multiple segments within the wet sand segment, than the value of bandwidth must be increased. If the separation of wet/dry sand does not correspond to any one of the segments, then the value of bandwidth must be decreased. According to the Parzen-Window density estimation theory and the distribution of the NIR and panchromatic band in wet sand areas, it should not be difficult to find a proper bandwidth for separating wet and dry sand. After successfully separating the wet and dry sand, by selecting the segment with the lowest average elevation, the wet sandy segment can be obtained. However, the shoreline on other types of terrain, such as sloped and vertical structures, did not benefit from this process.

Although LiDAR intensity has been widely used in data sets registration and
classification in topographic applications, we did not use it in any of the classification processes in this research. The reason that LiDAR intensity was not used in this research was the variety of LiDAR intensity on water surfaces. Figure 3.4 (a) shows a sandbar along the shore of the Lake Erie near Sandusky, Ohio. Figure 3.4 (b) shows the intensity map of the corresponding region. Red represents areas with strong intensity, and blue represents areas with weak one or areas without LiDAR points at all. The vertical black line in Figure 3.4 (b) represents the flight path of the airplane while the LiDAR data were acquired. The LiDAR intensity is very strong along the flight path on the water surface due to the water surface acting as a mirror while the laser pulse beamed toward the water surface. Under the flight path, the intensity is the strongest. Further away toward the perpendicular direction of the flight path, the intensity is weaker and eventually no LiDAR points were received on the water surface. The reason that the LiDAR points have this property on the water surface is because the LiDAR systems received the scattering laser pulses beamed from the LiDAR scanner. A mirror-like water surface will not scatter the laser beam. The laser beam is just reflected away from the water surface and toward the sky. Whether the receiver may or may not receive the reflected pulse is related with the scan angle of this LiDAR pulse. Hence the intensity is getting weaker while the receiver is getting less reflected photon from the laser pulse. Eventually, the LiDAR scan angle is further away from nadir till a limit where no signal is received. Figure 3.5 (a)–(d) show the profile of LiDAR point intensity along the four profile lines marked in Figure 3.4 (b). In Figure 3.5 (d) represents the intensity profile of the profile line 4, it’s near the boundary between land and water surface. If we could distinguish land and water surface using LiDAR intensity, there should be some dramatic variations where
land and water meet on this image. But in fact, the only indication of the boundary is the small undulation on the top in the graph around the 61m mark. It is almost impossible to be detected with any kinds of manual or automatic methods. On the other hand, elevation changes of coastal region are relatively significant even showing ripples on the water surface. Figure 3.6 shows the elevation profile along the flight path of the same location. However, the elevation profile also shows the significant systematic errors reside in our LiDAR point cloud. Figure 3.7 shows an elevation profile on the water surface perpendicular to the flight path. There is a significant slope on the water surface, the length of the slope is around 600 m, and the elevation difference is around 0.2 m.

Figure 3.4 Intensity map near Sandusky, Ohio. The vertical line is the flight path of the LiDAR platform.
Figure 3.5 The four intensity profiles corresponding to the four profiles in Figure 3.4. (a) Profile 1, (b) Profile 2, (c) Profile 3, and (d) Profile 4.
Figure 3.6 Elevation profile of the flight path. Elevation higher than 174.8 m belongs to land. Small undulation between 174.2~174.3 m is caused by wave. Elevation between 174.4~174.8 m is a breaking wave. The elevations stated here are manually inspected by comparing the LiDAR data with the orthophoto.

Figure 3.7 Elevation profile of Profile 1 in Figure 3.4
3.2.4 Shoreline Extraction from Classified LiDAR Point Cloud

LiDAR points were divided into water surface points and land points after the previous process. Since the LiDAR point cloud did not cover the entire water surface area, the shape of the shoreline may not be correct if we trace the shoreline using the boundary of water surface. Figure 3.8 shows an example of the distribution of LiDAR points on water surface. Consequently, shoreline is traced along the boundary of land LiDAR points. In this tracing process, modified convex hull algorithm was applied. This modified convex hull algorithm could follow the edge of LiDAR point cloud to obtain a clean boundary line. This boundary line represented the extracted shoreline.

Figure 3.8 The distribution of the LiDAR points on water surface. The LiDAR points are represented in blue dots. If the water surface not under the flight path during LiDAR acquisition, there may not be LiDAR points acquired.
3.3 Phase Two: Shoreline Extraction Based on Satellite Imagery

Although a shoreline can be extracted using LiDAR data, there are drawbacks such as the uneven distribution of LiDAR points, and LiDAR points may be missing partially on water surfaces, resulting in some of the narrow structures which should belong to land missing because the LiDAR point spacing is too wide. Furthermore, there is no way to control LiDAR points to always scan the edge of the land/water surface boundary. Precisely speaking, the extracted shoreline consists of the last land points scanned by LiDAR along the shore. As a result, the larger the LiDAR point spacing is, the larger the error would be. From the robustness point of view, a segment of shoreline may or may not be extracted just because one or few of the LiDAR points successfully scanned the critical areas. For example, if the area where the sea wall and the land meet is not scanned by the LiDAR, there will be no LiDAR points acquired around the critical area which is the root of the sea wall. Thus, the entire sea wall will not be extracted. Under these circumstances, this shoreline extraction procedure will not be considered as a robust algorithm. Additional information and procedures are needed. Aerial or satellite imagery on the other hand is raster data, the resolution is uniform, could provide an even distribution of the data. The uniform distribution of data points may help to lower the possibility of a highly accurate shoreline resulted from better “luck”. Satellite imagery provides multispectral information which an aerial photograph not always has. Multispectral information would help to identify the materials and determine the terrain types. In addition, it is also a data source of relatively low cost which needs less human labor for data pre-processing compared with aerial photogrammetry. Hence satellite
imagery is chosen by this research for shoreline extraction. The high resolution remote sensing satellite imagery used in this research is the World View 2 satellite. In Phase Two, a shoreline was extracted solely based on the WV2 satellite imagery. This process involves spectrum analysis, NDVI analysis, shadow analysis, segmentation and classification, and classification adjustments. Figure 3.9 is the flow chart of this process.

Figure 3.9 Flow chart of Phase 2: Shoreline extraction based on satellite imagery
This process could be considered as an image classification procedure. There are four classes of objects that are crucial to determine the high water mark for shoreline extraction. These classes are: shadow, water surface, vegetation, and sediment bank/manmade structure. There is no doubt that the water surface class would be the most important class for extracting the shoreline. However, other classes are needed for fine-tuning for the separation of water and land. Vegetation may be easily mixed up with water surface due to the similarity in reflectance with panchromatic band. If vegetation can be separated from water surface correctly, the accuracy of shoreline extraction would be improved. The shoreline is defined by the NOAA as the high water line, which is the wet/dry line on the beach (Leatherman, 2003; Floyd, 1995). The edges of coastal vertical structures such as piers and bulkheads are the distinct shorelines in manmade structure areas. As a result, the boundary of a sediment bank/manmade structure represents the shoreline better than the boundary of water surface within these areas. On the other hand, the need for the shadow class is totally different from the needs for other classes. Shadow may be on the water surface or on the land, the existence of shadow may have different meanings on different terrain types. For example, a north-facing vertical wall of the pier may create a shadow on the water surface if the sunshine comes from the south side. This shadow is considered as a part of water surface. Under other circumstances, such as a north facing bluff, blufftop creates a shadow on the sandy beach below, resulting in the sandy beach not being able to be identified. In this case, shadow is part of the land. Hence to determine the shoreline location within the shadow requires terrain type information.

The intensity of the satellite image is highly related to the geometry among the sun, topography, and the receiver (Hugli and Frei, 1983). Since the appearance of the
representation is not our concern, normalizing the intensity is not necessary. However, the same material may not show the same intensity in every multispectral band. As a result, performances of traditional classification algorithms classifying object using the intensity similarity may not be well. In this research, we performed classification using a supervised classification algorithm to identify the four classes that we described earlier. First, their separability is calculated using ENVI. The results show that all classes are separable, but the differences are smaller in shadow, vegetation, and sediment bank/manmade structure classes. Then, multiple algorithms are applied, including Parallelepiped, Minimum Distance, Mahalanobis Distance, Maximum Likelihood, Binary Encoding, Neural Network, and Support Vector Machine. The results show that the sediment banks may be classified relative well using these algorithms due to the intensity which is nearly constant in every band. But besides the result with sediment banks, other classification results are not so ideal. Water surface, shadow, and vegetation are not separated very well, and objects that should not belong to these four classes are classified into these classes. Figure 3.10 shows two locations of classification result. Figure 3.10 (a) and (c) uses WV2 multispectral band performing maximum likelihood supervised classification. (b) and (d) uses pen-sharpened WV2 multispectral band also performing maximum likelihood supervised classification. As can be seen from the image, water surface class, shadow class and sediment bank/manmade object class are mixed with one and other.

After this supervised classification experiment, it is found that a decision tree may be a better solution for classifying these four objects. Since we know the materials of the objects to be classified, proven indicators should be used. These indicators include NDVI.
and spectral matching. The threshold for NDVI was evaluated by this research and we developed our own criteria for spectral matching procedures. We also developed a new method for classifying shadow areas to separate water surfaces from shadows.

After assigning all image pixels into the four classes or the class of unclassified pixels, the classification result on the multispectral images is obtained. The locations of these classes are transformed from a lower resolution multispectral imagery onto a higher resolution panchromatic imagery. Since a further segmentation of the panchromatic image is needed, whether to perform the classification on the multispectral image of lower resolution or a processed pan-sharpened multispectral imagery does not matter. In order to reduce the computation time and preserve the original information, the original multispectral image is used. The detailed procedure is described in section 3.3.1~3.3.4. In addition, a classification refinement process based on physical meanings was developed and applied before the shoreline extraction (Section 3.3.5). The accuracy of the extracted shoreline was evaluated and the results are summarized in section 5.3.
Figure 3.10 Classification example using supervised classification algorithm. (a) Maximum likelihood classification on WV2 multispectral band. Shadow area and vegetation are mixed with water surface class. (b) Maximum likelihood classification on pen-sharpened WV2 multispectral band. Water surface band overflow to sediment bank/manmade object class. (c) Maximum likelihood classification on WV2 multispectral band. Shadow area and vegetation are mixed with water surface class. (d) Maximum likelihood classification on pen-sharpened WV2 multispectral band. Water surface band overflow to sediment bank/manmade object class and still mixed with shadow area.
Figure 3.10 continued

(b)

(c) Continued
3.3.1 Shadow Area Determination

There would be shadows on any passive remote sensing images taken in a bright sunny day. The influences of shadows would be one of the major drawbacks of remote sensing imagery, especially in shoreline extraction. Our test site Painesville is located at 41 degrees north latitude, by the south shore of Lake Erie. A higher latitude means a smaller sun angle, and mapping south shore of a lake means any obstacle along the shore (trees, buildings, coastal manmade structures, etc.) would lead to a shadow projected on the water surface. The worst case scenario would be the boundary between land and water surface is in the shade. The shoreline within shadows can hardly be identified, and even if in some area the shoreline could be slightly identified, the result is not reliable.
There are several kinds of features that may cause shadows along the shore: coastal vertical manmade structures, bluff tops, and trees along the shore. The shadows created by coastal vertical manmade structures such as bulkheads are usually on water surfaces. Without a proper design of the shoreline extraction method, edges of shadows instead of coastal vertical manmade structures (usually where the shoreline is located) might be considered as a part of the shoreline. Tops of the bluff may also create shadows on beaches and/or water surfaces down below shading the border between water and land. Shadows created by trees are even more complicated, they can be anywhere along the shore in different sizes. That's one of the reasons why most of the aerial photogrammetry projects are executed during the leaf-off season. However, the need for shoreline mapping is not always coincide with the leaf-off season.

Knowing the area is under shadows could be helpful when extracting the shoreline. With the knowledge of the objects surrounded, we can minimize the error caused by shadows as much as possible. Some of the examples of how shadows influence the shoreline extraction are describe below.

There are two images showing corresponding area in Figure 3.11~Figure 3.13. The images on the left are the WV2 panchromatic image; while on the right are the aerial orthophoto. In Figure 3.11, there are two dark regions around the center of the images. We can tell from the aerial imagery that the dark regions are the shadows of two trees. These shadows are partially on the land and partially on the water surface. Figure 3.12 shows a manmade vertical structure along the shore. Since the shoreline is under shadows, it's difficult to tell the exact location of the shoreline. Figure 3.13 shows a segment of the bluff line. There are usually sandy beaches below a bluff top, but it is hard to see them in
this image due to the shadows of the bluff top.

Figure 3.11 Shadows created by trees

Figure 3.12 Shadow created by vertical structure
From the examples above we can understand that separating shadow from water surface is the key of extracting an accurate shoreline, and it is also the most difficult task. Since shadows and water surfaces are usually the darker features in a panchromatic image, as a result, shadows and water surfaces are difficult to be distinguished and are often mixed with each other. We obtained the grey value difference between water surface and shadow on every band of the World View 2 satellite image in order to find the best band to distinguish water surface and shadow. After inspection, the largest intensity value difference between these two classes is on the Green and Yellow bands. On both of the bands, the intensity value of the water surface is about one hundred (100) higher than that of the shadows. The intensity value differences on all other bands are only about 20-50. According to this test, Green and Yellow bands are a very good candidate to distinguish water surfaces and shadows. The lower intensity of shadow and water surface areas are caused due to different reasons. The shadow is caused by the lack of sun light, the lower intensity would be uniform across all bands. On the contrary, the water surface
absorbs light and thus the water surface looks darker. However, the water surface absorbs more light in some electromagnetic wave band and less in others. Figure 3.3 shows that the reflectance rate of water surface is higher in-between 500~600 nm. This bandwidth of electromagnetic wave corresponds to the blue and green band in WV2. This also confirms our finding.

A solution which had been developed is to sum up both Green and Yellow bands and find the intensity value threshold manually. Finding a threshold can be done manually by selecting a relatively larger shadow area, and find the maximum intensity value within the shadow area. Meanwhile, the maximum intensity value of these shadow pixels shall not be mixed with intensity value of the water surface. These steps sum up the process of manually determining the shadow intensity value threshold. With our current data sources, the manually determined threshold is 350 after summing up the intensity value of Green and Yellow bands. The areas selected as shadow areas are shown in Figure 3.14, with the shadow areas marked in brown.
Figure 3.14 Multispectral image pixels are classified as shadow areas
Figure 3.14 continued
3.3.2 Water Surface and Vegetation Area Determination

Detecting water surfaces on multispectral satellite imagery has been widely studied in the remote sensing field. Algorithms such as Normalized Difference Water Index (NDWI), Modified NDWI (MNDWI), and Normalized Difference Pond Index (NDPI) are basically modifications of NDVI. These algorithms are modified for different data sources and different applications (Ji et al., 2009). In this research, all of the above-mentioned algorithms had been tested, the results showed no significant advantage compared to NDVI when with our data sources. Hence, NDVI was adopted to classify water surfaces and vegetations. The exact threshold for these two classes was determined by a trial and error process. First an empirical threshold was given, and the classification result was manually studied and then adjusted until all significant features were classified correctly. Vegetation is the area with NDVI values larger than the threshold 0.55, and water surfaces are areas with NDVI values smaller than the threshold -0.1.

After this process, water surface and vegetation classes were classified. However, water surface areas are easily mixed with shadow areas, especially calm water surfaces such as ponds or puddles. According to the National Climatic Data Center (NCDC) Climatological Data of Ohio at September, 2010, the WV2 imagery used in our research was taken two days after a rainy day with 0.19 inch of precipitation. Areas classified as water surfaces on the land are compared with the orthophoto and aerial photo, those areas are pit or uncovered drainage way. Hence, the water surface detected on land could be correctly classified, it is not always a false detected shadow area. After further investigation, the ratio between the Blue and Green bands could separate the water
surface and shadow more accurately. However, classification error may still reside in the final result. Figure 3.15 show the classification result of the water surfaces (light blue areas) and vegetation (green areas) areas.
Figure 3.15 Multispectral image pixels are classified as water surfaces and vegetation.
Figure 3.15 continued

(c)

(d)
3.3.3 Sediment Bank/Manmade Structure Area Determination by Spectrum Matching

Although World View 2 satellite is equipped with an 8 band multispectral sensor, it still cannot compare with hyperspectral imagery. The spectral resolution is not enough to distinguish every object in the image. However, we only need to identify several types of objects, and what we have implemented here to classify the WV2 image is the method of spectrum matching used in hyperspectral image classification. From the spectrum-reflectance chart which is a part of the spectrometer library we could know that the shape of sand on the spectrum-reflectance chart (spectral curve). When inspecting a manmade object such as a concrete building or a concrete surface, one must know that it also has a similar shape on the spectrum-reflectance chart within WV2 receiving spectrum (Figure 3.16 (a) and (b)). From the material point of view, silicon is the major compound in all these two objects. However, other objects such as water surfaces, vegetation, and building rooftop are totally in different shapes (Figure 3.16 (c) and (d)). This tells us that sandy beach might not be distinguished from concrete structures, but can from vegetation and water surfaces.
After the theoretical analysis of the spectrum, we then started to organize the criteria for finding sand and/or concrete on the WV2 imagery. First, we manually selected a set of sandy beach pixels and plotted the spectrum-reflectance chart (Figure 3.17). Then normalize the intensity of every band and plotted the normalized spectrum-reflectance.
chart (Figure 3.18). This reflectance chart may or may not be similar to the one created by spectrometer due to different sensors and different heights of locations of data acquisition. Second, we generalized the criteria from the normalized chart and tested on the satellite imagery. The blue line in Figure 3.18 represents the created criteria, which is the average normalized intensity of each band. The description of the criteria is listed as below:

- $I_{\text{Yellow}} > I_{\text{Green}}$
- $I_{\text{Green}} > I_{\text{Blue}}; I_{\text{Green}} > I_{\text{Coastal-Blue}}$
- $I_{\text{Yellow}} > I_{\text{Red-Edge}} > I_{\text{NIR-1}} > I_{\text{NIR-2}}$
- $I_{\text{NIR-1}} > I_{\text{Red}}$
- $I_{\text{Red-Edge}} > I_{\text{Coastal-Blue}}$

$I$ represent the intensity of the band in subscript. Any pixels in the image satisfy these criteria would be classified as sand.

![Figure 3.17 Spectral plot of the manually selected sand pixels](image)

Figure 3.17 Spectral plot of the manually selected sand pixels
Figure 3.18 After the normalization of Figure 3.17, the spectrum matching criteria are presented as the blue line.

After the developed criteria were applied in the spectrum matching process, several objects beside sand had been selected as well. These objects include sediment banks, concrete surfaces, sand piles, and partial concrete blocks (riprap revetments). Figure 3.19 show the spectrum matching result for a part of our study area. The areas marked in yellow represent the areas selected as a class by the spectrum matching process.
Figure 3.19 Multispectral image pixels are classified as sediment banks or manmade objects

Continued
Figure 3.19 continued

(c)

(d)
3.3.4 Classification Assignment on panchromatic image

After completing the three processes described previously, we now have 4 classified object classes on the multispectral imagery. The next step would be finding the corresponding regions of these classes on the panchromatic imagery. Since multispectral imagery is in 2m resolution and panchromatic is in 0.5m, assigning classes to panchromatic imagery is not a simple transformation or re-sampling of imagery. This process can be done using the pan-sharpened multispectral imagery to create the multispectral imagery of the same resolution as the panchromatic imagery. Since the goal of this procedure is to achieve the best classification in order to generate a data source for extracting an accurate shoreline. Thus, whether the classification result is accurate or not is not our major concern. Segmentation of the panchromatic image is needed to acquire a better boundary of the materials. Hence, whether to use the pan-sharpened multispectral imagery or to resample the original multispectral classification result does not affect the extracted shoreline. In this research, the original multispectral imagery is used to perform the classification result transformation from multispectral imagery to panchromatic imagery. This procedure involves in three major processes: (1) the image segmentation of the panchromatic imagery, (2) linking the segments with the classification results, and (3) stacking up the four classified segments.

The first process is the panchromatic image segmentation. The flow chart of the process is shown in Figure 3.20. In this process first the classification from multispectral imagery is adopted with the imagery re-sampled to the same resolution as the panchromatic image. Then the boundary of the water surface class is traced since this boundary line is the approximate shoreline. The boundary points on the edge of the image
are excluded because they do not belong to the approximate shoreline. A searching window with a predefined window size is created and the panchromatic image is cropped with this searching window then the Mean-Shift Segmentation is applied to this cropped panchromatic image. For every segment in the searching window, multispectral classification is overlapped onto the segment and the percentage of the overlapping area is calculated. A percentage threshold is predefined, which would be one of the parameters. And every pixel that belongs to this segment on the temporary occurrence map where the overlapping percentage is larger than the threshold is added by 1. This temporary occurrence map begins from a blank matrix, with zero for every index. After processing all of the segments in the searching window, the process continues with the next boundary point and its corresponding new searching window until every boundary point is processed and segmented. In the end, every index of the temporary occurrence map would be the number of times that pixel has been classified to a certain object class. There are definitely classification errors during this segmentation procedure. 10 is the tolerance we can be confident that the classification is robust. Hence all indices larger than ten are classified to the class on the panchromatic image.

The reason that the way this classification procedure is developed is to make sure that the available computer memory will not limit the size of the data. Despite the size of the image, this procedure can handle all sizes of data sets by adjusting the parameters. The reason that we focus on the computer memory issue is because the size and resolution of our satellite imagery. The total number of indices (pixels) of this panchromatic image is too large and the memory is not enough to perform the segmentation all at once. Base on the Mean-Shift theory, dividing the image into several
sections will be a conflict with the concept of "density-driven" since the density is created by the data themselves. If the image is divided, the density estimation on the edges of the divided image would be a problem. Under the current design, no matter how large the image is or how high its resolution is, computer memory will never be a problem, for it will take more computation time to exchange for memory usage while maintaining its robustness.
Classification of Multispectral Imagery

Re-sampling Classification into The Resolution of Panchromatic Image

Trace Water Surface Boundary

Panchromatic Satellite Imagery

Perform Mean-Shift Segmentation Within a Predefined Searching Window Along Boundary Points

For Each Segment Calculate Percentage of Overlapping Area

Percentage Larger Than Predefined Threshold

Yes

Add 1 To Every Pixel Belong To This Segment

No

Temporary Occurrence Map

Last Segment?

Yes

Temporary Occurrence Map > 10

Last Boundary Point?

Yes

Classification in Panchromatic Imagery

No

Figure 3.20 Flowchart of the panchromatic imagery classification

The image segmentation algorithm used in this process is still the mean-shift algorithm. However, the coding is done differently. In this study we used the EDISON as the core of the segmentation algorithm. There are two major parameters for the
mean-shift image segmentation: The spatial bandwidth, and color bandwidth. Beside the parameters from Mean-Shift algorithm, there are three other parameters from the procedure of our proposed process. Since the format of the WV2 imagery is the 16-bit TIFF file containing 11-bit data, the range of the image intensity is from 0 to 2048. However, the EDISON is designed to do the segmentation in a normalized LUV color space (0 to 1 for each band), and EDISON does not benefit from the 11-bit data depth when performing the segmentation. Hence, an equalization (stretching) of the 11-bit imagery would benefit the segmentation result. Several equalization methods had been tested, including "Histogram Equalize" and "Standard Deviation" of ArcGIS. The "Histogram Equalize" function is the classic equalization procedure of image processing. For each individual band, a histogram is created, minimum and maximum range of intensity is expended to 0 and 255. All intensity in-between is distributed proportionally (Szeliski, 2011). The concept of standard deviation stretching is to first calculate the mean and standard deviation (STD) of the intensity (grey value). The pixel values falling outside of the range of mean±2 STD will be set to either 0 or 255. Then the grey values are shifted and scaled in-between such that the mean value of the old distribution would be moved to 127 (Figure 3.21). The main difference between these two stretching methods is the upper and lower limits of the stretched image corresponding to different intensity values in the original image. The upper and lower limits in “Histogram Equalize” are the maximum and minimum intensity within the image, and those for “Standard Deviation” are mean+2 STD and mean-2 STD. The result shows that the "Standard Deviation" stretching is the best for the Mean-Shift segmentation algorithm. Hence, either "Standard Deviation" stretched image or the original image could be used for
The second parameter is the searching window size when cropping image patch from the panchromatic image. The larger the patch is, the more accurate the segmentation will be. However, it takes more computation time and memory to calculate. As a result, to find a number that can generate good result while maintaining reasonable computation time is the goal. Hence, a window size of 100 X 100 was adopted. The third parameter is the overlapping percentage threshold when determining if a segment belongs to this class.
The lower the percentage threshold, the more overflow area of the class there would be. However, some types of objects, such as the sediment banks, do need some overflow of the class in order to get a perfect result when extracting shoreline because according to the classification result in multispectral image not all the sediment banks were classified successfully. Some of the areas belong to sediment banks but do not showed up in the classification. But most of the classes are correct and do not need the overflow of segments, such as water surface and vegetation classes. All these classes require different parameter combinations due to their differences in properties. Although these parameters are determined manually, these parameters can remain the same if the same satellite data source is used.

Finally, after four classes had been classified in the panchromatic image separately, we need to combine them altogether to create one classification image. Since there is overlapping between the classes, how to stack them up while maintaining the correctness of the classification is the major concern. As we mentioned in the last paragraph, sediment banks may have some overflow problem by design, and usually are overflows to the water surface area. However, water surface classification result is relatively robust than sediment banks. Hence, sediment banks are placed at the bottom of the stack and water surface are stacked on top of the sediment banks. Meanwhile, shadow and vegetation usually accompany each other. Sometimes, shadow and vegetation areas overlap each other as well. We can assure that vegetation area is definitely a part of land but shadow area may be not. As a result, we stacked vegetation area on top of the shadow area. The water surface class may also overflow onto the shadow area due to the similarity of intensity on panchromatic image. However, we put in mind of this issue.
when classifying the shadow area. As a result, we need to make sure that the areas classified as shadow areas must be true shadow areas. Under this prerequisite, shadow areas may be stacked up on to the water surface class. The stacking sequence is shown in Figure 3.22.

![Figure 3.22 The stacking sequence generating the final classification from individual classifications.](image)

We did not develop a quantitative evaluation of the classification accuracy based on the following reason. Our concept of classification is to use proven indicators/algorithms to find the exact objects or materials that we are looking for, different from the supervised classification algorithms. Hence, the parameters or thresholds of these indicators/algorithms are empirical values with adjustments based on
manual inspection, not obtained from a certain manually identified training area. It is not practical to have a manually identified training area to evaluate the accuracy of classification in the production stage. The manual inspection is done by selecting one or multiple sensitive areas and making sure all image points within these areas are classified correctly. In other words, we have a semantic guideline instead of a quantitative one. This guideline will be described in section 5.3.

To determine these sensitive areas, a supervised classification is performed. After supervised classification, areas with incorrect classifications are considered as sensitive areas. Figure 3.10 shows two examples of sensitive areas.

### 3.3.5 Classification Adjustments and Refinements before Shoreline Extraction

The best case scenario of classification for shoreline extraction is the sediment bank/manmade objects or vegetation class immediately being next to the water surface class (left side of Figure 3.23). In this case, the shoreline could be clearly identified as the boundary line of the land. However, this is not always the case after classification. Some areas not classified as one of these four classes would be presented at the edge or next to the water surface. Unfortunately, we cannot create a new class for these areas because the insufficient of information, we have to stick with what we have and try to speculate what those areas are. To look on the bright side, we only need to determine if an area belongs to the land or the water surface. There is no doubt that sediment bank/manmade object and vegetation areas belong to the land. But shadow areas and unclassified areas are causing most of the issues. In this section, we need to identify all unclassified and shadow areas along the shore to determine what they are and whether they belong to the
water surface or the land. We have developed a rule-based procedure to determine what possible objects for unclassified and shadow areas may be. Figure 3.24 is the flowchart for the entire classification adjustment and shoreline extraction procedure.

Figure 3.23 In a sandy beach area, water surface and sediment bank classes are right next to each other. The water surface class is marked in blue, the sediment bank class in yellow, the shadow class in brown, and the vegetation in green. The uncolored area is the unclassified class. The areas further away from the water surface boundary will not include in the classification and marked as unclassified region.
Select an unclassified area from the classification

If northern side is shadow
Yes
Assign this unclassified area as land
No
Last unclassified area?
Yes
Select a shadow area

If southern side is sediment banks
Yes
Assign this shadow area as water surface
No
If surrounded by water surface
Yes
Split into two, half water surface half land
No
Last shadow area?
Yes
Select an unclassified area

If between water surface and sediment banks
Yes
Assign this unclassified area as water surface
No
If between water surface and vegetation/shadow area
Yes
Assign this unclassified area as land
No
Last unclassified area?
Yes
Trace the boundary of water surface

Water Surface

Land

Figure 3.24 Flowchart for the classification adjustment process.
The first process is to determine if an unclassified class belongs to the land. An unclassified area that is connected with water surface class is selected, and then the percentage of the immediate pixels on the opposite direction of sun light that belong to shadow areas is calculated. Light shine on an object on a higher elevation and its shadow is projected on a lower elevation surface. We are confident that if over 50% pixels of an unclassified area are shadow pixels on the opposite direction of sun light, that area would be on the land. Figure 3.25 shows an example of this situation.

Figure 3.25 The northern side of unclassified area is shadow class; this unclassified area will be adjusted as land. This is actually a gazebo on the pier.
The second process is to handle the shadow area. Shadow areas are the complicated objects of all classes. Since objects in shadow areas within our study area may be caused by manmade structures or bluff tops, two processes have been drawn. The first case is the shadows created by manmade structures (Figure 3.26), the other is those created by trees or bluff tops (Figure 3.27). For shadows created by manmade objects, direction of the sun light (sun located in south, sun light shines toward north in this satellite imagery) needs to be confirmed if there are sediment bank/manmade objects class next to this shadow area. If the percentage of the sediment bank/manmade object class is higher than 50%, then this shadow area belongs to the water surface. In the case of shadows created by trees or bluff tops, the possible cases are: shadows lying on the land, exactly on the shoreline, or on the water surface (Figure 3.28). If a shadow lies on the land, there should be a sediment bank/manmade object or unclassified class between the shadow and the water surface class. No adjustment is needed under this circumstance. If the shadow lies exactly on the shoreline or on the water surface, the shadow and the water surface class are right next to each other. Although we could detect the location where the shadow class and the water surface class met, there is no way to tell if the boundary is the shoreline. However, we could know for sure that the shoreline is within this shadow area. Under this circumstance, finding the exact location of shoreline is out of the question, instead, finding a line that can minimize the shoreline extraction error is the main goal. For long stretching forest along the shore or the bluff (Figure 3.27), the shadow created on the water surface is usually spike or cloud shaped areas depend on the shape of the trees that created it. The best choice of shoreline estimation would be the polyline that connecting the most inward water points within the shadow area. However,
for a single or small cluster of trees such as the examples shown in Figure 3.29, this could unveil the true classification of this shadow area by the surrounding classes. As a result, for each shadow area that is surrounded by water surface and/or other classes, a separation line could be drawn within the shadow area using the neighboring pixels classifications, to divide the shadow area into the water surface and land classes (Figure 3.30). These two shadow processes seem different in logic, but actually share the same procedure in programming. Both of them are dividing shadow area by creating polyline from the surrounding water pixels.

Figure 3.26 Shadows created by manmade objects such as vertical structures. (Image Credit: Bing Map)
Figure 3.27 Shadows created by trees and bluff tops. (Image Credit: Bing Map)

Figure 3.28 Possible conditions for a shadow-shoreline relation.
Figure 3.29 An example of a tree cluster. (Image Credit: Bing Map)

Figure 3.30 A shadow area split into the water surface and the land class based by the surrounding classes.
After completing the previous processes, the unclassified and shadow classes may be adjusted or reassigned. The final adjustment step would be based on the previous adjustments and re-evaluate the unclassified areas again. First, if unclassified areas are in-between water surface and sediment bank/manmade objects, they are assigned as water surface areas. Because these areas are more likely wet sediment banks on a flat beach caused by wave run-up. Figure 3.31 shows an example of this situation. This condition may also be caused by other situations as well, such as wetted manmade objects. However, there is no other information from the imagery itself for further distinguishing, and we have to accept this as a part of our type two error. However, LiDAR could provide elevation information to solve this issue. We will discuss this in the next chapter.

Secondly, if unclassified areas are in-between or contain sediment banks/manmade objects, vegetations and/or shadow areas, they are usually sloped structures. As a result, these areas are assigned as land. Figure 3.32 show examples of these unclassified areas.

After the entire process is completed, the classification is adjusted. The shoreline could be delineated by tracing the boundary of the water surface class to create the shoreline extracted from the WV2 satellite imagery.
Figure 3.31 An unclassified area in-between water surface class and sediment bank. This is usually the wet sand and should be classified as water surface.

Figure 3.32 Unclassified area in-between sediment banks, vegetations, and shadow areas.
3.4 Shoreline Comparison Method

To compare the extracted shoreline with the ground truth, three methods are used. The first method is comparing the distances between shorelines. The second method is comparing the area difference between land polygons created using the shoreline. The third method is the shoreline quality metrics developed by Ramirez (Ali, 2003).

3.4.1 Indicator of Shoreline Accuracy

Researchers in this field have developed several methods of line comparing. The most common way of calculating the distance between lines is as below (Figure 3.33):

1. Calculate the nodes on the extracted shoreline at predefined intervals.
2. For every node, find a point on the reference line that is closest to it.
3. Calculate the distance between the node and the point on the reference line.

Figure 3.33 The traditional method of calculating distances between shorelines
The accuracy obtained using this method represent the distance between shorelines very well, however, the closest distance may be correspondent to the peak point on the reference line and may not truly represent the maximum error between the shorelines. In this research, we developed another method of shoreline accuracy estimation. The key of this method is the concept of creating transecting lines perpendicular to the reference line at predefined intervals. The detailed procedure is described below. Figure 3.34 shows the procedure in the graphic form.

1. Select the manually digitized shoreline as the reference line;
2. Set an interval for calculating error distances (one-meter intervals were used in our experiment);
3. Create nodes on the reference line using the intervals defined in step 2;
4. On all the nodes draw transecting lines perpendicular to the reference line;
5. Search for the intersection point of the transecting line and the second shoreline; and
6. Calculate the distance between the intersection point and the node on the reference line.

There are fundamental differences between these two methods. The transecting line method averagely resulted in a longer distance, and the first method resulted in a shorter distance. This is because the first method is to find the closest point anywhere on the reference line, but the second method is to select the point exactly perpendicular to the reference line. On the other hand, where shoreline difference is large, a perpendicular
transecting line may not intersect with the reference line, resulting in a missing record of a huge value of distance. This may result in a bad shoreline with good accuracy in statistics. In this research, both methods were used. The first method was used in evaluating the entire shoreline, while the second method was used in evaluating local areas where we knew for sure there was no huge difference between shorelines.

1. Define reference line
2. Define interval
3. Create Node on Reference line
4. Draw perpendicular line
5. Find intersection point
6. Calculate distance

Figure 3.34 The method used when comparing extracted shoreline with the ground truth shoreline

3.4.2 Indicator of Area Difference on Land

The reason that the calculation of the land area was incorporated into this accuracy estimation is the blind spot the previous method had created. Under a normal circumstance, the shoreline closely follow the ground truth line, the statistics of distances
along the transecting line between the shorelines could represent the shoreline quality very well. However, in some regions, shorelines may be far apart, such as groins, breakwaters, and riprap revetments. In such regions, extracted shorelines could miss some of the area or missed out entirely. This inaccurateness did not occur with the previous statistics because the created transecting lines are not intersecting with the extracted shoreline.

The area of the land was calculated by predefining a rectangle boundary and then calculating the land area within this boundary. The indicator directly shows the area difference in square meters. An area larger than the ground truth is presented in a positive value, while that smaller than the ground truth is presented in a negative. Figure 3.35 shows how the process is done.
Figure 3.35 An extracted shoreline (in blue) compared with the ground truth shoreline (in red). Although the red area is not extracted as a shoreline, the accuracy is very good with the transecting shoreline comparison method. With the area indicator accompanying with distance indicator, the evaluation of shoreline quality is more reliable.

### 3.4.3 Shoreline Quality Metrics

This method consists of four indicators representing different aspects of the shoreline quality measures. These indicators are Generalization Factor (GF), Distortion Factor (DF), Bias Factor (BF), Fuzziness Factor (FF), and Overall Quality Factor (OVF). The Generalization Factor represents how much details are shown with a shoreline. A shoreline of higher resolution is with more details of the shore, and it would result in
longer length of a shoreline of lower resolution (Figure 3.36). The Distortion Factor represents the quality of length distribution of a shoreline (Figure 3.37). The Bias Factor shows if one of the shorelines is with a systematic shift. This factor is 1 when no bias exists between shorelines (Figure 3.38). Fuzziness Factor represents the accuracy of endpoints. Circles created on the endpoint using the maximum distance between the end points as radius, and circles with larger overlaps mean the end points are more accurate (Figure 3.39) (Ali, 2003). The FF factor did not participate in our shoreline accuracy estimations since we could not recreate the calculated value using the equation provided by Ali (2003).

Figure 3.36 The concept of Generalization Factor (Ali, 2003)
Figure 3.37 The concept of Distortion Factor (DF) (Ali, 2003)

Figure 3.38 The concept of the Bias Factor (BF) (Ali, 2003)
Figure 3.39 The concept of the Fuzziness Factor (FF) (Ali, 2003)
Chapter 4. Shoreline Integration Utilizing Scenario Analysis

The shorelines extracted from LiDAR data and satellite imagery have their own advantages and disadvantages. Their accuracies are different with different types of terrain, and different sensor types may result in fundamental differences of extracted shorelines. Furthermore, they represent shorelines in different time periods. With all these differences and advantages in mind, we came up with a procedure to integrate these two shorelines while making up for each other’s weakness and evaluate the shoreline changes all at once. Before explaining the entire process of the shoreline integration, the concept of this process needs to be described. This integration procedure is based on the scenario of the Painesville coastline, in other locations with similar terrain type can utilize this procedure directly. However, with other types of coastal terrain, this procedure may need to be modified and/or extended.

Figure 4.1 shows the flow chart of the shoreline integration procedure. The fundamental process in this shoreline integration procedure is the erosion/accession analysis. It is assumed that the shoreline extracted from LiDAR point cloud is obtained using an older data set, and the WV2 satellite imagery is extracted more recently. In this case, for shoreline areas with erosion, the shoreline extracted from WV2 is more inland than the shoreline extracted from LiDAR data. On the contrary, for shoreline areas in accession, the shoreline extracted from LiDAR data is more inland than the shoreline
extracted from WV2. Incorporating these definitions into the process, we can divide the entire shoreline into three types based on conditions, which are shoreline with erosion, shoreline in accession, and shoreline without changes. Since these two shorelines were acquired and extracted using completely different techniques and may result in different accuracies, expecting a true evaluation of erosion/accession is not quite practical. A range of threshold is defined based on the shoreline accuracy obtained with the LiDAR data. The three following sections will illustrate the procedures of handling three different statuses of shoreline changes: shoreline with erosion, shoreline in accession, and shoreline without significant changes.
4.1 Shoreline Integration for Manmade Structure Areas

This section of the chapter focuses on the areas that do not have any significant erosion or accession situation in progress. These areas are most likely manmade structures especially in Painesville. Among the seven types of terrain that had been identified in Painesville, five may remain constant. These types are vertical structures, sloped structures, groins, breakwaters, and pier structures. Needless to say, all of them are...
manmade structures. Piers, however, does not have the same property as other terrain types due to the presence of boats or yachts. Fishing or recreational boats are available in our study area and they are usually small in size and relatively low in height. As a result, they may not be identified in LiDAR shorelines and it may look like there is nothing there along the pier. However, on satellite images, any boats larger than half of a pixel size will be shown on the image. Due to the materials of boats, it would be highly possible to identify them as sediment banks/manmade structures, or unclassified areas other than water surfaces. Handling pier structures will be discuss in the later section and are thus excluded from this process.

Among these four types of terrain, using WV2 we could extract a better shoreline with vertical structures but not as robust as LiDAR shorelines with other types of terrain. Groins and breakwaters are basically sloped structure in shape, with only differences in location and function. Hence, sloped structure may represent all three types of terrain. In this process, LiDAR shoreline will be used to be combined with the WV2 shoreline within areas with these three types of terrain by averaging the shorelines. However, LiDAR and WV2 data sets are acquired during different time periods, and combining two shorelines with different water levels doesn't make any sense at all. Hence, water level differences need to be investigated. This procedure can be further divided into three process: 1. Determining water level for LiDAR data set and satellite imagery, 2. determining sloped structure areas using the LiDAR data set, and 3. combining LiDAR and satellite extracted shoreline within the sloped structure area (Figure 4.2).

In order to retrieve the water level of the data sets, the exact date and time when the data set was acquired is needed. Then, the date and time is used looking up the nearest
gauge station water level records to obtain the water level when the data sets are acquired. We were targeting on sloped structures, usually with 35~45-degree slopes to water. As we know, half of a pixel size is the smallest distance we can detect from satellite imagery and half a pixel size in WV2 imagery is 25cm. In a 45-degree slope, water level drops 25cm vertically, shoreline extends 25cm horizontally, and vice versa. As a result, we used ±25cm as the threshold of the water level differences between LiDAR and WV2 data sets. If a water level difference between data sets is smaller than the threshold, water levels are similar and we treat them as if they are the same. Under this circumstance, shorelines can be combined directly by averaging them and no other processes are needed. If LiDAR data set water level is lower than the WV2 imagery water level, a new shoreline must be generated. First, DEM is generated using the LiDAR point cloud, then a contour line of the WV2 water level height is created and considered as a shoreline. All these processes can be done with ArcGIS. After this new shoreline is created, the shoreline within the sloped structure areas can then be calculated by combining the WV2 shoreline and this new created shoreline. On the other hand, if the LiDAR data set water level is higher than the WV2 imagery one, there is no way to obtain the WV2 imagery water level shoreline from the LiDAR data set. Hence, no process is needed, but the errors corresponding to the WV2 shoreline will remain with the final integrated shoreline.
Figure 4.2 Flow chart for performing the shoreline integration within sloped structure areas.

Detecting where sloped structures are located is another critical process in this procedure. Sloped structure areas can be detected using the following procedure (Figure
4.3). First a TIN is created using the Delaunay triangle algorithm based on LiDAR point cloud. Then by searching along the LiDAR data extracted shoreline it can be determined if the two points of a shoreline segment belong to a same TIN triangle and if the third point is on the land. For segments with two points belonging to the same triangle, the highest and lowest points within this triangle are obtained and the slope along this high/low triangle edge can be calculated. For points on the shoreline connected with only one point of a TIN triangle, one of the triangles that is connected with this point is found while two other points are on the land side with the highest elevation, then the slope along this triangle edge is calculated. If the slopes calculated with both of the methods are between 25 and 50 degrees, then this segment or point is on top of a sloped structure. The slope calculated with a triangle with only one point connected to the shoreline may not lead to robust results compared with the one having an edge of a triangle being a part of the shoreline. This may cause noises when calculating slopes. Since slopes are usually the same within a neighboring area, median filters may be applied to the calculated slope data set to remove the extreme values within the neighboring shoreline points.
4.2 Shoreline Integration for Areas with Erosion

There are two possible conditions for an area with erosion, the shoreline is truly with erosion or the shoreline delineation is incorrect. Incorrect delineation of shoreline can only happen in one situation, which is when WV2 classifies a land area as an
unclassified area, and this mistake is not corrected by the shoreline adjustment procedure described in section 3.3.5. In order to determine which case it is for every area with erosion, WV2 classification results must be incorporated with the shorelines. First, the eroded area (the area between LiDAR and WV2 shorelines) intersect with the area in WV2 classified as water surface. If most of the pixels (80% is used as thresholds) within the eroded area belong to the water surface class, then this eroded area is truly an area with shoreline erosion. If this eroded area is not overlapping with the water surface class, there can be only one condition, which is that these areas would be treated as water surface during the WV2 shoreline adjustment: these areas belong to an unclassified area. Hence, unclassified area within eroded area reassigned to the land class, and then the WV2 shoreline is regenerated.

This incorrect shoreline delineation may happen where a land is stretching into a water body, such as groins or breakwaters. These structures are usually wet due to waves, and making it impossible to classify them as sediment banks/manmade structures which is the class it should be. This may create serious error in shoreline extraction when using satellite imagery. For example, if two lands connected by a breakwater are not classified as sediment banks/manmade structures, one side of the shoreline may not be extracted. Because the shoreline tracing process is from one end of the image to the other end, without the area in-between the land, the other side will never be reached.

4.3 Shoreline Integration for Areas with Accession

There are only two possibilities if a coastal area is truly with accession: a sediment bank is in the process of sedimentation, or new manmade objects are built on
the shore and change the shoreline. However, for both of the cases, shorelines can be extracted by WV2 perfectly, and no further adjustment is needed. Other situations may create illusions of shoreline accession which in fact are not. These cases are piers with boats or yachts and shadows are projected onto water surface and the boundary of water and land cannot be determined on imagery. In order to solve these two types of shoreline issues, two processes are designed. The first process is to extract manmade piers and the second process is to replace the WV2 shoreline within shadow areas with the shoreline extracted from LiDAR data.

The procedure of extracting shoreline on piers is depicted in Figure 4.4. The shoreline segments that are been identified as areas with accession is imported to the procedure. First, we calculated the length of this segment. Since we like the result to be robust, the longer the segment is, it is more robust that the area truly belongs to a pier. Hence, we defined the threshold of the length of single pier to be at least 15m. Then, we created a bounding box of this shoreline segment and extended this box with another threshold, 10m in this case, as a searching window. Within this searching window, we found a straight line larger than 15m then extended the new searching window with the new bounding box containing this new straight line. After the new searching window was created, we calculated the equation of this new line and the line accuracy compared with the LiDAR shoreline. If the accuracy was less than 3m in RMSE (the accuracy of LiDAR extracted shoreline), this line was added to the shoreline candidate pool. The lines larger than 3m in accuracy were discarded. Then, we used the line equation we just calculated, and found parallel lines within this new searching window. If parallel lines were found, a new searching window was created and the process of searching new lines needed to be
done repeatedly. If there was no new parallel line found, then we had to try to search perpendicular lines, and then other possible lines. In the end, all straight lines possibly parts of the shoreline were identified. Candidate lines with end point close to another line's end point were found and then these lines were connected. If there were several end points also close to each other, multiple combinations of line linking sequences could be established. Then an accuracy estimation process can be applied to all of these linked lines, the line with the highest accuracy would be the shoreline representing the pier. The line sequence selection in this process may require human inspection in order to maintain the best shoreline representation.

The other possibility for an area seeming to have accession but it is actually not is the area where water surface and shadow are overlapped. In section 3.3.5, we have discussed this matter but there is no perfect solution with only satellite imagery. With the help from LiDAR data, this problem may be solved to a certain degree. LiDAR shorelines are not affected by shadows. Hence, LiDAR shorelines can be used to correct water/land boundaries in shadow areas of WV2 satellite imagery. However, not every water/land boundary in a shadow area can be corrected due to changes of shorelines through time. For manmade areas, whether there are shadows or not does not matter. Shorelines can be extracted perfectly with satellite imagery, hence no correction is needed. For areas truly with erosion, since LiDAR represents a shoreline in the past, correcting the WV2 shoreline with LiDAR shoreline is not a reasonable method. For areas other than sediment banks or manmade structures with shoreline accession, in a heavily eroded bluff region like Painesville, it represent that this accession is a deception. Since WV2 shoreline did not recess, representing this shoreline remains unchanged. Thus,
LiDAR shoreline can be used to replace WV2 shoreline within these shoreline segments.

Figure 4.4 Flow chart of the shoreline refinement procedure at piers.
Chapter 5. Experiment and Accuracy Analyses of the Extracted Shorelines

5.1 Descriptions of Study Area, Data sources, and Data sets

Painesville Township, Ohio is the chosen study area for this research. It is located 25 miles east of Cleveland, along the south shore of Lake Erie, in Lake County. The data sets we used contain the data of the shoreline of the entire township including Fairport Harbor. The total length of the shoreline is around 14 km (Figure 5.1). Since 1988, the Ohio Department of Natural Resources (ODNR) has begun to monitor the erosion along Lake Erie and designate the Coastal Erosion Area (CEA) to indicate the areas that are anticipated to be lost. According to the ODNR CEA map updated in 2010, 51% of the shoreline within our study area was designated as CEA (information derived from ODNR, 2010). The maximum recess rate within the study area was 10.5 ft/year (3.2 m/year) between 1990 and 2004 (ODNR, 2010) (Figure 5.2). CEA was measured based on bluffline erosion, and shoreline was also recessed along with the bluffline as well. However, the recession rate and shape of the shoreline may be different from those of the bluffline, since eroded soil from the bluff usually creates a sediment bank between bluff toe and water surface, and the shoreline erosion is measured along the sediment bank.
Figure 5.1 The location and the orthophoto of Painesville Township.

Figure 5.2 The CEA frame with the largest recess rate within study area (ODNR, 2010)
There are some features of shoreline that were commonly present along coastal region within our study area. From the west, sandy beach of Headlands Beach State Park stretches for around 1km (Figure 5.3 a). The next feature was a light house (Fairport Harbor Light House) built on breakwater (Figure 5.3 b). Then, there were manmade vertical structures along the Grand River (Figure 5.3 c). There were also recreational docks (Figure 5.3 d) and sloped structures built with ripraps (Figure 5.3 e). And there were also features including eroded bluffs mixed with sandy beach, building materials, and vegetation (Figure 5.3 f). It is more complicated and difficult to classify the terrain type under this situation because of the inhomogeneity of the materials.
Figure 5.3 Shoreline features along the Painesville coastline (Image Credit: Bing Map). (a) Sandy beach at Headland Beach State Park, (b) Fairport Harbor Light House, (c) Manmade vertical structures along the shore and the Grand River, (d) Recreational Dock along the Lake Erie shore, (e) Riprap revetment sloped structures protecting the shoreline, and (f) Bluff erosion area, sandy beach, and building materials are mixed within this area.
Figure 5.3 continued

(c)

(d) Continued
The datasets we used in this study included the public domain LiDAR data provided by Ohio Geographically Referenced Information Program (OGRIP) and the Worldview-2 satellite images acquired by DigitalGlobe.
OGRIP is a program under the Ohio office of Information Technology (OIT). The Ohio Statewide Imagery Program (OSIP) organizes the aerial photogrammetry, topographic LiDAR, and Digital Elevation Models (DEMs) for the entire State. OSIP datasets are totally free for public access and OSIP is a continuing project for periodically acquiring data for the entire State of Ohio. The dataset of the first phase of OSIP is acquired during 2006-2008 and released to the public completely in 2009. In 2011, the second phase of the OSIP began and new acquired dataset have not been released yet. In the OSIP phase 1 dataset, North American Datum 83 (NAD83) High Accuracy Reference Network (HARN) was used as the horizontal datum. North American Vertical Datum of 1988 is used as vertical datum. The datasets are provided to users in the form of county by county or 5000-foot grids in TIFF (image) and LAS (LiDAR point cloud) formats.

The dataset for our study area, Painesville, contains Aerial orthophotos, Aerial LiDAR, and DEM generated by aerial LiDAR systems. In this research, we only used the LiDAR data from this dataset; the orthophoto was used as a reference and not involved in the shoreline extraction procedure. The resolution of the LiDAR point cloud is about 7 foot (2.1m) point spacing and the accuracy is about 1 foot (0.3 m) in both horizontal and vertical directions (Figure 5.4). The LiDAR data were acquired by the Woolpert Inc. and the Horizons Inc. using Leica ALS50 with the flying altitude of 7300 ft (2225 m) at 170 knots on April 11th, 2006 (16:43~21:01 EDT). The information of the provider, datum, resolution, accuracy, and acquisition date is summarized in Table 5.1. The aerial photographs were also acquired by the Woolpert using Leica ADS40 with the flying altitude of 4800 ft (1463 m).
Figure 5.4 The distribution of the LiDAR point cloud. The yellow areas are covered by LiDAR points.

The WorldView-2 satellite image used for this research contains an area which is 8 km long and 4.5 km wide. UTM 17N was used as the datum for this satellite image when we received it. There were two images in the package, the 0.5m resolution panchromatic image (Figure 5.5) and the 2.0m resolution multispectral image with 8 bands (Figure 5.6). This dataset was acquired on Sep. 14th, 2010 (16:45 UTC). The information regarding the provider, datum, resolution, accuracy, and acquisition date is summarized in Table 5.1.
Figure 5.5 Panchromatic band in the World View 2 imagery in Painesville

Figure 5.6 Multi-spectral band (R, G, B) in the World View 2 imagery in Painesville
In order to evaluate the accuracy of the extracted shoreline, the ground truth shoreline needed to be determined. The ground truth shoreline was delineated manually following the guidelines described in section 2.3 based on the panchromatic WorldView-2 satellite imagery. However, the aerial orthophotos was used as a secondary data source where areas were difficult to determine the shoreline in the grey scale satellite imagery. The total length was around 14 km and the terrain types were also determined while delineating the ground truth. This ground truth shoreline would be used during all the phases, and all the extracted shoreline would be compared against this ground truth shoreline. Table 5.1 lists all the datasets used in the experiments.

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Provider</th>
<th>Datum</th>
<th>Resolution</th>
<th>Accuracy</th>
<th>Acquisition Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>World View 2 Panchromatic</td>
<td>DigitalGlobe</td>
<td>UTM 17N</td>
<td>0.5 m</td>
<td>6.5 m</td>
<td>2010.9.14</td>
</tr>
<tr>
<td>World View 2 Multi-Spectral</td>
<td>DigitalGlobe</td>
<td>UTM 17N</td>
<td>2.0 m</td>
<td>6.5 m</td>
<td>2010.9.14</td>
</tr>
<tr>
<td>LiDAR</td>
<td>OGRIP</td>
<td>NAD 83 HARN OHIO NORTH; NAVD 88</td>
<td>7 ft. point spacing (2.1 m)</td>
<td>Horizontal About 1 ft. (0.3 m) Vertical 1 ft. (0.3m)</td>
<td>2006.4.11</td>
</tr>
<tr>
<td>Orthophoto</td>
<td>OGRIP</td>
<td>NAD 83 HARN OHIO NORTH</td>
<td>1 ft. (0.3 m)</td>
<td>5 ft. (1.5 m)</td>
<td>2006.3.18-2006.5.7</td>
</tr>
</tbody>
</table>

5.2 Strategy and Procedure to Perform Shoreline Extraction

The proposed shoreline extraction approach consists of three phases. In Phase One, a shoreline is extracted by classifying LiDAR points into water surface points and
land points then tracing the boundary of land points. In Phase Two, a shoreline is extracted by classifying satellite imagery into four classes and then tracing the boundary of water surface class. In Phase Three, the two shorelines extracted in the two previous phases are integrated based on their physical topography and properties. To perform the shoreline extraction procedure, we should follow the sequence of Phase One, Phase Two, and Three. All approaches that we have evaluated in this chapter are provided with accuracy estimations. These accuracy estimations are the RMSEs of the shoreline accuracy, and this method is the first method described in section 3.4.1.

In Phase One, LiDAR data and satellite imagery are used. We could follow the procedure described in section 3.2 to perform the process of Phase One. There are two parameters needed to be determined in Phase One, which are the bandwidths used for mean-shift segmentation. One of these bandwidths is for elevation segmentation, and the other is for NIR and panchromatic band segmentation. As discussed in section 3.2.3, the bandwidth used for elevation segmentation is rather easy to be determined. However, the bandwidth for NIR and panchromatic band segmentation is still easy to find, but some testing is needed. Empirical values could first be used, if the performance is not as expected, other values should be applied. The procedure of parameter selection is described in section 3.2.3. Accuracy estimation is performed with the extracted shoreline and the results are described in section 5.3. From the accuracy estimation, we have discovered an issue within the classification decision tree. A solution is proposed and described in section 5.3 as well.

In Phase Two, only the satellite images are used to extract the shoreline. The shoreline extraction procedure is described in section 3.3. There are a total of five
parameters within this procedure. Unlike the bandwidths in Phase One, the parameters could not be determined easily. In this phase, three of the parameters (CB, SB, SOP) are in a case-by-case situation, and different parameters could be applied to each class. In section 5.4.1, we are going to demonstrate how the parameters are determined one class at a time. In section 5.4.2, the accuracy of the extracted shoreline is estimated.

In Phase Three, the two shorelines extracted in the two previous phases are integrated using rule-based classification based on scenario analysis. The integration process is described in Chapter 4. There are no parameters within this procedure, however, when if extracting shorelines outside our study area, other scenarios may need to be included and adjusted for the classification. The accuracy of the integrated shoreline is estimated in section 5.5.

In section 3.4, we have already discussed several methods to evaluate the quality of shorelines. After comparing the shorelines extracted using these approaches, some suggestions and conclusions related to these shoreline quality indicators are included in section 5.6.

Every algorithm has its advantages and disadvantages, and some of which are related to the data sources we used and we would know that from the beginning of the research. Some of them are observed from the results after the algorithm is developed. The issues we have found after the development stage are discussed in section 5.7. For most of the issues discussed, possible solutions are provided. The detailed analyses of these solutions will be left for future researchers.
5.3 Accuracy Estimation of the Extracted Shoreline in Phase One

The shoreline extracted in Phase One was compared with the ground truth shoreline, and the result showed that the largest error was 26 meters, the average error was 2.56 meters, and the RMSE was 3.26 meters (Table 5.2). After conducting a further analysis of the cause of errors, it was found that there were shoreline changes between the acquisition times of the two datasets (the LiDAR data was acquired on April 11th, 2006, and satellite imagery was acquired on Sep. 14th, 2010). Figure 5.7~Figure 5.9 show examples of shoreline accession in a sandy beach area, the accession was caused by sediment transport. Figure 5.10~Figure 5.11 show examples of shoreline recession in a bluff area, the recession was caused by bluffline erosion. The reason that sometimes the extracted shoreline is representing the shoreline instance of 2006, and sometime 2010 is the design of the decision tree for classifying LiDAR points. Figure 5.12 shows the decision tree described in section 3.2. Where sedimentation process is in progress, if water surface points are acquired by the LiDAR system, this water surface will be classified as water surface in the process of “Mean-shift segmentation of elevation” in Figure 3.2 since their elevations are similar as that of the water surface. The extracted shoreline is the shoreline in LiDAR dataset instead of the satellite imagery. However, if these water surfaces are not showing in the LiDAR data (due to reflection of water surface described in section 3.2.3), the extracted shoreline may represent the shoreline depicted in the satellite imagery. Areas where the shoreline has been eroded away, such as the condition described in Figure 5.10, are further analyzed. These LiDAR points are within the areas used to be the sediment bank but now belonging to water surfaces. Thus,
the attributes of these LiDAR points show the properties of water surfaces. There is no process in the decision tree to classify points into water class using the point attribute. This is one of the drawbacks existing in the approach we proposed, and the following modifications in the decision tree can be done to resolve this problem: including the NDVI indicator to classify the water surface class. As shown in Figure 5.13, NDVI is added as a feature to the mean-shift segmentation along with the elevation. As a result, the largest segment and the segment with the lowest elevation will only include the LiDAR points that are on the water surface in both LiDAR data and satellite imagery. With this process, the problem of accessed land being classified as water surface can be solved. Next, a decision process is added in-between the mean-shift segmentation processes to determine if the NDVI of the unclassified LiDAR points are similar to the NDVI values of the LiDAR points already identified as water surface. The similarity matching could be done either using maximum and minimum boundaries or a normal distribution calculated from the NDVI values of the points on water. This process is designed to classify the LiDAR points on land and now on water surface.
Table 5.2 Shoreline accuracy before manual inspection

<table>
<thead>
<tr>
<th>Shoreline Type</th>
<th>RMSE (m)</th>
<th>Max Error (m)</th>
<th>Average Error (m)</th>
<th>Length (m)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediment Bank</td>
<td>5.037</td>
<td>26.018</td>
<td>3.410</td>
<td>3183</td>
<td>24</td>
</tr>
<tr>
<td>Sloped Structure</td>
<td>2.352</td>
<td>23.211</td>
<td>2.031</td>
<td>5174</td>
<td>39</td>
</tr>
<tr>
<td>Vertical Structure</td>
<td>1.449</td>
<td>15.423</td>
<td>1.747</td>
<td>2651</td>
<td>20</td>
</tr>
<tr>
<td>Bluff</td>
<td>2.834</td>
<td>12.986</td>
<td>3.471</td>
<td>1732</td>
<td>13</td>
</tr>
<tr>
<td>Groin</td>
<td>0.868</td>
<td>3.522</td>
<td>1.305</td>
<td>68</td>
<td>1</td>
</tr>
<tr>
<td>Breakwater</td>
<td>4.789</td>
<td>17.724</td>
<td>3.784</td>
<td>111</td>
<td>1</td>
</tr>
<tr>
<td>Pier</td>
<td>0.673</td>
<td>2.970</td>
<td>1.048</td>
<td>240</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>3.258</td>
<td>26.018</td>
<td>2.564</td>
<td>13159</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 5.7 A large error on the sandy beach caused by the shoreline sedimentation. In (a) is the aerial photo acquired in 2006, (b) base map is the WV2 panchromatic image acquired in 2010. The black line is the same line in both images, which is the shoreline extracted in Phase One. There is a sedimentation process occurring during these four years and the shoreline has been accessed.
Figure 5.8 The sedimentation process around the center of the image. In (a) is the aerial photo acquired in 2006, (b) base map is the WV2 panchromatic image acquired in 2010. There are shoreline extraction errors near the left side of the image. The extracted shoreline is not consistent with either of the datasets.
Figure 5.9 The sediment bank is accumulated near the center of the image showing on (b). This sediment bank is not successfully shown in the extracted shoreline. In (a) is the aerial photo acquired in 2006, (b) base map is the WV2 panchromatic image acquired in 2010.
Figure 5.10 The bluff and the sediment bank were eroded away during the time frame of LiDAR and the satellite imagery acquisition. In (a) is the aerial photo acquired in 2006, (b) base map is the WV2 panchromatic image acquired in 2010. The extracted shoreline represents the topography in LiDAR data.
Figure 5.11 A clear evidence of bluffline erosion. In (a) is the aerial photo acquired in 2006, (b) base map is the WV2 panchromatic image acquired in 2010. The shoreline is eroded along with the bluffline. The extracted shoreline represents the LiDAR data.
Figure 5.12 Decision tree of classifying LiDAR points with multispectral information as attribute.
After the manual inspection, it was found that 8% (1 km) of the entire shoreline was in significant erosion or accession. Since these errors were not originated by the procedure, excluding these errors could lead to a better evaluation of the extracted shoreline accuracy. The excluded areas are shown in Figure 5.7, Figure 5.9, and Figure 5.10. After excluding these shoreline errors within the erosion/accession area, the
shoreline accuracy was improved. The average error was 2.07 meters and the RMSE was 2.152 meters (Table 5.3). We divided the shoreline into multiple sections with different terrain types and estimated their accuracies independently. The performance of the extracted shoreline was the best in vertical structure and pier areas (1.45m and 0.67m RMSE, respectively) (Figure 5.15) and very good in sloped manmade structures (Figure 5.14). In pier and vertical manmade structure areas, the extracted shoreline result was robust in finding the shoreline in the vicinity but the extracted shoreline creates a zigzagged line instead of a straight line due to the resolution (point spacing) of the LiDAR point cloud.

After excluding the error points caused by the shoreline changes, the resultant accuracy would be the true accuracy of our shoreline extraction procedure. The result showed high accuracy in sediment bank and sloped structure, and lowest accuracy on groin and breakwater. The extracted shoreline in sediment banks, sloped structures, and vertical structures was robust and accurate. The achieving of segmenting wet/dry line would be primarily contributed by the panchromatic and NIR intensity of satellite imagery. The areas where the shoreline is in-between bluff toe and the water surface, these areas are classified as bluff. In these areas, a small strip of sediment bank can usually be found. Although these areas are made of the same materials, the shoreline accuracy shows totally different characteristics. We have mentioned that in Figure 5.3 (f) that there are different types of obstacles on top of the sediment bank, making classification complicated. In addition, while excluding the areas with shoreline changes, some areas with minor bluffline erosion are not completely removed, such as the area shown in Figure 5.11. Since the area where the shoreline has changed resides in the
extracted shoreline, the shoreline accuracy estimation within the bluff areas may be biased.

Table 5.3 Shoreline accuracy after excluding shoreline changes between datasets

<table>
<thead>
<tr>
<th>Shoreline Type</th>
<th>RMSE (m)</th>
<th>Max Error (m)</th>
<th>Average Error (m)</th>
<th>Length (m)</th>
<th>% Original Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediment Bank</td>
<td>2.005</td>
<td>18.290</td>
<td>1.774</td>
<td>2253</td>
<td>71</td>
</tr>
<tr>
<td>Sloped Structure</td>
<td>1.963</td>
<td>16.347</td>
<td>1.868</td>
<td>5061</td>
<td>98</td>
</tr>
<tr>
<td>Vertical Structure</td>
<td>1.449</td>
<td>15.423</td>
<td>1.747</td>
<td>2651</td>
<td>100</td>
</tr>
<tr>
<td>Bluff</td>
<td>2.757</td>
<td>12.986</td>
<td>3.335</td>
<td>1669</td>
<td>96</td>
</tr>
<tr>
<td>Groin</td>
<td>0.868</td>
<td>3.522</td>
<td>1.305</td>
<td>68</td>
<td>100</td>
</tr>
<tr>
<td>Breakwater</td>
<td>4.789</td>
<td>17.724</td>
<td>3.784</td>
<td>111</td>
<td>100</td>
</tr>
<tr>
<td>Pier</td>
<td>0.673</td>
<td>2.970</td>
<td>1.048</td>
<td>240</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>2.152</td>
<td>18.290</td>
<td>2.073</td>
<td>12074</td>
<td>92</td>
</tr>
</tbody>
</table>

Figure 5.14 The shoreline extraction with good performance in the sloped manmade structure.
5.4 Parameter Determination and Accuracy Estimation of the Shoreline Extracted in Phase Two

5.4.1 Parameter Determination

The accuracy of the extracted shoreline of our proposed procedure based on the satellite imagery was affected by the selection of parameters. There were five parameters needed to be defined to classify each class. These parameters are spatial bandwidth (SB) and color bandwidth (CB) from the original mean-shift image segmentation algorithm, and panchromatic satellite image stretching method (IS), segmentation window size (SWS), and segment overlapping percentage (SOP) from the procedure that we developed. These parameters are described in detail in Section 3.3. In order to find the best set of parameters to extract the most accurate shorelines we needed to evaluate the
impact of each and every parameter. Then, a different parameter set was selected for each class in order to satisfy different properties of the terrain types. After the shoreline was extracted from the combination of the classes, then this shoreline accuracy was evaluated and compared with other shorelines created using other parameter sets. Since we are developing shoreline delineation algorithms and evaluating extracted shoreline accuracy, we manually delineate the ground truth shoreline for comparison. Users in the future would not have the ground truth shoreline for accuracy comparison. In addition, they may work on other locations and/or use satellite images other than WV2. Under these circumstances, it is impossible for us to provide a set of parameters to accommodate all types of satellite sensors and locations. A proper solution is to provide a procedure of finding the best set of parameters for future users. However, without a ground truth shoreline, the accuracy of the shorelines extracted using different sets of parameters may not be estimated precisely. Instead of finding a parameter set creating the most accurate shoreline, we found the parameter sets leading to the robust and reliable shoreline accuracy which is more substantial. Hence the average and standard deviation of the RMSE accuracy for every parameter combination is calculated. For example, while a value is used for parameter A and a range of values are used for parameter B, not only the accuracy created by each parameter set is estimated, the average and standard deviation are calculated representing the performance of this value of parameter A. The higher the average accuracy value, the more accurate shoreline this parameter value would lead to. The better the standard deviation value, the more stable and dependable this parameter setting.

The first parameter here to evaluate was the method of panchromatic satellite
imagery stretching (IS). This process was actually a pre-processing step for the panchromatic image. The original satellite imagery was created directly by quantifying the photons received by the sensor. Adding an equalization process to the pre-process of the original image could distinct the separation of the classes. This parameter defined the use of the original image or the image stretched through the standard deviation stretching procedure. Since the water surface class contributes most to extracting the shoreline, the water surface class was used to evaluate the impact of the parameters. For other parameters which did not play an important role in this parameter evaluation, default values were used. Table 5.4 shows the RMSE of the accuracy estimation (first method described in section 3.4.1) using the water surface boundary mimicking the shoreline with panchromatic image preprocessed using the standard deviation stretching. For parameter CB and SOP, we evaluated most of the possible values. CB was ranged from 0.5 to 6.5, and SOP was ranged from 0.1 to 0.9. For other parameters such as SB and SWS, the default values 7 and 100 by 100 were used, respectively. The reason that the combination of parameters IS, CB, and SOP are first compared, and a single value is used for SB and SWS is because IS and CB are interrelated. The IS parameter may influence the color intensity, and CB defines the intensity bandwidth of mean-shift segmentation. SOP on the other hand, is a classifier; with no default value. Hence the parameter SOP exists in every comparison in section 5.4. SB would influence the estimation of density function in the mean-shift algorithm, resulting in a different number of classes, and SWS would influence the continuity of a segment. Hence, these two parameters are independent and could be evaluated separately. After a further inspection of Table 5.4 (a), we found that the SOP value above 0.7 did not lead to an accurate result or even the
shoreline was not traceable from the beginning to the end (represented by -1). On the other hand, the SOP values of 0.1 and 0.2 led to the best accuracy. Meanwhile, the performance was better with CB parameter between 0.5–5.5. Table 5.4 (b) shows the shoreline accuracies using the original panchromatic imagery without any preprocessing. Since SOP values larger than 0.7 did not lead to better results base on the previous evaluation, they were not included in the table. In general, the performance of the panchromatic imagery preprocessed using the standard deviation stretching method was better in accuracy and robustness, hence standard deviation stretched panchromatic image would be used in the latter evaluation.
Table 5.4 The RMSE of shoreline accuracy with respect to the CB and SOP parameters. (a) Standard deviation stretched panchromatic imagery, and (b) Original panchromatic imagery.

(a) Standard Deviation Image Stretching

<table>
<thead>
<tr>
<th>Segment Overlapping Percentage (SOP)</th>
<th>Color Bandwidth (CB)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AVG</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.5</td>
<td>5.5</td>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>5.655</td>
<td>4.348</td>
<td>3.161</td>
<td>2.445</td>
<td>2.179</td>
<td>2.184</td>
<td>2.249</td>
<td>3.175</td>
</tr>
<tr>
<td>0.2</td>
<td>4.988</td>
<td>3.680</td>
<td>2.961</td>
<td>2.438</td>
<td>2.461</td>
<td>2.730</td>
<td>2.374</td>
<td>3.090</td>
</tr>
<tr>
<td>0.3</td>
<td>4.486</td>
<td>3.282</td>
<td>2.718</td>
<td>2.620</td>
<td>2.766</td>
<td>2.803</td>
<td>4.784</td>
<td>3.351</td>
</tr>
<tr>
<td>0.4</td>
<td>4.265</td>
<td>3.333</td>
<td>3.220</td>
<td>4.987</td>
<td>5.009</td>
<td>3.034</td>
<td>4.860</td>
<td>4.101</td>
</tr>
<tr>
<td>0.5</td>
<td>4.582</td>
<td>3.343</td>
<td>3.261</td>
<td>5.029</td>
<td>5.038</td>
<td>5.060</td>
<td>4.881</td>
<td>4.457</td>
</tr>
<tr>
<td>0.6</td>
<td>5.001</td>
<td>5.016</td>
<td>4.969</td>
<td>5.035</td>
<td>5.137</td>
<td>5.066</td>
<td>4.956</td>
<td>5.026</td>
</tr>
<tr>
<td>0.7</td>
<td>4.638</td>
<td>-1</td>
<td>4.528</td>
<td>4.701</td>
<td>4.971</td>
<td>5.116</td>
<td>4.969</td>
<td>4.820</td>
</tr>
<tr>
<td>0.8</td>
<td>-1</td>
<td>-1</td>
<td>4.424</td>
<td>4.695</td>
<td>4.846</td>
<td>5.160</td>
<td>5.249</td>
<td>4.875</td>
</tr>
<tr>
<td>0.9</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>5.152</td>
<td>5.464</td>
<td>5.332</td>
<td>5.526</td>
<td>5.369</td>
</tr>
<tr>
<td>STD</td>
<td>0.459</td>
<td>0.704</td>
<td>0.847</td>
<td>1.226</td>
<td>1.324</td>
<td>1.317</td>
<td>1.221</td>
<td>2.206</td>
</tr>
</tbody>
</table>

(b) Original Image

<table>
<thead>
<tr>
<th>Segment Overlapping Percentage (SOP)</th>
<th>Color Bandwidth (CB)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AVG</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.5</td>
<td>5.5</td>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>11.408</td>
<td>10.904</td>
<td>9.651</td>
<td>7.994</td>
<td>6.171</td>
<td>4.032</td>
<td>4.733</td>
<td>7.842</td>
</tr>
<tr>
<td>0.2</td>
<td>11.316</td>
<td>10.704</td>
<td>9.378</td>
<td>7.727</td>
<td>5.8075</td>
<td>2.812</td>
<td>4.598</td>
<td>7.477</td>
</tr>
<tr>
<td>0.3</td>
<td>11.232</td>
<td>10.523</td>
<td>9.122</td>
<td>7.411</td>
<td>5.365</td>
<td>2.943</td>
<td>5.333</td>
<td>7.419</td>
</tr>
<tr>
<td>0.4</td>
<td>11.134</td>
<td>10.343</td>
<td>9.017</td>
<td>7.327</td>
<td>5.271</td>
<td>3.095</td>
<td>5.448</td>
<td>7.377</td>
</tr>
<tr>
<td>0.5</td>
<td>10.125</td>
<td>9.594</td>
<td>8.596</td>
<td>7.341</td>
<td>5.018</td>
<td>3.324</td>
<td>5.344</td>
<td>7.049</td>
</tr>
<tr>
<td>0.6</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>6.327</td>
<td>5.265</td>
<td>5.260</td>
<td>5.231</td>
<td>5.521</td>
</tr>
<tr>
<td>0.7</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>4.659</td>
<td>4.828</td>
<td>5.212</td>
<td>4.900</td>
</tr>
<tr>
<td>AVG</td>
<td>11.043</td>
<td>10.413</td>
<td>9.152</td>
<td>7.354</td>
<td>5.365</td>
<td>3.756</td>
<td>5.129</td>
<td>5.052</td>
</tr>
<tr>
<td>STD</td>
<td>0.523</td>
<td>0.504</td>
<td>0.396</td>
<td>0.567</td>
<td>0.497</td>
<td>0.972</td>
<td>0.328</td>
<td>4.214</td>
</tr>
</tbody>
</table>
The second parameter for evaluation is the window size used when performing the mean-shift segmentation. In this experiment, three types of window sizes had been tested. These sizes were 100 by 100, 150 by 150, and 200 by 200. The reason that we selected these three window sizes is the trade-off between computation time and segmentation robustness. If the window size is too small, the continuity of the segments may be affected. If the window size is too large, the computation time would significantly increase. These three sizes are chosen mainly based on the manageability of computation time. Table 5.5 shows the accuracy differences when different window sizes were used. The parameter set leading to the best shoreline accuracy was 1.5 for CB and 0.1 for SOP. There was no accuracy improvement when increasing the window size and it was the same for the standard deviation of accuracy. The result of this experiment showed that the increase of the segmentation window size did not result in better shoreline extraction. Since reducing window size could significant lower the computation time, the window size of 100 by 100 was selected for our procedure.

Next, the parameters CB, SB, and SOP were investigated. From the previous experiments, ranges that lead to better shoreline accuracy for parameter CB and SOP are acquired. However, we would like to know if the ranges of CB and SOP would remain the same after adding SB to the equation. A table with the six sections representing the accuracies of the extracted shoreline using the SOP parameter values of 0.1~0.6 was created (Table 5.6). Within each section, accuracies were estimated with the shorelines extracted using the CB parameter values ranged from 0.5 to 6.5 and the SB values ranged from 5 to 11. The goal of this experiment was to find range of CB and SB values respectively that could help to extract an accurate and robust shoreline. Figure 5.16
depicts the accuracy of shorelines in a graph with the data from Table 5.6. The results showed that when the SB value was 5, the most accurate shoreline could be obtained. As for the parameter CB, the generated shoreline could maintain relatively accurate and robust with the value between 1.5 and 4.5. Meanwhile, the parameter SOP could also be determined, the value of 0.1, 0.2, and 0.3 worked better with the water surface class.

After the investigation of the table, a manual inspection of the classification and extracted shoreline was also performed. The result of the manual inspection clearly showed that several parts of lands were missing in the shoreline extracted using the SOP value of 0.1 (Figure 5.17). We motioned this matter in section 3.3.5 that it would not be possible to remove these shoreline errors in the latter adjustment process. As a result, the SOP value of 0.1 could not be used even though the corresponding accuracies were among the best.

This section demonstrated the process of determining the parameter set and the result showed that it would be better to preprocess the panchromatic satellite imagery with the standard deviation image stretching, with the window size (SWS) of 100 by 100 when the performing mean-shift segmentation. These two parameters can be fixed and do not need to be evaluated even when other datasets or data sources are used. The other three parameters are on a case-by-case situation. CB and SD are data source dependent; in this case, they are tied to the WV2 satellite imagery that we used. Within the scope of our research, the mean-shift color bandwidth (CB) parameter value is between 0.5 and 4.5, and the mean-shift spatial bandwidth (SB) parameter value is 5. The SOP parameter is class dependent, and for the water surface class, the segment overlapping percentage (SOP) should be between 0.2 and 0.4.
There were totally four classes needed to be classified before extracting the shoreline. These classes were water surface, vegetation, shadow, and sediment bank/manmade object. Different parameter sets could be applied to these four classes in order to acquire a better classification. Based on the suggested parameter sets described in the previous paragraph, we plotted every classification using these parameter sets for each class, and manually selected the best classification. The key of evaluating the classification quality was the separation of water surfaces and shadow areas. The water surfaces should classify correctly, and should not overflow to other class types. Meanwhile, the shadow areas must correctly indicate the areas on water surfaces where shadows were projected. On the other hand, the sediment bank/manmade object class should cover as many actual dry sandy beaches and dry concrete structures as possible. Finally, the vegetation class should not overflow to other classes. In addition, where the vegetation class was present on the multispectral image must also be present on the classified image. Following these guidelines, one set of parameter was determined for each class.
Table 5.5 Shoreline accuracy with respect to the SWS, CB, and SOP parameters

<table>
<thead>
<tr>
<th>Color Bandwidth (CB)</th>
<th>SWS</th>
<th>SOP</th>
<th>6.5</th>
<th>5.5</th>
<th>4.5</th>
<th>3.5</th>
<th>2.5</th>
<th>AVG</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 by 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td>5.655</td>
<td>4.348</td>
<td>3.162</td>
<td>2.445</td>
<td>2.179</td>
<td>2.184</td>
<td>2.249</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td>4.988</td>
<td>3.680</td>
<td>2.961</td>
<td>2.438</td>
<td>2.461</td>
<td>2.730</td>
<td>2.374</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td>4.486</td>
<td>3.282</td>
<td>2.718</td>
<td>2.620</td>
<td>2.766</td>
<td>2.803</td>
<td>4.784</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td>4.265</td>
<td>3.334</td>
<td>3.220</td>
<td>4.987</td>
<td>5.009</td>
<td>3.034</td>
<td>4.860</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td>4.582</td>
<td>3.344</td>
<td>3.261</td>
<td>5.029</td>
<td>5.038</td>
<td>5.060</td>
<td>4.881</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td>5.002</td>
<td>5.016</td>
<td>4.969</td>
<td>5.035</td>
<td>5.137</td>
<td>5.066</td>
<td>4.956</td>
</tr>
<tr>
<td>STD</td>
<td>0.497</td>
<td>0.704</td>
<td>0.803</td>
<td>1.380</td>
<td>1.433</td>
<td>1.258</td>
<td>1.323</td>
<td>1.121</td>
<td></td>
</tr>
<tr>
<td>150 by 150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td>8.078</td>
<td>5.932</td>
<td>3.742</td>
<td>2.689</td>
<td>2.336</td>
<td>2.257</td>
<td>2.275</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td>7.214</td>
<td>4.792</td>
<td>2.777</td>
<td>2.512</td>
<td>2.494</td>
<td>2.519</td>
<td>2.740</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td>7.072</td>
<td>4.674</td>
<td>2.493</td>
<td>2.602</td>
<td>2.792</td>
<td>2.820</td>
<td>2.808</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td>6.979</td>
<td>4.637</td>
<td>3.473</td>
<td>2.766</td>
<td>5.001</td>
<td>3.021</td>
<td>3.254</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td>6.829</td>
<td>4.642</td>
<td>3.437</td>
<td>3.737</td>
<td>5.114</td>
<td>5.133</td>
<td>4.927</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td>6.627</td>
<td>5.703</td>
<td>4.945</td>
<td>3.824</td>
<td>5.153</td>
<td>5.210</td>
<td>4.938</td>
</tr>
<tr>
<td>STD</td>
<td>0.505</td>
<td>0.592</td>
<td>0.859</td>
<td>0.955</td>
<td>1.404</td>
<td>1.326</td>
<td>1.160</td>
<td>1.625</td>
<td></td>
</tr>
<tr>
<td>200 by 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td>9.512</td>
<td>6.576</td>
<td>4.432</td>
<td>2.705</td>
<td>2.314</td>
<td>2.199</td>
<td>2.370</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td>8.484</td>
<td>5.517</td>
<td>2.706</td>
<td>2.755</td>
<td>2.732</td>
<td>2.715</td>
<td>2.777</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td>8.453</td>
<td>5.444</td>
<td>2.820</td>
<td>2.610</td>
<td>2.769</td>
<td>2.809</td>
<td>4.723</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td>8.411</td>
<td>5.392</td>
<td>2.854</td>
<td>2.799</td>
<td>5.070</td>
<td>4.929</td>
<td>4.836</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td>8.364</td>
<td>5.369</td>
<td>3.246</td>
<td>3.703</td>
<td>5.113</td>
<td>5.019</td>
<td>4.909</td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td>8.576</td>
<td>6.186</td>
<td>4.789</td>
<td>5.108</td>
<td>5.177</td>
<td>5.085</td>
<td>4.988</td>
</tr>
<tr>
<td>STD</td>
<td>0.436</td>
<td>0.509</td>
<td>0.906</td>
<td>0.980</td>
<td>1.387</td>
<td>1.351</td>
<td>1.193</td>
<td>2.029</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.6 Shoreline accuracy with respect to the (a) SOP (0.1–0.3), (b) SOP (0.4–0.6), CB, and SB parameters

<table>
<thead>
<tr>
<th>SOP</th>
<th>SB</th>
<th>Color Bandwidth (CB)</th>
<th>AVG</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>7</td>
<td>5.655 4.348 3.162 2.445 2.179 2.184 2.249</td>
<td>3.175</td>
<td>1.347</td>
</tr>
<tr>
<td>0.1</td>
<td>9</td>
<td>5.531 4.665 3.613 2.519 2.257 2.253 2.274</td>
<td>3.302</td>
<td>1.341</td>
</tr>
<tr>
<td>0.1</td>
<td>11</td>
<td>6.357 5.375 3.851 3.102 2.388 2.171 2.241</td>
<td>3.641</td>
<td>1.654</td>
</tr>
<tr>
<td></td>
<td>AVG</td>
<td>5.838 4.594 3.384 2.575 2.240 2.195 2.268</td>
<td>3.299</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>0.364 0.589 0.426 0.372 0.110 0.039 0.030</td>
<td>1.370</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>5.125 3.055 2.319 2.284 2.440 2.663 2.436</td>
<td>2.903</td>
<td>1.014</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>4.988 3.680 2.961 2.438 2.461 2.730 2.374</td>
<td>3.090</td>
<td>0.952</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>5.103 4.017 2.929 2.489 2.449 2.351 2.377</td>
<td>3.102</td>
<td>1.062</td>
</tr>
<tr>
<td></td>
<td>AVG</td>
<td>5.271 3.885 2.909 2.516 2.434 2.524 2.399</td>
<td>3.134</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>0.403 0.723 0.455 0.240 0.034 0.201 0.029</td>
<td>1.069</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>5</td>
<td>4.488 2.620 2.203 2.182 2.749 2.822 2.892</td>
<td>2.851</td>
<td>0.776</td>
</tr>
<tr>
<td>0.3</td>
<td>7</td>
<td>4.486 3.282 2.718 2.620 2.766 2.803 4.784</td>
<td>3.351</td>
<td>0.906</td>
</tr>
<tr>
<td>0.3</td>
<td>9</td>
<td>4.798 3.756 2.820 2.660 2.687 2.765 4.763</td>
<td>3.464</td>
<td>0.975</td>
</tr>
<tr>
<td>0.3</td>
<td>11</td>
<td>5.657 4.492 3.484 2.826 4.847 2.705 2.788</td>
<td>3.828</td>
<td>1.175</td>
</tr>
<tr>
<td></td>
<td>AVG</td>
<td>4.857 3.537 2.806 2.572 3.262 2.774 3.807</td>
<td>3.374</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>0.553 0.789 0.526 0.275 1.057 0.051 1.117</td>
<td>0.980</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 5.6 continued

(b)

<table>
<thead>
<tr>
<th>SOP</th>
<th>SB</th>
<th>Color Bandwidth (CB)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AVG</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td></td>
<td>6.5</td>
<td>5.5</td>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.269</td>
<td>2.588</td>
<td>2.318</td>
<td>2.395</td>
<td>4.972</td>
<td>3.031</td>
<td>4.862</td>
<td>3.491</td>
<td>1.175</td>
</tr>
<tr>
<td>9</td>
<td>4.709</td>
<td>3.719</td>
<td>2.779</td>
<td>3.404</td>
<td>2.831</td>
<td>2.908</td>
<td>4.792</td>
<td>3.592</td>
<td>0.861</td>
</tr>
<tr>
<td>11</td>
<td>5.409</td>
<td>4.378</td>
<td>3.411</td>
<td>2.881</td>
<td>4.969</td>
<td>3.021</td>
<td>4.806</td>
<td>4.125</td>
<td>1.014</td>
</tr>
<tr>
<td>STD</td>
<td>0.539</td>
<td>0.748</td>
<td>0.487</td>
<td>1.125</td>
<td>1.077</td>
<td>0.061</td>
<td>0.036</td>
<td>0.980</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>6.5</td>
<td>5.5</td>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.999</td>
<td>5.397</td>
<td>5.183</td>
<td>5.055</td>
<td>5.108</td>
<td>5.042</td>
<td>4.928</td>
<td>4.959</td>
<td>0.448</td>
</tr>
<tr>
<td>7</td>
<td>4.582</td>
<td>3.344</td>
<td>3.261</td>
<td>5.029</td>
<td>5.038</td>
<td>5.060</td>
<td>4.881</td>
<td>4.456</td>
<td>0.805</td>
</tr>
<tr>
<td>9</td>
<td>4.457</td>
<td>3.541</td>
<td>2.814</td>
<td>3.519</td>
<td>4.978</td>
<td>5.069</td>
<td>4.825</td>
<td>4.172</td>
<td>0.878</td>
</tr>
<tr>
<td>11</td>
<td>4.922</td>
<td>4.390</td>
<td>3.397</td>
<td>3.121</td>
<td>5.050</td>
<td>5.094</td>
<td>4.911</td>
<td>4.412</td>
<td>0.824</td>
</tr>
<tr>
<td>STD</td>
<td>0.382</td>
<td>0.937</td>
<td>1.043</td>
<td>1.007</td>
<td>0.053</td>
<td>0.022</td>
<td>0.045</td>
<td>0.772</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td>6.5</td>
<td>5.5</td>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.330</td>
<td>4.803</td>
<td>4.723</td>
<td>5.101</td>
<td>5.135</td>
<td>5.076</td>
<td>4.958</td>
<td>5.018</td>
<td>0.207</td>
</tr>
<tr>
<td>7</td>
<td>5.002</td>
<td>5.016</td>
<td>4.969</td>
<td>5.035</td>
<td>5.137</td>
<td>5.066</td>
<td>4.956</td>
<td>5.026</td>
<td>0.062</td>
</tr>
<tr>
<td>9</td>
<td>5.374</td>
<td>4.769</td>
<td>4.871</td>
<td>5.019</td>
<td>5.056</td>
<td>5.105</td>
<td>4.910</td>
<td>5.015</td>
<td>0.196</td>
</tr>
<tr>
<td>11</td>
<td>4.754</td>
<td>4.553</td>
<td>5.012</td>
<td>5.077</td>
<td>5.038</td>
<td>5.155</td>
<td>4.912</td>
<td>4.929</td>
<td>0.210</td>
</tr>
<tr>
<td>AVG</td>
<td>5.115</td>
<td>4.785</td>
<td>4.894</td>
<td>5.058</td>
<td>5.092</td>
<td>5.100</td>
<td>4.934</td>
<td>4.997</td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td>0.292</td>
<td>0.190</td>
<td>0.128</td>
<td>0.038</td>
<td>0.052</td>
<td>0.040</td>
<td>0.026</td>
<td>0.174</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.16 Shoreline accuracies with different combinations of CB, SB, and SOP
Figure 5.17 An example of the shorelines extracted with the SOP parameter value of 0.1. In the upper-right corner where the Fairport Harbor Lighthouse is located, the water surface class (areas in blue) overflows to the shadow and manmade object classes and stops the continuity of the land. The extracted shoreline prematurely ended at the lighthouse and the shoreline of the breakwater is missing.

### 5.4.2 Accuracy Estimation

The values of these parameters used are listed in Table 5.7. The extracted shoreline using this set of parameters was evaluated for its accuracy and quality. Conducted evaluations included evaluations of accuracy, area differences, fuzziness factor, generalization factor, bias factor, and distortion factor comparing against the ground truth shoreline. These quality estimates are listed in Table 5.8. After completing this parameter set selection process, the accuracy of the shoreline extracted with this classification result could reach 1.8 m. According to the results of the individual terrain type shoreline accuracy analysis (Table 5.9), the accuracy of the sediment bank areas
were the best among all terrain types (0.8m). The accuracy of the groin areas were actually the highest of all terrain types. However, the total length of the groin areas was 69m and might not be significant statistically to conclude that the groin areas are having the best accuracy.

This accuracy estimation shows that the proposed shoreline extraction from the satellite imagery approach performed well with sediment banks, vertical structures, and groins. The accuracy corresponding to the piers is the worst. However, pier is one of the terrain types that our shoreline integration process targeted upon; hence we will leave the discussion for later. The accuracy corresponding to breakwaters is relatively low in this experiment. The breakwaters are basically riprap revetments on shallow water. Beaten by waves, they are usually wet concrete blocks and difficult to identify in a satellite image. However, the distinct materials they are made of can still help to separate them from the water surface class as an unclassified class in our classification process. While performing the scenario analysis, we did not consider the topographic of breakwaters. This breakwater scenario should be taken into consideration for the classification adjustment and refinement process (section 3.3.5) and potentially improve the accuracy within breakwaters.
Table 5.7 The parameter set chosen according to the parameter determination guidelines

<table>
<thead>
<tr>
<th>Class</th>
<th>Water Surface</th>
<th>Vegetation</th>
<th>Shadow</th>
<th>Sediment Bank / Manmade Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS</td>
<td>STD</td>
<td>STD</td>
<td>STD</td>
<td>STD</td>
</tr>
<tr>
<td>SWS</td>
<td>100*100</td>
<td>100*100</td>
<td>100*100</td>
<td>100*100</td>
</tr>
<tr>
<td>CB</td>
<td>4.5</td>
<td>4.5</td>
<td>5.5</td>
<td>4.5</td>
</tr>
<tr>
<td>SB</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>SOP</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 5.8 The quality estimates of the extracted shoreline

<table>
<thead>
<tr>
<th>Location</th>
<th>RMSE (m)</th>
<th>Maximum Error (m)</th>
<th>Average Error (m)</th>
<th>Area Difference (m²)</th>
<th>GF</th>
<th>BF</th>
<th>DF</th>
<th>OVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>East (Painesville)</td>
<td>1.900</td>
<td>14.540</td>
<td>1.619</td>
<td>5756.692</td>
<td>1.221</td>
<td>1.812</td>
<td>71.230</td>
<td>24.754</td>
</tr>
<tr>
<td>West (Fairport Harbor)</td>
<td>1.170</td>
<td>9.901</td>
<td>0.886</td>
<td>-95.291</td>
<td>1.100</td>
<td>1.723</td>
<td>63.488</td>
<td>22.103</td>
</tr>
<tr>
<td>Overall</td>
<td>1.796</td>
<td>14.540</td>
<td>1.464</td>
<td>5661.4</td>
<td>1.160</td>
<td>1.767</td>
<td>67.359</td>
<td>23.429</td>
</tr>
</tbody>
</table>

Table 5.9 The accuracies of the extracted shorelines with individual terrain types

<table>
<thead>
<tr>
<th>Shoreline Type</th>
<th>Length (m)</th>
<th>Maximum Error (m)</th>
<th>Average Error (m)</th>
<th>RMSE (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediment Bank</td>
<td>3099</td>
<td>7.555</td>
<td>0.665</td>
<td>0.782</td>
</tr>
<tr>
<td>Sloped Structure</td>
<td>4822</td>
<td>10.019</td>
<td>1.667</td>
<td>1.673</td>
</tr>
<tr>
<td>Vertical Structure</td>
<td>2839</td>
<td>14.540</td>
<td>0.839</td>
<td>1.477</td>
</tr>
<tr>
<td>Bluff</td>
<td>1621</td>
<td>7.343</td>
<td>1.703</td>
<td>1.395</td>
</tr>
<tr>
<td>Groin</td>
<td>69</td>
<td>1.651</td>
<td>0.660</td>
<td>0.468</td>
</tr>
<tr>
<td>Breakwater</td>
<td>96</td>
<td>7.560</td>
<td>2.551</td>
<td>2.058</td>
</tr>
<tr>
<td>Shadow</td>
<td>789</td>
<td>10.666</td>
<td>3.041</td>
<td>2.391</td>
</tr>
<tr>
<td>Pier</td>
<td>311</td>
<td>11.726</td>
<td>4.766</td>
<td>3.758</td>
</tr>
<tr>
<td>Total</td>
<td>13650</td>
<td>14.540</td>
<td>1.464</td>
<td>1.796</td>
</tr>
</tbody>
</table>
5.5 Accuracy Estimation of the Shoreline Integrated in Phase Three

This part of the experiment followed the procedures described in Chapter 4. The first step of this shoreline integration process is to handle the sloped structure areas. The information needed before determining if this step could be performed was the water level differences between the acquisition time of the LiDAR and the satellite imagery. The water level data were acquired from the NOAA National Ocean Service (NOS) gauge station at Fairport Harbor (The location of the gauge station is shown in Figure 5.1). The water level while the LiDAR point cloud data were acquired was around 571.43m, and the water level while the WV2 satellite imagery data were acquired was 571.09m. The result showed that the water level corresponding to the LiDAR data was 0.34m higher than the satellite imagery data; as a result, it was not possible to integrate the shoreline extracted from LiDAR data with the WV2 shoreline within the sloped structure areas described in section 4.1.

The second step was to handle the areas which appear to be in the process of erosion. The major area that benefited from this process is the location under the condition described in Figure 3.35. After this process, the areas in the satellite image that should have been classified as land but actually classified as water surface instead are identified and corrected. The overall accuracies of the extracted shoreline were improved. The major areas affected by this step are shown in Figure 5.18(a) at the location of Fairport Harbor Lighthouse and on the right side of the manmade riverbank in Figure 5.18(b). This type of error usually occurs with the vertical structure terrain type.

The third step was to handle the areas which appear to be in the process of accession. The major area corrected by this step was the piers within the study area. The
locations and boundaries of the piers were detected by an autonomous procedure, and as well as the line segments constructing the piers. However, the linking process of these line segments was inspected and adjusted manually after the line segments were extracted. In this process, pier and vertical structure among the terrain types are benefited; the accuracy of these terrain types was in fact improved. The accuracy estimates are listed in Table 5.10 and Table 5.11 and the extracted shorelines are demonstrated in Figure 5.18.

In Figure 5.18 (a), the shoreline extracted from LiDAR data misses a great length of shoreline on the sediment bank (sandy beach) due to the sedimentation of the beach. In Figure 5.18 (b), some of the shoreline did not extract correctly from the satellite imagery (marked in yellow polyline), shoreline extracted from LiDAR data corrected the errors and improved the accuracy of the integrated shoreline. In Figure 5.18 (d), the integrated shoreline within the pier is significantly improved from the satellite imagery extracted shoreline. In Figure 5.18 (g), integrated shoreline within the shadow areas is still having large errors. In Figure 5.18 (i), the shoreline integration procedure successfully corrected the sloped structure areas that are not crystal clear in the satellite imagery (upper-right where yellow line is present). In Figure 5.18 (j) and (k), the shoreline erosion could be clearly identified.

The result showed that the shoreline accuracy improved from 1.796m to 1.495m after this shoreline integration process. The accuracies with the following terrain types were significantly improved: sloped structure (from 1.673m to 1.556m), vertical structure (from 1.477m to 0.870m), and pier (from 3.758m to 0.856m) areas. The improvements were reflected exactly on the terrain types we targeted.

What we would like to discuss next is the shadow areas. Although we have
procedures (described in section 3.3.5) targeted on shadow areas, the corresponding accuracy is still the worst of all. Since LiDAR is an active scanning system, the formation of the shadow is totally different. Hence, the shoreline within shadow areas could be replaced by the LiDAR shoreline. The reason that we did not perform this process is because we needed to determine the terrain type within the shadow areas. Based on our accuracy analysis, the accuracy of the shoreline extracted in Phase One is the lowest. If we use the shoreline with worse accuracy to replace the shoreline within the shadowed areas in the satellite imagery, we might end up with an even worse accuracy than before the shoreline replacement. More testing and analyses are needed before we can provide any substantial suggestions to solve this shadow-related issue.

Table 5.10 The quality estimates of the integrated shoreline

<table>
<thead>
<tr>
<th>Location</th>
<th>RMSE (m)</th>
<th>Maximum Error (m)</th>
<th>Average Error (m)</th>
<th>Area Difference (m²)</th>
<th>GF</th>
<th>BF</th>
<th>DF</th>
<th>OVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>East (Painesville)</td>
<td>1.579</td>
<td>11.957</td>
<td>1.438</td>
<td>6906.192</td>
<td>1.205</td>
<td>2.124</td>
<td>102.443</td>
<td>35.257</td>
</tr>
<tr>
<td>West (Fairport Harbor)</td>
<td>0.978</td>
<td>9.750</td>
<td>0.758</td>
<td>662.709</td>
<td>1.142</td>
<td>1.539</td>
<td>7.696</td>
<td>3.459</td>
</tr>
<tr>
<td>Overall</td>
<td>1.495</td>
<td>11.957</td>
<td>1.288</td>
<td>7568.900</td>
<td>1.173</td>
<td>1.832</td>
<td>55.070</td>
<td>19.358</td>
</tr>
</tbody>
</table>
Table 5.11 The accuracies of integrated shoreline with individual terrain types

<table>
<thead>
<tr>
<th>Shoreline Type</th>
<th>Length (m)</th>
<th>Maximum Error (m)</th>
<th>Average Error (m)</th>
<th>RMSE (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediment Bank</td>
<td>3549</td>
<td>7.439</td>
<td>0.658</td>
<td>0.767</td>
</tr>
<tr>
<td>Sloped Structure</td>
<td>6376</td>
<td>9.802</td>
<td>1.615</td>
<td>1.556</td>
</tr>
<tr>
<td>Vertical Structure</td>
<td>3088</td>
<td>9.665</td>
<td>0.585</td>
<td>0.870</td>
</tr>
<tr>
<td>Bluff</td>
<td>1953</td>
<td>7.343</td>
<td>1.707</td>
<td>1.400</td>
</tr>
<tr>
<td>Groin</td>
<td>79</td>
<td>1.639</td>
<td>0.659</td>
<td>0.468</td>
</tr>
<tr>
<td>Breakwater</td>
<td>19</td>
<td>7.668</td>
<td>2.647</td>
<td>2.119</td>
</tr>
<tr>
<td>Shadow</td>
<td>889</td>
<td>11.957</td>
<td>3.084</td>
<td>2.468</td>
</tr>
<tr>
<td>Pier</td>
<td>322</td>
<td>4.111</td>
<td>0.869</td>
<td>0.856</td>
</tr>
<tr>
<td>Total</td>
<td>16381</td>
<td>11.957</td>
<td>1.288</td>
<td>1.495</td>
</tr>
</tbody>
</table>
Figure 5.18 Extracted shoreline of the entire study area. Ground truth shoreline in red, phase 1 (LiDAR) shoreline in magenta, phase 2 (WV2) shoreline in yellow, integrated shoreline in blue. Extracted shorelines along (a) Headlands Beach State Park, (b) along Grand River, (c) Fairport Harbor Lakefront Park, (d) recreational dock, (e) vertical and sloped structures, (f) sloped structures, (g) sloped and vertical structures, groins, and bluffs, (h) bluffs, and sloped structures, (i) sloped structures and breakwaters, (j) bluffs and sediment banks, and (k) bluffs.
Figure 5.18 continued
Figure 5.18 continued
Figure 5.18 continued
Figure 5.18 continued
Figure 5.18 continued
Figure 5.18 continued
Figure 5.18 continued

(i) Continued
Figure 5.18 continued
Figure 5.18 continued

(k)
5.6 Evaluation of the Shoreline Quality Metrics

Developing algorithms to extract shoreline is one thing, evaluating quality of the extracted shoreline is totally another. In this research, we incorporated several existing shoreline evaluating methods and two methods we developed. These methods are:

I. Shoreline accuracy estimation:
   a) Periodic nodes are created along the extracted shoreline, and the closest point on the ground truth line is found. The distance is measured from the node on the extracted shoreline to the closest point on the ground truth line. It is the most popular method for shoreline evaluation.
   b) Periodic perpendicular lines are created along the ground truth line that intersects the extracted shoreline. The distance is measured from the ground truth line to the extracted shoreline along the perpendicular line. This method was developed by this research.

II. Area difference of land areas. This method was developed by this research.

III. Shoreline quality metrics
   a) Generalization Factor (GF)
   b) Bias Factor (BF)
   c) Distortion Factor (DF)
   d) Overall Factor (OVF)

During the previous investigations, we have tried the I.a, I.b and II accuracy
estimation methods. Method I.a provides a robust estimation of the shoreline accuracy. It would not create impractical intersection points prolonging the distance of shoreline errors as method I.b would. Hence method I.a would be the most reliable overall shoreline accuracy indicator. The RMSE, Maximum Error, and Average Error in Table 5.12 were calculated based on I.a. From any aspect, these accuracy indicators made some accuracy improvements from Phase One, Phase Two, and Phase Three. The area difference indicator could represent large quantities of missing or exceeding lands. However, the overall accuracy could not be represented by this indicator. The area difference column in Table 5.12 shows the results of this indicator. In Phase One location west, this value is significant lower than the corresponding value in other datasets. The cause is the shoreline accession and our approach in Phase One extracted the shoreline based on the LiDAR dataset (this phenomenon is shown in Figure 5.7 and discussed in section 5.3). In Phase Two location west, this value is still lower than the value in Phase Three, due to the phenomenon shown and presented in Figure 5.17. Using this indicator to compare with other shorelines, you would notice that this indicator may show if there are issues within these extracted shorelines, though you have to dive in and figure out what exactly the problems are.

The generalization factor (GF) from the shoreline quality metrics measured the length ratio of the extracted shoreline and the ground truth shoreline. Ideally, this value should be close to 1. The length of shoreline was highly related to the type of equipment used (data sources and datasets) to create the shoreline. For example, a higher resolution satellite image would create a longer shoreline than LiDAR point cloud since the resolutions of data source were different. However, a longer shoreline did not necessarily
representing higher accuracy. In Table 5.12, the GF column in Phase One and Phase Two confirms the statement we given previously. The values are very close within the same set of shoreline, but significantly larger while comparing between the shorelines extracted in Phase One and Phase Two. Furthermore, this indicator was not significantly improved while shoreline accuracy was. The bias factor (BF) was to measure the ratio of extracted shoreline which fell on either side of the ground truth shoreline. If this factor was significantly larger or smaller than one (1), there might be a bias between these two shorelines. However, there was no normalization for the shoreline length, either. For example, if the distance of the extracted shoreline which falls on the left side of the reference shoreline is twice the distance on the right side, and the left side distance is used as a denominator and the right side distance as a numerator, the value come up with is 0.5. If other way around, the value will become 2.0. This is not a proportional indicator. In Table 5.12, the BF values are smaller than 1 in Phase One and larger than 1 in Phase Two. We can imagine from the values that in Phase One the shoreline is closer toward inland, and in Phase Two the shoreline is closer to the water surface. This is in fact a correct representation of the extracted shoreline, since in Phase One the task is to extract the last LiDAR points on land, and in Phase Two is to extract the water surface boundary. Distortion factor (DF) was also to measure the length difference between extracted shoreline and ground truth shoreline. The factor assumed that the length of the shoreline was equally distributed in proportion. Ideally, this value should be 0. However, the length of shoreline was affected by data sources, terrain types, objects along the shoreline, etc. For example, a sloped structure made by riprap revetment on a higher resolution data source would lead to a longer shoreline; however, a vertical structure would not change
the shoreline length significantly while switching the resolutions of data sources. Without considering these factors, the longer the shoreline stretched, the larger the error would be. Besides, there was no normalization for the shoreline length either. In our test area, the eastern shoreline (10.8 km) is significantly longer than the western shoreline (3.2 km). Every DF value of the eastern shoreline is significantly larger than that of the western shoreline in Table 5.12. Considering the overall DF indicators, it seems that the integrated shoreline would lead to the best DF value. But after further investigation into the eastern and western shorelines, you would certainly know that the reason is not because the DF can well represent the shoreline accuracy. The shoreline accuracy in Phase One west is the worst, but the DF value is actually small. Finally, the overall factor (OVF) was just an average of the three shoreline quality metric factors, with no normalization of any kind. From the previous statements it is understandable that most of the indicators are not normalized, and thus a direct averaging of these indicators does not make any sense. In Table 5.12, although OVF seems to get smaller while the accuracy improves, in fact, the value is dominated by DF. As a result, we do not advice using OVF as a shoreline quality indicator. Based on the theory and observation of these shoreline quality indicators, indicators I.a, II, and III.b were useful to represent the quality of shorelines.
Table 5.12 Shoreline quality indicators with respect to the shorelines extracted in Phase One, Phase Two and Phase Three

<table>
<thead>
<tr>
<th>Shoreline Extraction Approach</th>
<th>Location</th>
<th>RMSE (m)</th>
<th>Maximum Error (m)</th>
<th>Average Error (m)</th>
<th>Area Difference (m²)</th>
<th>GF</th>
<th>BF</th>
<th>DF</th>
<th>OVF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase One (Shoreline Erosion/ Accession Included)</td>
<td>East (Painesville)</td>
<td>2.475</td>
<td>18.290</td>
<td>2.410</td>
<td>6778.345</td>
<td>1.026</td>
<td>0.803</td>
<td>140.916</td>
<td>47.582</td>
</tr>
<tr>
<td></td>
<td>West (Fairport Harbor)</td>
<td>5.033</td>
<td>26.018</td>
<td>3.081</td>
<td>-5611.72</td>
<td>1.026</td>
<td>0.920</td>
<td>21.659</td>
<td>7.868</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>3.258</td>
<td>26.018</td>
<td>2.564</td>
<td>1166.625</td>
<td>1.026</td>
<td>0.861</td>
<td>81.287</td>
<td>27.725</td>
</tr>
<tr>
<td>Phase Two</td>
<td>East (Painesville)</td>
<td>1.900</td>
<td>14.540</td>
<td>1.619</td>
<td>5756.692</td>
<td>1.221</td>
<td>1.812</td>
<td>71.230</td>
<td>24.754</td>
</tr>
<tr>
<td></td>
<td>West (Fairport Harbor)</td>
<td>1.170</td>
<td>9.901</td>
<td>0.886</td>
<td>-95.291</td>
<td>1.100</td>
<td>1.723</td>
<td>63.488</td>
<td>22.103</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>1.796</td>
<td>14.540</td>
<td>1.464</td>
<td>5661.4</td>
<td>1.160</td>
<td>1.767</td>
<td>67.359</td>
<td>23.429</td>
</tr>
<tr>
<td>Phase Three Integrated Shoreline</td>
<td>East (Painesville)</td>
<td>1.579</td>
<td>11.957</td>
<td>1.438</td>
<td>6906.192</td>
<td>1.205</td>
<td>2.124</td>
<td>102.443</td>
<td>35.257</td>
</tr>
<tr>
<td></td>
<td>West (Fairport Harbor)</td>
<td>0.978</td>
<td>9.750</td>
<td>0.758</td>
<td>662.709</td>
<td>1.142</td>
<td>1.539</td>
<td>7.696</td>
<td>3.459</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
<td>1.495</td>
<td>11.957</td>
<td>1.288</td>
<td>7568.900</td>
<td>1.173</td>
<td>1.832</td>
<td>55.070</td>
<td>19.358</td>
</tr>
</tbody>
</table>
5.7 Summary and Discussion

According to the accuracy requirements for NGS shoreline mapping projects, the horizontal accuracy at the 95% confidence level is 1m in harbors, ports, channels, etc; 3m in approach areas to ports; and 5m in open coastal areas (NGS, 2011). “Accuracy reported at the 95% confidence level means that 95% of the positions in the dataset will have an error with respect to true ground position that is equal to or smaller than the reported accuracy value (FGDC, 1998)”. According to this definition, the accuracies of 95% confidence level are calculated and summarized in Table 5.13. Corresponding to the terrain types we have defined, most of harbors, ports, channels are coastal vertical structures. Approach areas to ports are usually coastal vertical structures or coastal sloped structures. The accuracy of our integrated shoreline at the 95% confidence level is 4.452m. The accuracy within vertical structures is 2.280m and 4.911m within sloped structure areas. The accuracy does not satisfy the accuracy requirement of 1m within harbors, ports, and channels, and 3m within approach areas to ports defined by the NGS, NOAA.
Table 5.13 The accuracies of integrated shoreline with individual terrain types including accuracy at 95% confidence level

<table>
<thead>
<tr>
<th>Shoreline Type</th>
<th>Length (m)</th>
<th>Average Error (m)</th>
<th>RMSE (m)</th>
<th>Accuracy at 95% Confidence Level (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sediment Bank</td>
<td>3549</td>
<td>0.658</td>
<td>0.767</td>
<td>2.339</td>
</tr>
<tr>
<td>Sloped Structure</td>
<td>6376</td>
<td>1.615</td>
<td>1.556</td>
<td>4.811</td>
</tr>
<tr>
<td>Vertical Structure</td>
<td>3088</td>
<td>0.585</td>
<td>0.870</td>
<td>2.280</td>
</tr>
<tr>
<td>Bluff</td>
<td>1953</td>
<td>1.707</td>
<td>1.400</td>
<td>4.567</td>
</tr>
<tr>
<td>Groin</td>
<td>79</td>
<td>0.659</td>
<td>0.468</td>
<td>1.445</td>
</tr>
<tr>
<td>Breakwater</td>
<td>19</td>
<td>2.647</td>
<td>2.119</td>
<td>6.773</td>
</tr>
<tr>
<td>Shadow</td>
<td>889</td>
<td>3.084</td>
<td>2.468</td>
<td>7.779</td>
</tr>
<tr>
<td>Pier</td>
<td>322</td>
<td>0.869</td>
<td>0.856</td>
<td>2.991</td>
</tr>
<tr>
<td>Total</td>
<td>16381</td>
<td>1.288</td>
<td>1.495</td>
<td>4.452</td>
</tr>
</tbody>
</table>

In order to meet the NGS accuracy requirements, the accuracies within vertical structures, sloped structures, and piers need to be improved. The accuracies of these three types of terrains obtained using our proposed approach are already the best, but they still fall short of the NGS standards. In the meantime, the shoreline within the harbors, ports and channels can be manually delineated to replace the shoreline extracted by the proposed approach in order to achieve the required accuracy. On the other hand, further investigations to improve the shoreline extraction accuracy are needed. Within the vertical structure areas, most of the inaccuracy is caused by shadows (Figure 5.18 (g) the coastline near the baseball field). For the sloped structure areas, we can apply a
shoreline integration process to improve the shoreline accuracy specifically for this terrain type. Unfortunately, the water level at the time of the satellite image acquisition is lower than that at the time of the LiDAR data acquisition. Thus, the integration process that is anticipated to improve the accuracy for sloped areas cannot be performed. Another set of data is needed to evaluate the performance of the integrated shoreline within the sloped areas. In the pier areas, the berthed boat at the piers still cannot be 100% identified and removed from the satellite imagery. The two piers in Figure 5.18 (c) (one on the left side of the figure along the river and one near the center of the figure) can clearly show the problems the boats have caused. In other terrain types, after inspecting the shoreline in Figure 5.18, several issues are still visible. In Figure 5.18 (a) on the upper-right corner where the light house is located, there is a patch of shadow that was not successfully removed, resulting in an arrow-shaped segment of shoreline. In Figure 5.18 (g) and (k), the error caused by shadow is still significant. Other than the issues described here, there are still several issues regarding our approach that we have discussed in sections 5.3~5.5 including the modification of the decision tree for Phase One, adding a breakwater scenario into the Phase Two process, and the shoreline accuracy improvements in the shadow areas. If these three issues can be resolved, the performance of this proposed approach may be further improved.

The approach developed by this research is designed to work not only on the Painesville datasets but also datasets of other locations. Our study area is within a non-tidal water body. The major doubt would be if through this approach shoreline extraction could be performed with tidal coastlines. Tidal effect is actually similar to erosion/accession when seeing it from another point of view. The main reason that we
decided to analysis erosion/accession as our integration process is we know handling tidal effect would be a necessity capability for shoreline mapping. There would be no significant difference between performing shoreline extraction using the proposed approach in tidal areas and doing so in non-tidal areas. When a new dataset is used, some adjustment to the procedure is needed depending on the condition of the data sources. For LiDAR data, an inspection whether there are LiDAR points on water surface must be done. If there are no LiDAR points on water surface, the process of mean-shift segmentation of elevation described in Figure 3.2 has to be skipped. For satellite imagery, the thresholds described in section 3.3.1 must be re-evaluated. And the process of determining the parameters described in section 5.4 needs to be performed as well. After these processes are performed, shorelines could be extracted from both Phase One and Phase Two. In the shoreline integration process, no parameter needs to be determined, however, water level information is required to evaluate the water level difference between datasets. We may not have taken all terrain types in the world into consideration, some scenarios may need to be included in the process while performing the shoreline integration. However, we have covered most of the terrain types in Painesville. If the location that you would like to work on is of a topography which is similar to Painesville, there should be no problem to applying our proposed approach to your case without any further modification.
Chapter 6. Conclusion and Future Research

6.1 Conclusions

The objective of this research was to develop a new shoreline delineation approach in order to incorporate relatively lower cost data sources and reduce human labor while maintaining reasonable shoreline accuracy. The data sources used including the LiDAR data, satellite imagery, and gauge station water level information. The accuracy of the extracted shoreline could reach an accuracy of 1.495 m RMSE, or 4.452m at the 95% confidence level. A further investigation of the accuracy in different terrain type was performed. The contributions of this dissertation are listed below:

1) Two shoreline extraction approaches were developed. One extracts shoreline from LiDAR points with multispectral information from satellite imagery. The other extracts shoreline solely from satellite imagery. These two approaches could be performed independently and not necessarily integrate the extracted shoreline.

2) The first approach which extracts shoreline from LiDAR points with multispectral information from satellite imagery could tolerate LiDAR systematic errors residing in LiDAR points. Since empirical values are used for parameters, this approach could be considered as a fully autonomous shoreline extraction approach. The extracted shoreline could reach the accuracy of 2.152 m RMSE.

3) The second approach extracts shoreline solely from satellite imagery utilizing
object-oriented concepts to classify object materials and terrain types. The usage of shadow effect for analyzing terrain topography for image classification is also a brand-new concept. The extracted shoreline could reach an accuracy of 1.796 m RMSE.

4) A shoreline integration approach was developed to integrate the extracted shoreline from these shoreline extraction approaches. The integration of two shorelines from two time periods is a brand-new solution for shoreline integration. With the knowledge of the terrain types from the second shoreline extraction approach, the information to predict how the shoreline may change through time is provided. By comparison with the evidence regarding the extracted shorelines from previous approaches, the complete picture of what the shoreline type is and how the shoreline has changed can be revealed. Based on this knowledge, shorelines from different time periods can be integrated. The accuracy of the integrated shoreline could reach 1.495 m RMSE.

5) Among the data sources used in this approach, the only one with costs is the satellite imagery. The need of manpower is significantly less compared to ground surveying or aerial photogrammetry. The processes that needed human attention are the parameter determination in second approach and the shoreline delineation within piers in the shoreline integration process.

6.2 Future Works

Within our scope of research, there are several issues that our proposed approach did not perform well or could be improved. These issues include:
1) The accuracy of the extracted shoreline at the 95% confidence level does not satisfy the accuracy requirement defined by the NGS, NOAA. The accuracies of the extracted shoreline within the areas of sloped structures, vertical structures, and piers needed to be improved in order to meet the standards (details of the analysis are described in section 5.7).

2) While extracting shoreline from LiDAR points with multispectral information, the extracted shoreline may sometimes represent the shoreline of LiDAR points and sometimes represent the shoreline of satellite imagery (details of the analysis are described in section 5.3). A possible solution is described and should be tested in the future.

3) Shoreline accuracy performed worse on breakwater structures for there were no scenarios established for breakwaters in satellite imagery shoreline extraction approach (details of the analysis are described in section 5.4.2). A lead for solving this issue is provided and should be tested in the future.

4) The shoreline accuracy performance was the worst on shadow areas. Although we have the procedures specific for improving shoreline extraction within shadow areas, the accuracy still performance was still the worst among all terrain types (details of the analysis are described in section 5.5 and 5.7). This issue should be further investigated in the future.

5) The water levels are different when the LiDAR data and the satellite imagery are acquired. However, the water level with the LiDAR data is higher than the satellite imagery. As a result, we cannot perform the integration procedure specific for integrating sloped structures. Other datasets should be acquired to test
the performance of this integration process (details of the analysis are described in section 4.1 and 5.5).

Beyond our scope of research, the ultimate goal of our research should be the extraction of tide-coordinated shoreline. Tide-coordinated shorelines are the most meaningful and useful shorelines in applications. Finding a way to produce tide-coordinated shorelines from the instantaneous shorelines extracted using the approach developed by this research will broaden applications significantly. The possible way of extracting a tide-coordinated shoreline from instantaneous shoreline would be extracting multiple instantaneous shorelines while water levels are different within a short period of time. Each instantaneous shoreline represents a contour line on the terrain at the corresponding water level. Hence, DEM can be created using these multiple contour lines. After the DEM is created, tide-coordinated shorelines at any water level within the range between the highest and the lowest instantaneous shorelines can be generated using this DEM. However, there may be some practical issues with the process of acquiring multiple satellite images within a short period of time. For example, most of the remote sensing satellites are in sun-synchronous orbits, a satellite passes the same location approximately the same local time every time. Within a short time period, water level at the same local time every day is also about the same. Further research is definitely needed to overcome the issues and eventually extract tide-coordinate shorelines form the approach developed by this research.

Shoreline mapping has always been a time consuming task for either aerial photogrammetry based or LiDAR based approach. It is impossible to acquire a shoreline whenever you desire. However, the approach proposed by this dissertation provides new
possibilities for extracting shorelines. For example, with disasters like hurricane Katrina, or tsunami in Japan or south Asia, it is difficult to have an up to NGS standard aerial photogrammetry or LiDAR datasets acquired during the time of occurrence, since every resource would be put into assisting victims and GPS reference stations may not be available. If there are existed LiDAR dataset, our approach could be adopted to create shorelines in just a few hours after the satellite imagery is received with the shoreline accuracy satisfying the NGS standard. However, the accuracy of the extracted shoreline first needs to be improved and meet the NGS standard, and different satellite sensors and different point spacing LiDAR datasets are needed to test the ability of this proposed approach before we can put this approach into practice.
References


Choung, Y., 2009. Extraction of bluffs from 2.5 dimensional Delaunay triangle mesh using LiDAR data. Master thesis, The Ohio State University, Columbus, Ohio.


Crombaghs, M., E. De Min and R. Bruegelmann, 2000. On the Adjustment of


Spatially Delineating a Contiguous, estuarine shoreline of North Carolina, North Carolina of Coastal Management.


Leica, 2012. Leica ADS80Datasheet. URL: 


Fiorentina, Florence, Italy, pp. 81 – 95.


Ohio Department of Natural Resources (ODNR), 2010. Lake Erie Coastal Erosion Area. URL: http://www.dnr.state.oh.us/tabid/7959/Default.aspx (last accessed: 07/17/12)

Ohio Geographically Referenced Information Program (OGRIP), 2012. Ohio Statewide Imagery Program. URL: http://ogrip.oit.ohio.gov/ProjectsInitiatives/StatewideImagery.aspx (last accessed: 07/17/12)

Ramirez, J. 2000, Quality Evaluation of Linear Features, A white paper submitted to NIMA.


Toth, C., and D. A. Grejner-Brzezinska, 2009. Airborne LiDAR Reflective Linear Feature Extraction for Strip Adjustment and Horizontal Accuracy Determination, The Ohio State University, Columbus, Ohio.


