NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

URBAN CLASSIFICATION TECHNIQUES USING THE FUSION OF LIDAR AND SPECTRAL DATA

by

Justin E. Mesina

September 2012

Thesis Advisor:	Richard C. Olsen
Second Reader:	Fred A. Kruse

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704–0188) Washington DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Urban Classification Techniques Using the Fusion of LiDAR and Spectral Data		**5. FUNDING NUMBERS**	
6. AUTHOR Justin E. Mesina			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number NA.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		**12b. DISTRIBUTION CODE** A	

13. ABSTRACT (maximum 200 words)

Combining different types of data from varying sensors has the potential to be more accurate than a single sensor. This research fused airborne LiDAR data and WorldView-2 (WV-2) multispectral imagery (MSI) data to create an improved classification image of urban San Francisco, California. A decision tree scenario was created by extracting features from the LiDAR, as well as NDVI from the multispectral data. Raster masks were created using these features and were processed as decision tree nodes resulting in seven classifications. Twelve regions of interest were created, then categorized and applied to the previous seven classifications via the maximum likelihood classification. The resulting classification images were then combined. A multispectral classification image using the same ROIs was also created for comparison. The fused classification image did a better job of preserving urban geometries than MSI data alone and suffered less from shadow anomalies. The fused results however, were not as accurate in differentiating trees from grasses as using only spectral results. Overall the fused LiDAR and MSI classification performed better than the MSI classification alone but further refinements to the decision tree scheme could probably be made to improve final results.

14. SUBJECT TERMS Fusion, Multi-Source, Hyperspectral, Multispectral, LiDAR, Urban Classification			**15. NUMBER OF PAGES** 85
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01–280–5500 Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release, distribution is unlimited

URBAN CLASSIFICATION TECHNIQUES USING THE FUSION OF LIDAR AND SPECTRAL DATA

Justin E. Mesina
Civilian, Department of the Navy
M.B.A., Avila University, 2010
B.S., University of North Texas, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN REMOTE SENSING INTELLIGENCE

from the

NAVAL POSTGRADUATE SCHOOL
September 2012

Author: Justin E. Mesina

Approved by: Richard C. Olsen
Thesis Advisor

Fred A. Kruse
Second Reader

Dan C. Boger
Chair, Department of Information Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Combining different types of data from varying sensors has the potential to be more accurate than a single sensor. This research fused airborne LiDAR data and WorldView-2 (WV-2) multispectral imagery (MSI) data to create an improved classification image of urban San Francisco, California. A decision tree scenario was created by extracting features from the LiDAR, as well as NDVI from the multispectral data. Raster masks were created using these features and were processed as decision tree nodes resulting in seven classifications. Twelve regions of interest were created, then categorized and applied to the previous seven classifications via the maximum likelihood classification. The resulting classification images were then combined. A multispectral classification image using the same ROIs was also created for comparison. The fused classification image did a better job of preserving urban geometries than MSI data alone and suffered less from shadow anomalies. The fused results however, were not as accurate in differentiating trees from grasses as using only spectral results. Overall the fused LiDAR and MSI classification performed better than the MSI classification alone but further refinements to the decision tree scheme could probably be made to improve final results.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AARA - American Recovery and Reinvestment Act

ALTM - Airborne Laser Terrain Mapper

AVIRIS - Airborne Visible/Infrared Imaging Spectrometer

BIL - Band Interleaved by Line

BSP - Binary Space Partitioning

BSQ - Band Sequential

E3De - Environment for 3D Exploitation

ENVI – The Environment for Visualizing Images

EO - Electro-optical

ERTS - Earth Resources Technology Satellite

FLAASH - Fast Line-of-Sign Atmospheric Analysis of Spectral Hypercubes

GPS - Global Positioning System

HSI - Hyperspectral Imaging

IDL - Interactive Data Language

IMU - Inertial Measurement Unit

JPL - Jet Propulsion Laboratory

LAS - Laser File Format

LiDAR - Light Detection and Ranging

LVIS - Laser Vegetation Imaging Sensor

MPiA - Multi-Pulse in Air

MSI - Multispectral Imaging

NAD83 - North American Datum of 1983

NASA - National Aeronautics and Space Administration

NDVI - Normalized Difference Vegetation Index

NGA - National Geospatial-Intelligence Agency

NGVD88 - North American Vertical Datum of 1988

PDOP - Positional Dilution of Precision

QTM - Quick Terrain Modeler

Radar - Radio Detection and Ranging

SAR - Synthetic Aperture Radar

SHARE - SpecTIR Hyperspectral Airborne Rochester Experiment

UFD - Urban Feature Data

USGS - U.S. Geological Survey

UTM - Universal Transverse Mercator

VIS - Visual Information Solutions

WGS84 - World Geodetic System of 1984

WV-2 - WorldView-2

ACKNOWLEDGMENTS

I would like extend thanks to my thesis advisor Prof. Richard C. Olsen and my second reader, Prof. Fred A. Kruse for their support and guidance this past year both in coursework and with this thesis. Thanks also the Remote Sensing Center staff, particularly Mrs. Jean Ferreira, Mrs. Dione Martinsen, and Mrs. Bree Winn for lending a helpful ear when needed and keeping things in perspective.

To all my friends and colleagues in the Remote Sensing Intelligence Curriculum, especially my office-mates Maj. Ralph Shoukry (USAF), Ms. Catherine Darby, and Ms. Michelle Sandersfeld, thank you for the reassurance and unity as we pushed through those long hours of study sessions and lab work.

A warm thanks to CAPT Carol O'Neal (USN ret.) and all of the Del Monte Brass for allowing me to have a musical getaway every week during my studies at NPS and allow me to be part of some great military events on campus.

And of course, thanks to my entire family for their love and support: my mother-Ofelia Mesina, father-Hernando Mesina, sister-Abigail Yu, and brother-in-law-Clinton Yu. Thank you for all your encouragement and visits out to Monterey.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PURPOSE OF RESEARCH

Two of the latest remote sensing technologies include light detection and ranging (LiDAR) and hyperspectral imaging (HSI). LiDAR is an active system similar to that of radar but sends visible and infrared pulses to calculate distances and produce a 3-dimensional point cloud of ground structures. LiDAR has a unique advantage of being able to penetrate through foliage to capture some aspects within and below vegetation. Hyperspectral imaging is a passive system that captures distinct spectra of ground features, exploiting electronic characteristics and molecular vibrations to identify and classify materials.

The purpose of this research was to look into techniques to fuse LiDAR and spectral data to classify urban environments. These two datasets were expected to complement each other and optimize classification capabilities. This thesis utilized WorldView-2 (WV-2) imagery. While technically an 8-band multispectral imaging (MSI) system, this imagery was chosen due to its higher spatial resolution and availability.

Although highly capable in their own right, LiDAR and spectral information do lack certain details. LiDAR provides detailed information regarding geometries such as spatial distances, heights, and canopy penetration but lacks any information concerning the particularities in the electromagnetic spectrum. Spectral provides highly detailed electromagnetic information to the point of material identification, but it is limited to two dimensions without spatial information in the ‘z’ or height dimension. These technologies are uniquely matched to lead to fusion opportunities.

Classification techniques ranging from building extraction to vegetation species identification are all available for comparison and combination. Although lacking the spectral resolution of a true hyperspectral sensor, the WorldView-2 satellite from DigitalGlobe has a unique advantage in being accessible to federal government, local

government, and private organizations. This thesis looked at the fusion of LiDAR and WorldView-2 data, but techniques developed here should be applicable to imaging spectroscopy.

B. OBJECTIVE

The primary objective of this thesis was to use the fusion of LiDAR and multispectral data to classify the urban environment of downtown San Francisco, California. The LiDAR data were collected as part of the American Recovery and Reinvestment Act's (ARRA) Golden Gate LiDAR Project (GGLP) in the summer of 2010. The WorldView-2 data were acquired via DigitalGlobe with satellite imagery collected in autumn of 2011.

Downtown San Francisco is an area which includes a variety of ground materials ranging from coastal waters, beaches, and parks to urban housing and large skyscrapers. The final fused product is a classified urban image based upon criteria from both of the datasets. The goal is to create a LiDAR and MSI fused classified urban image that is more representative of reality than a classified urban image based on multispectral data alone.

In the background chapter there is an overview of LiDAR and electro-optical (EO) imaging presented along with information on previous work using single-source and multi-source fusion techniques. The Problem and Methods sections provide further information regarding the study area, software, methodologies used, and the actual application of the technique. The Evaluation and Summary section offer conclusions from the process and assesses the results.

II. BACKGROUND

This chapter briefly looks at the fundamental operations as well as classification methods of a LiDAR imaging system and a multispectral system. This also takes an in-depth look at the variety of techniques that have been previously used that take advantage multi-source fusion. Considering what has been accomplished in the past, it then discusses some of the theory behind this project. The last part of the Background chapter discusses the features of the area of interest, San Francisco, California.

A. LIGHT DETECTION AND RANGING

1. LiDAR Fundamentals

Light detection and ranging is a remote sensing technique that works similar to radio detection and ranging (radar). These systems are known collectively as active imaging systems, as they emit electromagnetic pulses and time their return in order to detect an object's distance. LiDAR uses ultraviolet, visible, or near infrared wavelength laser pulses rather than microwaves. LiDAR systems can be terrestrial, airborne, or space-borne. Commonly, terrestrial systems are used for 3D modeling, whereas air and space systems are used for wide area mapping. This paper focuses on airborne systems (Crutchley & Crow, 2009).

When a laser pulse is emitted, it hits the surface of an object and is backscattered. Some scattered light is then returned towards the originating sensor and detected by a photo-detector. LiDAR sensors are also equipped with highly accurate position detection systems. Using both the Global Positioning System (GPS) constellation as well as an on-board Inertial Measurement Unit (IMU), the LiDAR system can achieve an accurate absolute position and sensor orientation with respect to the Earth. The returned pulse is received as a waveform, or a long pulse with differing rates of intensities. Waveform information is typically measured amongst a set of thresholds and is broken into a set of distinct returns. Combining this information with the time difference information, points are generated with a latitude, longitude, and elevation. See Figure 1 for the typical data exchange from an airborne system (Crutchley & Crow 2009).

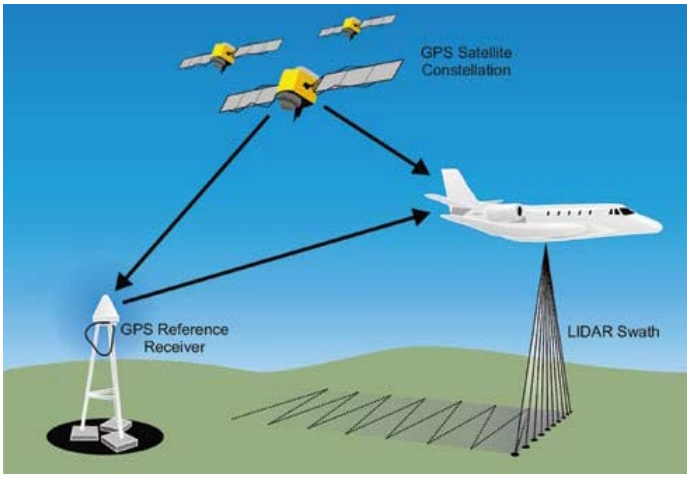

Figure 1. LiDAR data flow in an airborne system (From Holden et al., 2002)

After scanning an area, multiple returns and points are combined with each other in what is known as a point cloud. Point clouds are representative models of an area and are processed further to create products such as a digital surface model (DSM) and digital elevation model (DEM). Figure 2 shows an example of point cloud results after data processing.

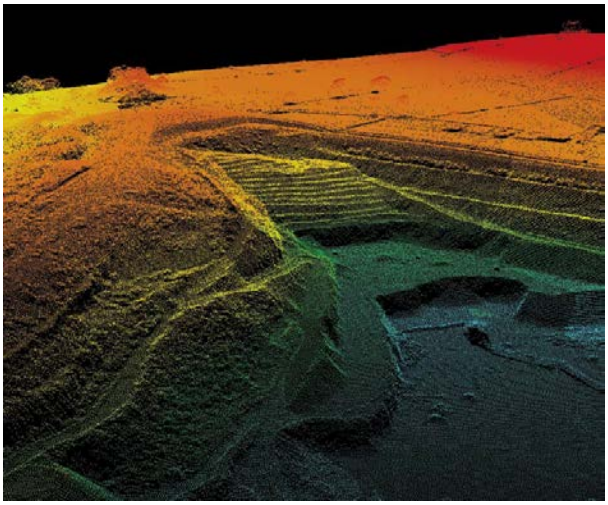

Figure 2. Example of a point cloud after processing and colored by height (From Cambridge University, 2006)

2. LiDAR Classification

With the values LiDAR provides of elevation and intensity, classification is possible with the point cloud alone. In a study by the University of Cambridge and University of Wales, they created land cover type classification employing elevation, intensity, and also point distribution frequency. Their study area included the meandering areas of the Garonne and Allier rivers in France. It was determined that clear water had the lowest reflectance of 0–10%, vegetation was about 50%, and soils were up to 57% with the highest reflectance. The classification method used a series of criteria based on height, intensity, and distribution which was then processed in the geographic information system ArcGIS and the programming languages C++ and MATLAB. When classifying land types, they achieved an accuracy of 95% and 94%. When classifying riparian forests, their accuracy varied from 66% and 98%. The study area consisted of natural and rural environments. Figure 3 shows one of their results near the Chatel Meander on the Allier River.

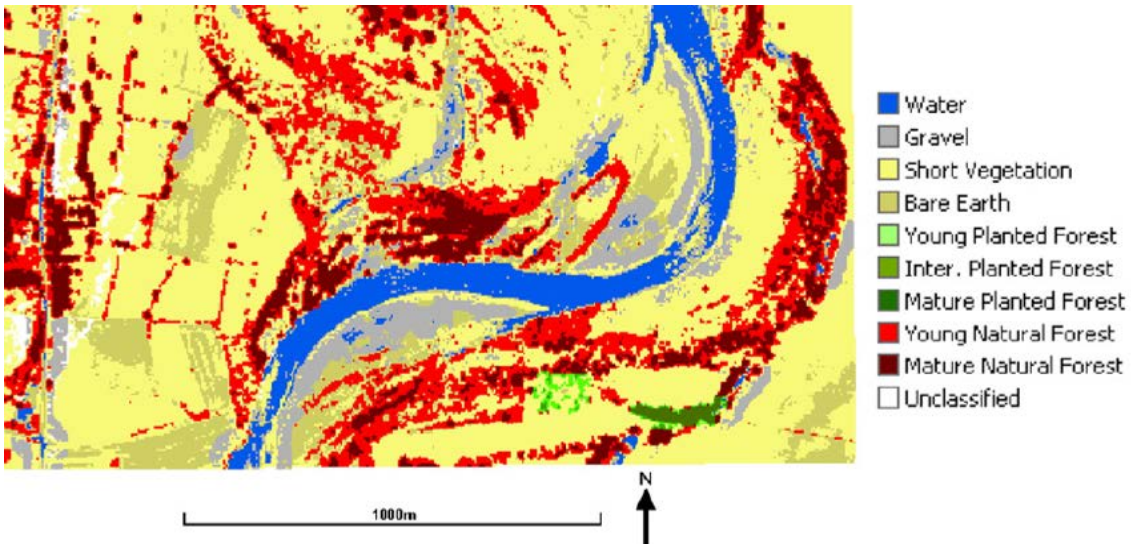

Figure 3. Results from LiDAR-only classification near the Chatel Meander of the Allier River in France (From Antonarakis et al., 2008)

In a thesis from the Naval Postgraduate School, LiDAR data was used to identify tree vegetation in the Elkhorn Slough of central California. With known vegetation characteristics of the study site, identification could be accomplished. QuickBird

multispectral imagery was used to identify regions of interest with Eucalyptus, Scrub Oak, Live Oak, and Monterey Cyprus trees. Tree types such as Eucalyptus and Oak trees were separated by differing return data. It was found that the Monterey Cyprus and Eucalyptus trees were similar in dimension and were separated by foliage density based on LiDAR return intensities. Density characteristics were analyzed as well as LiDAR intensity characteristics of the regions of interest. The conclusion was that LiDAR could be used to identify vegetation; however a detailed knowledge of the vegetated area must be collected and known via on-site surveys. Figure 4 shows the composite results for the LiDAR vegetation classification (Helt, 2005).

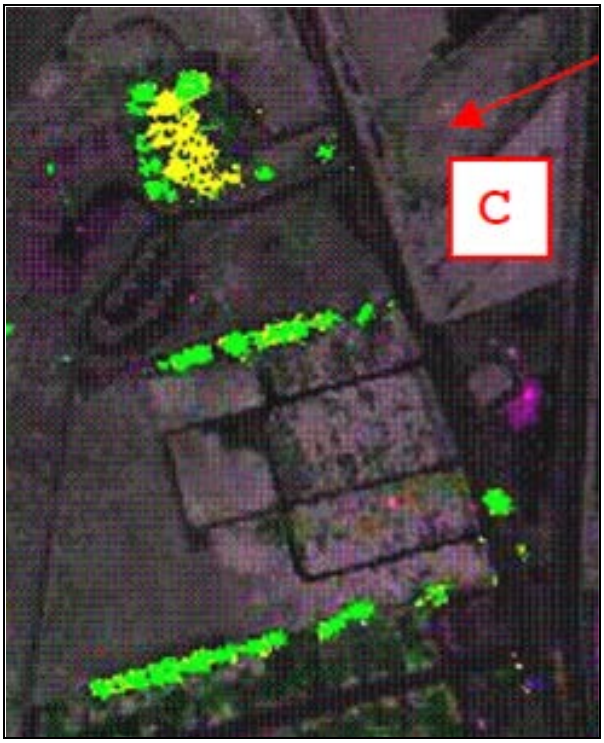

Figure 4. Results from Elkhorn Slough LiDAR classification: yellow-areas characteristics of Eucalyptus; green-areas with characteristics of Monterey Cyprus (After Helt, 2005).

B. SPECTRAL IMAGING

1. Spectral Fundamentals

Electro-optical sensors are a type of optical sensor that passively collects spectral radiance from a scene. The common types of EO sensors are panchromatic, multispectral, and hyperspectral. For remote sensing purposes, these sensors are deployed on an aircraft or satellite. Multispectral imaging sensors usually contain less than about 20 distinct spectral bands measuring energy at a few wavelengths. Hyperspectral imaging usually have hundreds of bands, which create a contiguous spectrum that can be formed into a hypercube. Although hyperspectral sensors are capable of excellent spectral resolution, they usually suffer from poorer spatial resolutions than their multispectral counterparts (Stein et al., 2002).

Hyperspectral imaging sensors are also known as imaging spectrometers. One such sensor is AVIRIS (The Airborne Visible/Infrared Imaging Spectrometer) which is flown by The National Aeronautics and Space Admiration's Jet Propulsion Laboratory (NASA JPL). This sensor has a spectral resolution of 10 nanometers covering the 0.4 to 2.5 micrometer range in 224 spectral bands. Multispectral imaging provides synoptic spatial coverage but does not allow for the same precision of identification. Figure 5 shows the spectral resolution differences between a MSI and HSI sensor (Kruse, 2007).

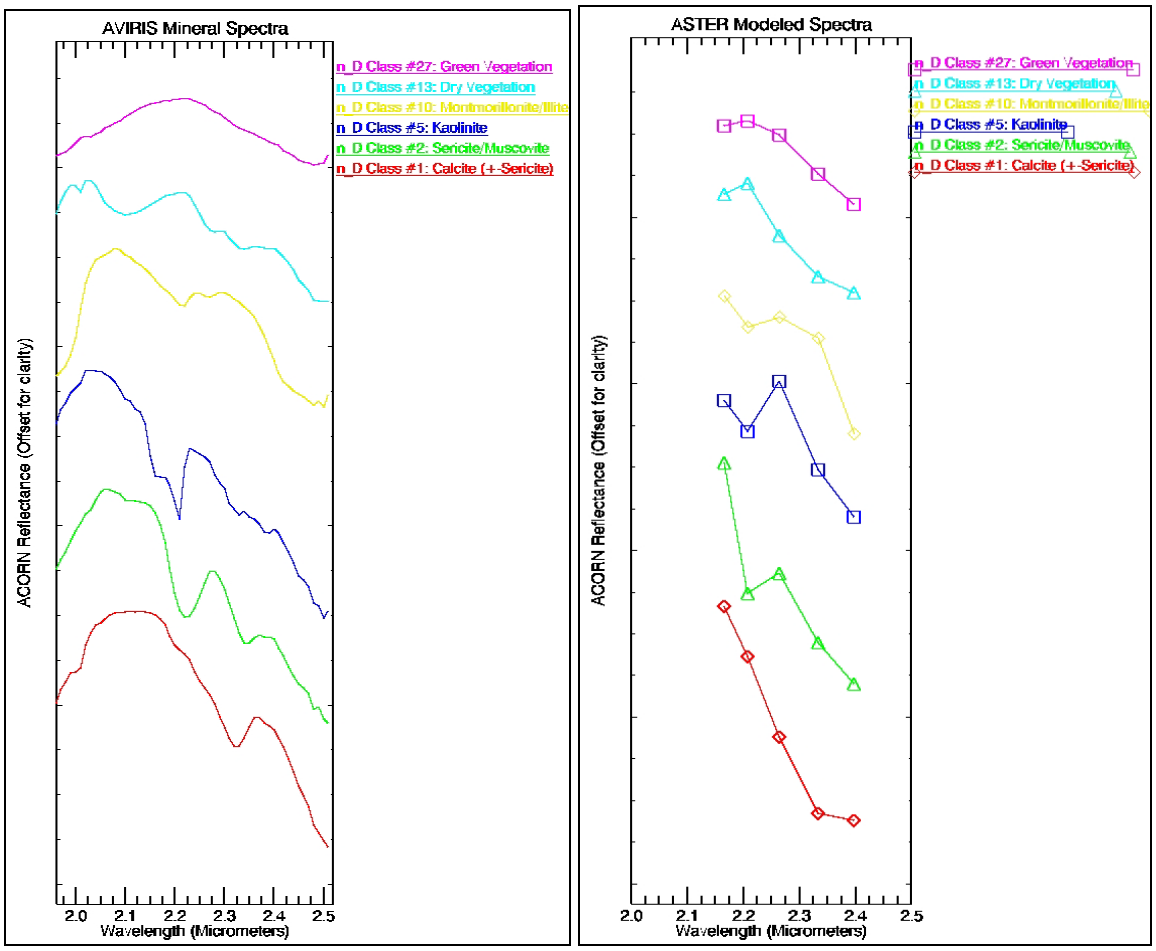

Figure 5. Comparison of AVIRIS (left) hyperspectral spectra and ASTER (right) multispectral spectra for selected minerals, dry, and green vegetation (From Kruse, 2007).

2. Multispectral Classification

This section describes some of the multispectral satellites in use today as well as some of the classification methods used with the data that they provide. The systems discussed are Landsat, IKONOS, and WorldView-2. This section also discusses two multispectral techniques used for this project: the normalized difference vegetation index and maximum likelihood classification.

a. Landsat

The Landsat program began in the early 1970s as an earth observing program for use in applications such as agriculture, geology, and forestry. The first

satellite was originally named the Earth Resources Technology Satellite (ERTS) and was a joint effort between the U.S. Geological Survey and the National Aeronautics and Space Administration. Out of the seven satellites that have been launched in the program, Landsat 5 and Landsat 7 are the systems which remain operational. Landsat 7 has eight spectral bands with varying spatial resolution from 15 meters (panchromatic), 30 meters (multispectral), 60 meters (long-wave infrared), and 90 meters (thermal infrared). See Figure 6 for a timeline of Landsat imaging and Figure 7 for Landsat 7 spectral band ranges (USGS, 2012).

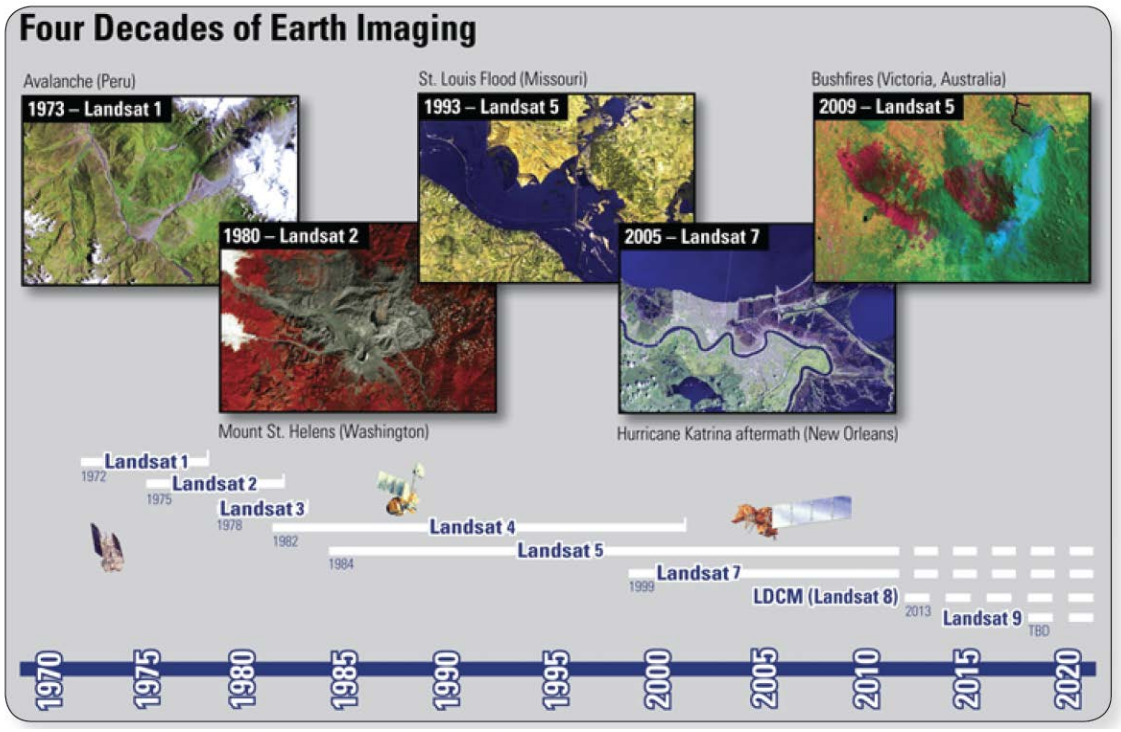

Figure 6. Landsat timeline and imaging samples (From USGS, 2012)

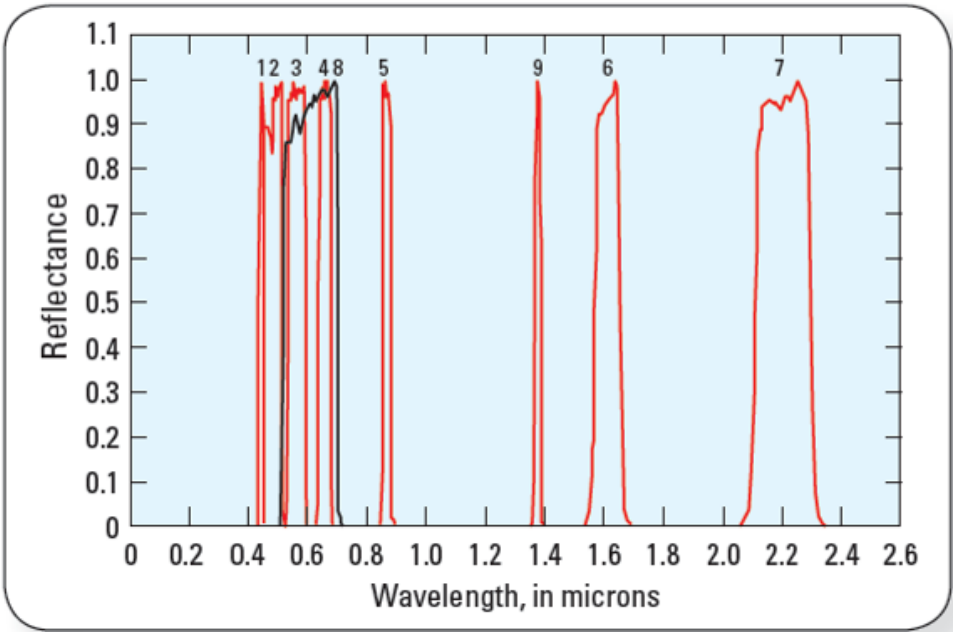

Figure 7. Landsat 7 spectral band ranges (From USGS, 2012)

In a study by the University of Minnesota, land cover classification and change were analyzed utilizing Landsat imagery around the Twin Cities area in Minnesota. They used data from the Landsat Thematic Mapper for 1986, 1991, 1998, and 2002. A hybrid supervised-unsupervised classification technique was developed that clustered the data into subclasses then applied the maximum likelihood classifier. Their results showed that urban land development increased from 23.7% to 32.8% while rural land types decreased from 69.6% to 60.5%. Figure 8 shows the change detected from the four classification maps created (Yuan et al., 2005).

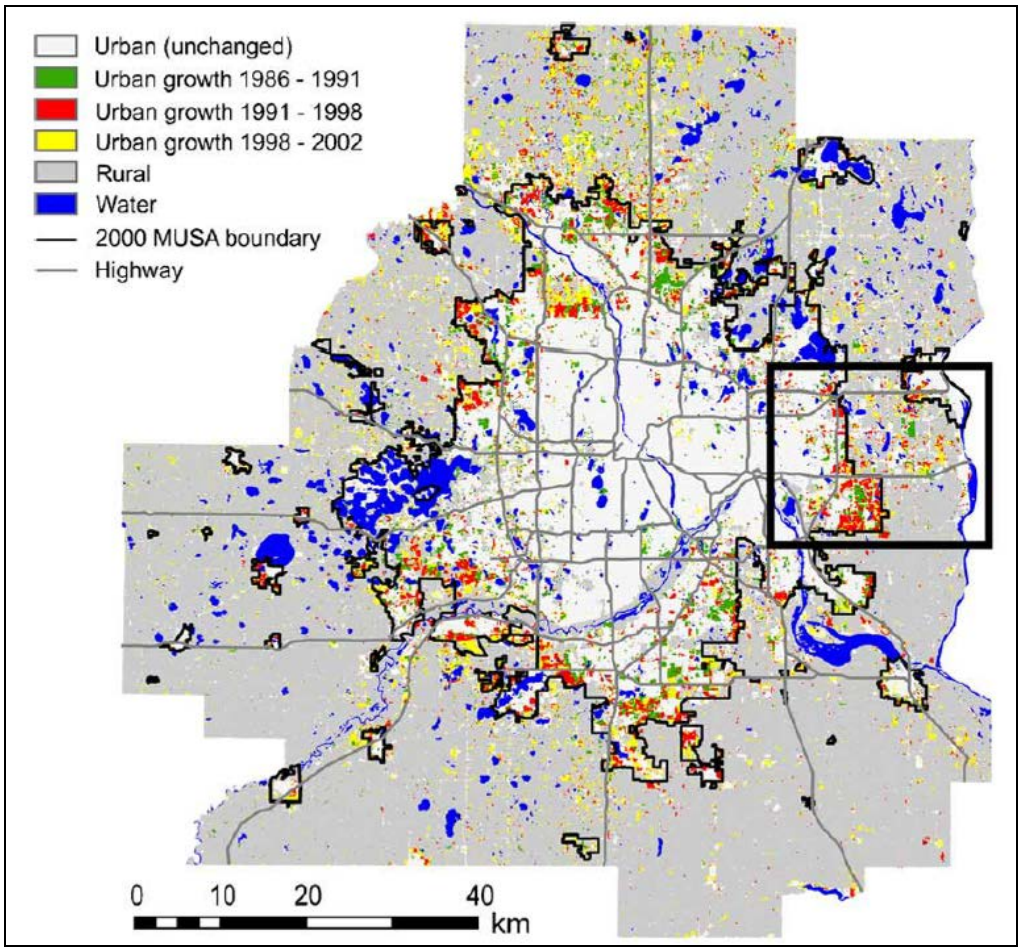

Figure 8. Land cover changes from Landsat in the Twin Cities from 1986 to 2002 (From Yuan et al., 2005)

b. IKONOS

IKONOS is a satellite system launched by the commercial company GeoEye. It was launched in 1999 and was the first satellite launched to offer sub-meter panchromatic images. Optimal spatial resolution is 0.82 meter (panchromatic) and 3.28 meter (multispectral). It orbits at an altitude of 423 miles and has a revisit time of three days, with downlinks to multiple ground stations. It has applications from military intelligence to community mapping and has been used in stereo imaging and environmental monitoring. Figure 9 shows the IKONOS spectral response and Figure 10 shows an example of a stereo pair collection (Dial & Grodecki, 2003).

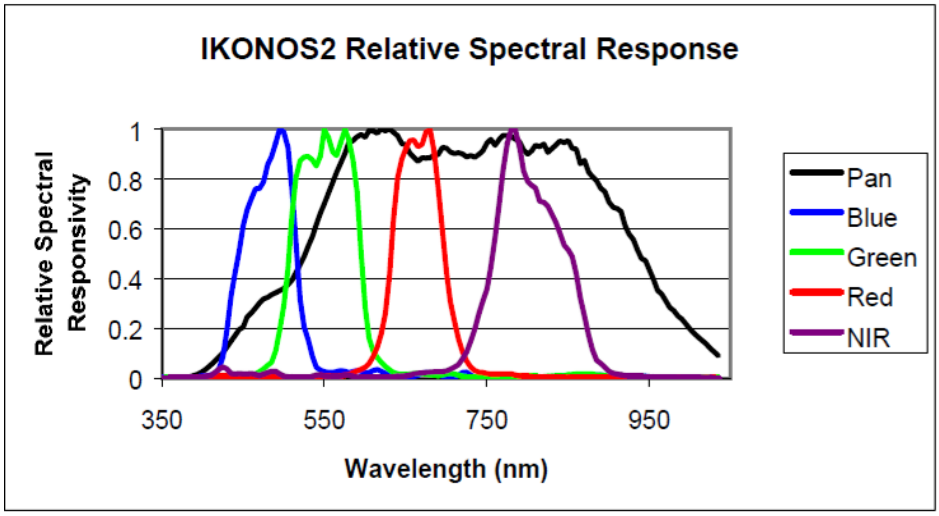

Figure 9. IKONOS spectral response bands (From Dial & Grodecki, 2003).

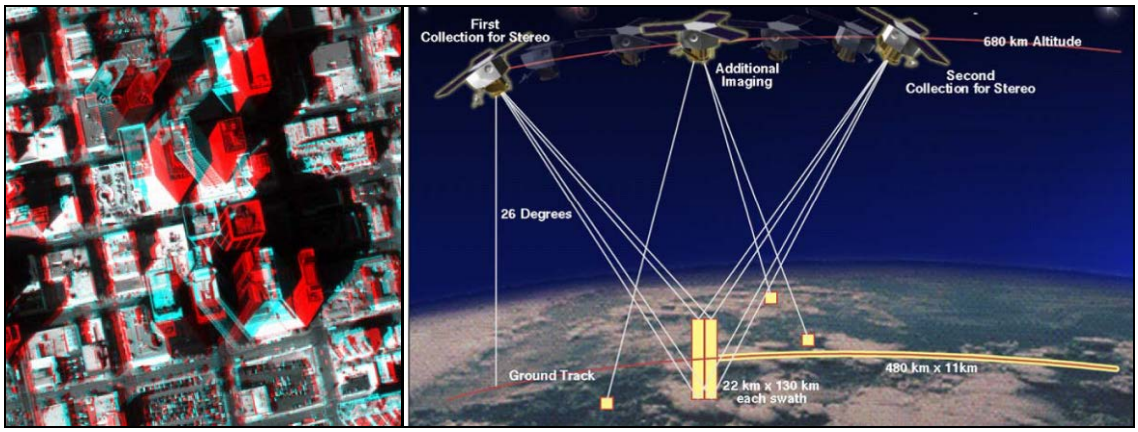

Figure 10. An example of an IKONOS visualization of a stereo pair and the satellite pass to obtain it (From Dial & Grodecki, 2003).

The forest area around Flanders, Belgium was analyzed by Ghent University utilizing IKONOS imagery and object-based classification. Their algorithm divided features into three categories of features: spectral type, shape, and texture. It was a three step process that involved image segmentation, feature selection by genetic algorithms, and joint neural network based object classification. The project was initiated to show the potential of their techniques when there was a limited set of training data. The project was also demonstrated as a way to update the Flemish Forest Map with a regularly operational method. Figure 11 shows one of their results next to the current

Flemish Forest Map with forest areas marked and their results. They showed significantly higher classification accuracy when compared to a strategy without feature selection and joint network output.

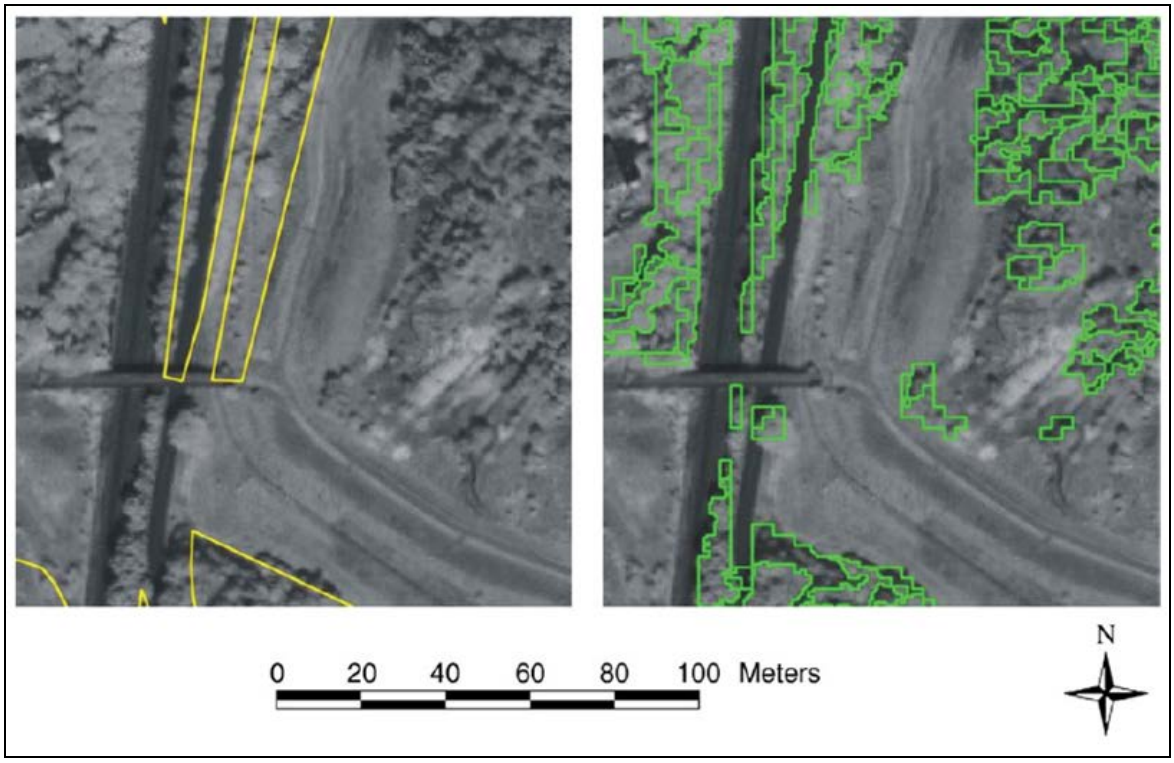

Figure 11. Forest mapping results from IKONOS over Flanders, Belgium: left-Flemish Forest Map forest cover in yellow outline; right-genetic algorithm forest over in green outline (From Coillie et al., 2005)

c. WorldView-2

The spectral imagery used in this project was obtained by the WorldView-2 satellite, operated commercially by DigitalGlobe. The system was launched on October 8, 2009. WorldView-2 has a panchromatic resolution of 46 centimeters, a swath width of 16.4 kilometers at nadir, and an average revisit period of 1.1 days. The satellite orbits at an altitude of 770 kilometers and can collect 975,000 square kilometers a day. WorldView-2 has 8 multispectral bands and is the highest commercially available at the time of this writing. The bands include a coastal (400–450 nm), blue (450- 510 nm), green (510- 580 nm), yellow (585–625 nm), red (630–690 nm), red edge (705–745 nm), near infrared (770–895 nm), and near infrared 2 (860–1040 nm). The 8 bands give

WorldView-2 imagery an advantage over other MSI systems as their additional bands can lead to more specific classification and feature extraction results. Please see Figure 12 for the spectral band locations of WorldView-2 and Figure 13 for spectral radiance response of the system (DigitalGlobe, 2011).

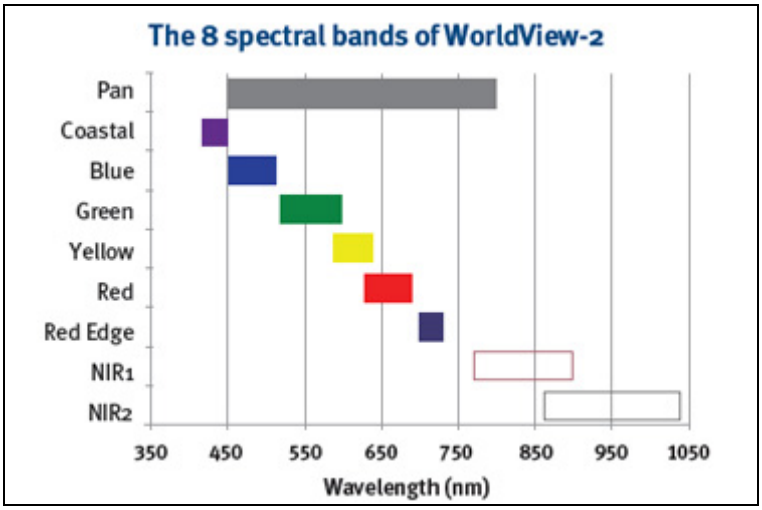

Figure 12. The wavelength ranges of WorldView-2 (From DigitalGlobe, 2011)

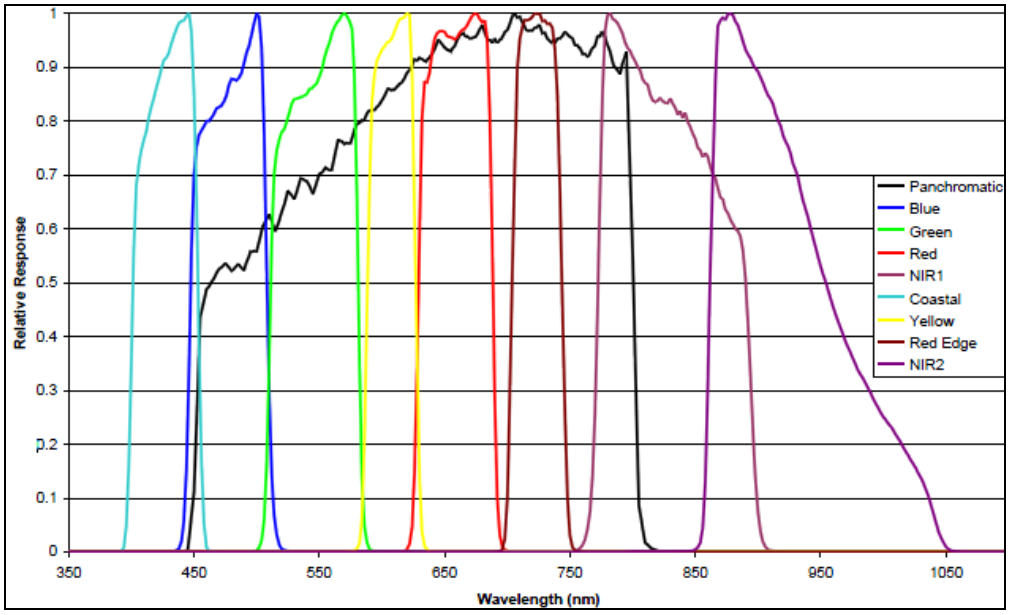

Figure 13. The Relative Spectral Radiance Response of WorldView-2 (From DigitalGlobe, 2011)

d. NDVI

NDVI stands for the Normalized Difference Vegetation Index. It is calculated from the red and near infrared values. The equation is as follows:

$$NDVI = \frac{NIR - \text{Red}}{NIR + \text{Red}}$$

Equation 1: Normalized Difference Vegetation Index

Like other materials, radiation emitted onto leaves can be absorbed or scattered as a function of wavelength. Green leaves absorb most of the radiation in the visible from 0.4 to 0.7 microns and reflects most of the near infrared from 0.7 to 1.05 microns. Vegetation also has a strong red absorption band from 0.62 to 0.68 microns which has been correlated with biomass production. This reflectivity in the near infrared increases with increased photosynthetic activity. NDVI is a good indicator of vegetation. NDVI values range from -1.0 to +1.0 with typical healthy vegetation ranging from 0.2 to 0.8 (Santos & Negri, 1996).

e. Maximum Likelihood Classification

Maximum Likelihood classification is one of the major tools for classifying pixels in a spectral image. It is a supervised technique that requires training pixels which are used define each classification. The classifier is based on multivariate normal distribution theory and works to find the maximum for a given statistic. It assumes a normal distribution in each class. In normal distributions, the likelihood function $P(x|k)$ can be expressed as:

$$L_k(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\sum_k|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_k)^T \sum_k^{-1} (x-\mu_k) \right)$$

Equation 2: Likelihood function from maximum likelihood classifier

Where x is the vector of a pixel with *n* bands and $L_k(x)$ is the likelihood memberships function of x belonging to class *k*. Figure 14 shows an example of the maximum likelihood classification applied to a Landsat image (Liu et al., 2010).

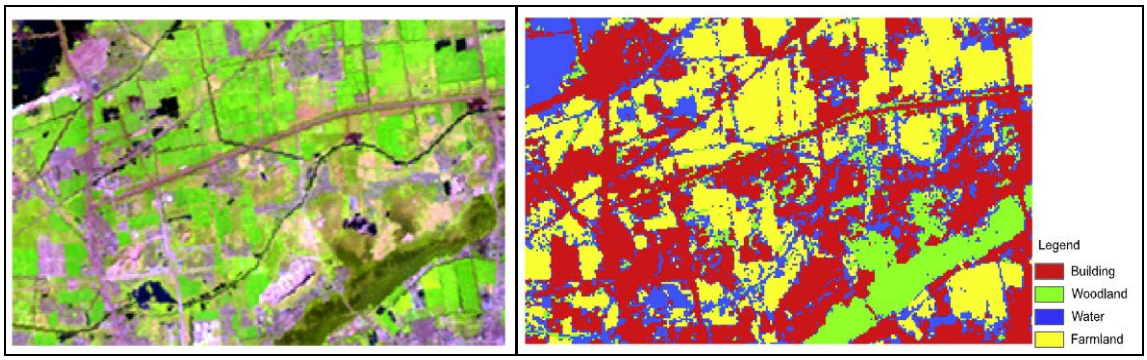

Figure 14. Sample of a Maximum Likelihood Classifier (From Liu et al., 2010)

C. MULTI-SOURCE FUSION LITERATURE REVIEW

There have been many different approaches to analyzing the fusion of LiDAR and spectral data. Some approaches utilized multispectral imagery and others utilized hyperspectral imagery. This section takes a look at previous work done in the field in natural and urban environments.

1. Fusion Vegetation Analysis

In a joint study conducted by members of the University of Maryland, the University of California, and the Goddard Space Flight Center, LiDAR and hyperspectral data fusion was examined to observe biomass and stress in the Sierra Nevada. Waveform LiDAR data were collected by the Laser Vegetation Imaging Sensor (LVIS) and hyperspectral data were collected by AVIRIS. HSI image spectral endmembers were collected from green vegetation, non-photosynthetic, vegetation, soil, and shade. LVIS metrics, AVIRIS spectral indices, and their endmembers were analyzed. A correlation was found between shade fractions and LVIS calculated canopy height. Their study showed that biomass errors found with fusion and without fusion were different, but not statistically significant, particularly amongst hardwood trees and pine trees. It was found that the confidence intervals were narrowed with the fusion method relative to the individual data analyses. Overall, LiDAR was better suited for biomass estimation, with hyperspectral imagery used to refine predictions and determine canopy state and stress. Figure 15 shows the results of their project as a tilted 3D model (Swatantran et al., 2011).

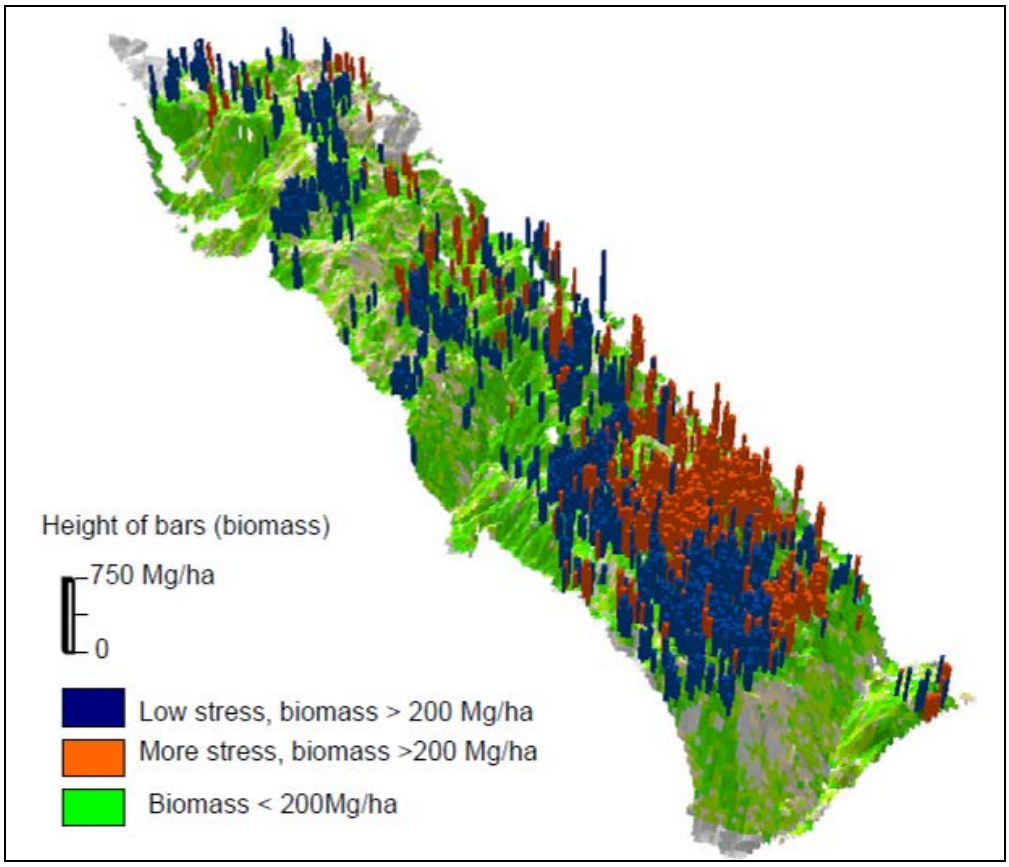

Figure 15. The fused results of biomass calculations in the Sierra Nevada (From Swatantran et al., 2011)

2. Fusion and Shadowed Features

A paper from the Rochester Institute of Technology analyzed how to leverage LiDAR data to aid in hyperspectral target detection (Ientilucci, 2012). They analyzed how illuminations can be obtained by processing LiDAR to estimate varying illumination of targets within a scene. The data they used were from the SpecTIR Hyperspectral Airborne Rochester Experiment (SHARE) program tested over Rochester, New York. The study showed how the spectrum of a blue felt target panel varied slightly because of background but was significantly altered and reduced when shaded. They performed a match filter detection algorithm and showed that the shaded spectrum was not just a magnitude change but actually made the material look spectrally different to the sensor. They created a forward physics based model with LiDAR data, that when used as a match filter found their targets in both shaded areas and in the open. Many improvements still

need to be made, however, as the process was not able to detect all targets in a single pass. Figure 16 shows some of the LiDAR processing that was done in order to automate shadow detection (Ientilucci, 2012).

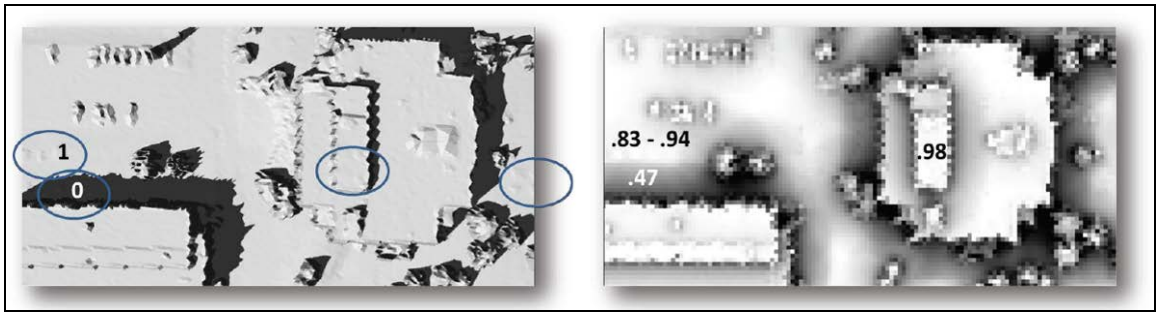

Figure 16. Shadow map (left) and an illumination map (right) created from LiDAR images at the Rochester Institute of Technology (From Ientiluccci, 2012)

3. Fusion Feature Detection

The National Geospatial-Intelligence Agency (NGA) performed a study over Kandahar, Afghanistan to use multi-source fusion to create 2D and 2.5D data to portray the dynamic urban landscape. Their study indicated that nearly 15% of the buildings required vegetation detection in order to be successfully validated. Their study also analyzed temporal change detection at the object level and addressed issues involving building features such as balconies, TV dishes, domes, and other attributes. The study included NGA's Urban Feature Data (UFD) vector information, LiDAR from the U.S. Army's Buckeye collection system, and WorldView-2 multispectral imagery. Temporal change detection was possible, as they had multiple Buckeye collections spaced six weeks apart. Using a combination of geometric analysis and NDVI calculations, they were able to create a framework to maintain a database to validate and update 3D urban features using the tools of the military and sensor communities. Figure 17 shows a sample of 3D temporal change detection that was observed (Arrington et al., 2012).

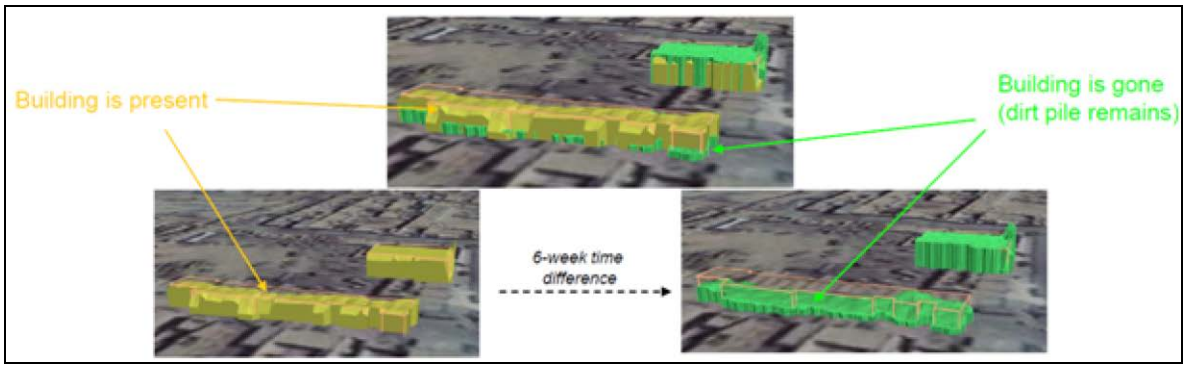

Figure 17. Temporal building changes in Kandahar, Afghanistan
(From Arrington et al., 2012)

A project by the University College London used fusion data to improve methods of building extraction to achieve higher levels of accuracy and quality by using height and geometry information in conjunction with NDVI indices of the area. They utilized a tool called the Binary Space Partitioning (BSP) tree which merged convex polygons and divided extracted lines to create full building outlines. The analysis utilized pan-sharpened multi-spectral imagery from IKONOS in conjunction with LiDAR. Their study area was a subset of an industrial area in the Royal Borough of Greenwich, London, United Kingdom. The process was a two-step procedure that included building detection and description; first detecting dominant features and then isolating them from the background. They compared their results with the Ordnance Survey and rated their accuracy at 90.1%. In the error analysis, they predicted that false positives and false negatives could be reduced with a more evenly distributed point cloud at a higher density. Figure 18 shows the result of their building extraction process and comparisons to their sources (Sohn & Dowman, 2007).

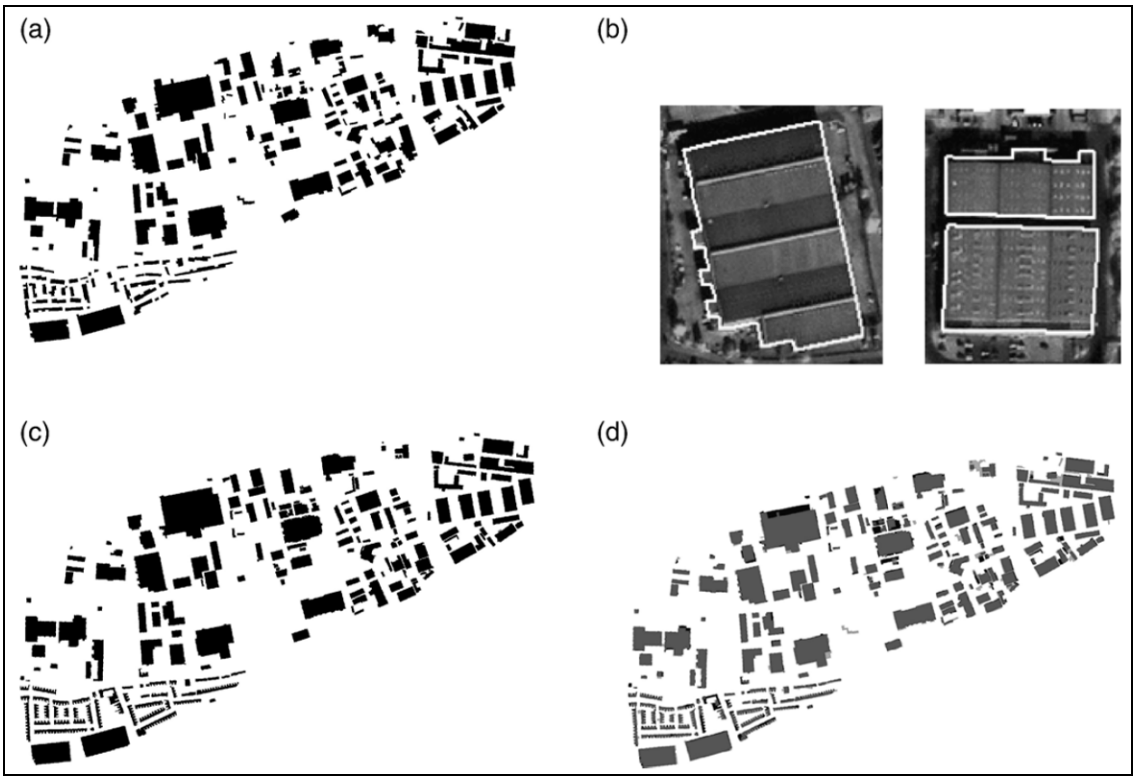

Figure 18. Building Extraction in Greenwich: (a) Building Map; (b) extraction results subset; (c) Ordinance Survey; (d) extraction errors (light grey: true positives; dark grey: false positives; false negatives)
(From Sohn & Dowman, 2007)

A study at the Naval Postgraduate School looked at the fusion of LiDAR and spectral data using methods that would be meaningful to city planners and emergency responders (Kim et al., 2012). Their research goal was to detect building rooftops, which in turn detected building footprints. The study area was over Monterey, California, and utilized LiDAR collected from the ALTM (Airborne Laser Terrain Mapper) Gemeni system and spectral data from WorldView-2. The process involved a series of extractions, masks, and exceptions. With the LiDAR data, statistics were found on local neighborhoods and flat surfaces were extracted from the rest of the background. LiDAR based masks were then used to differentiate points that were considered ground and points that were considered vegetation based on multiple returns. Exclusions occurred based on area sizes. Areas less than ten square meters were likely false alarms and areas larger than thirty square meters were likely highways. NDVI was calculated using the

spectral image. An NDVI threshold of 0.35 and higher mapped healthy vegetation, which could then be removed. The results were an effective method for extracting rooftops based on LiDAR/MSI fusion, which would be difficult without both data sets. Figure 19 shows some of the results from their process (Kim et al., 2012).

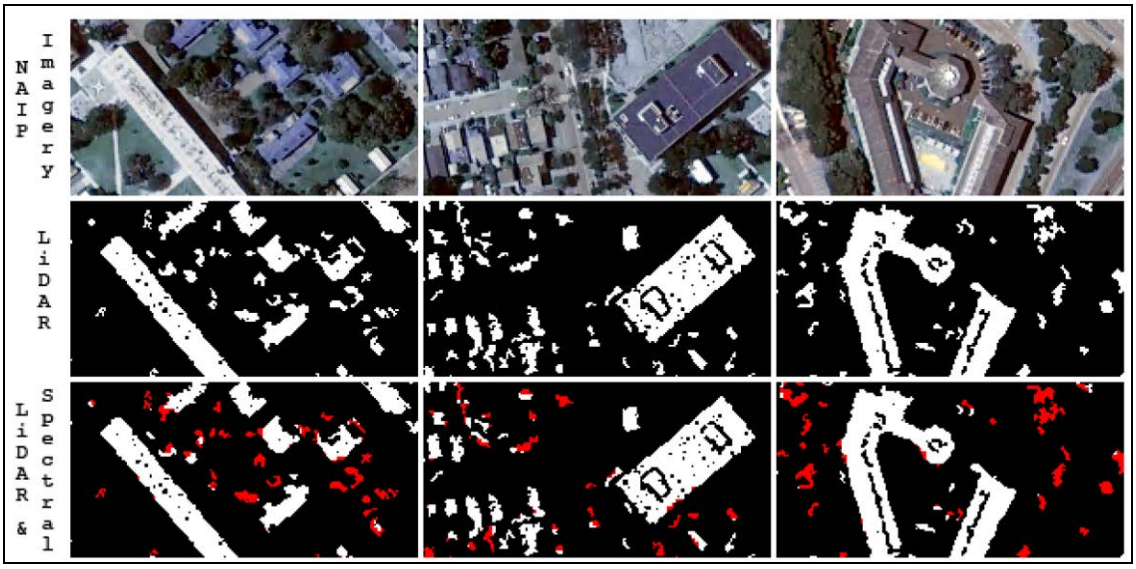

Figure 19. Rooftop extraction results in Monterey, CA; the bottom row shows fused (LiDAR and WV-2) extraction results in white with red showing false alarms from the LiDAR only extraction (From Kim et al., 2012)

D. THEORY

The core of this research utilized a rule based classifier as a type of decision tree. Decision trees form a multistage or hierarchical decision scheme similar to the branches of a tree. They begin with a root of all the data and branch to internal nodes that create a series of splits that end up at terminal nodes. Each of the nodes is a binary decision that sets it as one class or keeps in in the remaining classes, eventually moving through the tree to the end nodes. Rather than creating one complex decision, the decision tree technique breaks it down into a series of simpler choices (Xu et al., 2005).

The approach to this thesis combined some of the previous efforts' techniques in deriving data through the point cloud and multispectral image, creating nodes in the form of masks. Some of the other works focused on detecting specific target types using fusion. This project focused on combining the efforts in order to classify an entire urban

scene as best as possible. While spectral signatures generally do well at identifying materials, fusion techniques impose many more requirements that need to be met before a pixel is classified as a particular material.

One of the consistent themes from the literature review was the need to differentiate vegetation from non-vegetation. Spectral differentiation of these is important, as some vegetation and man-made objects can appear geometrically similar. NDVI was calculated for this process to determine vegetation, and it was masked early in the tree process.

Other masks were derived from LiDAR. Distinctions were made via number of returns, above ground level, and intensity as well as also utilizing some of the pre-set LiDAR classifications provided by the vendor.

The terminal nodes were created through mask combinations, more specific types of material were isolated and regions of interest were dedicated to those subdivisions. The images were then classified using the maximum likelihood classifier. The results of the classifiers were multiple classified images with masked out areas. In order to create a complete image, these sets were then compiled together.

E. STUDY AREA: SAN FRANCISCO, CALIFORNIA

The study area for this project was San Francisco California. San Francisco is located in northern California near where the Pacific Ocean meets the San Francisco Bay and Golden Gate strait. It is situated at about North 37.759880 latitude and West 122.437393 longitude. The area of the city is about 47 square miles with a population density of about 17,200 persons per square mile. The population was estimated at about 813,000 in 2011 (U.S. Census, 2010).

Because of its unique location, San Francisco is an ideal location for this project. The area features everything from beaches and parks to large bridges and skyscrapers all in relatively close proximity to each other. This diverse mix of urban and natural landscapes was beneficial for assessing the effects of LiDAR and multispectral fusion. Figure 20 shows San Francisco County in relation to the rest of California.

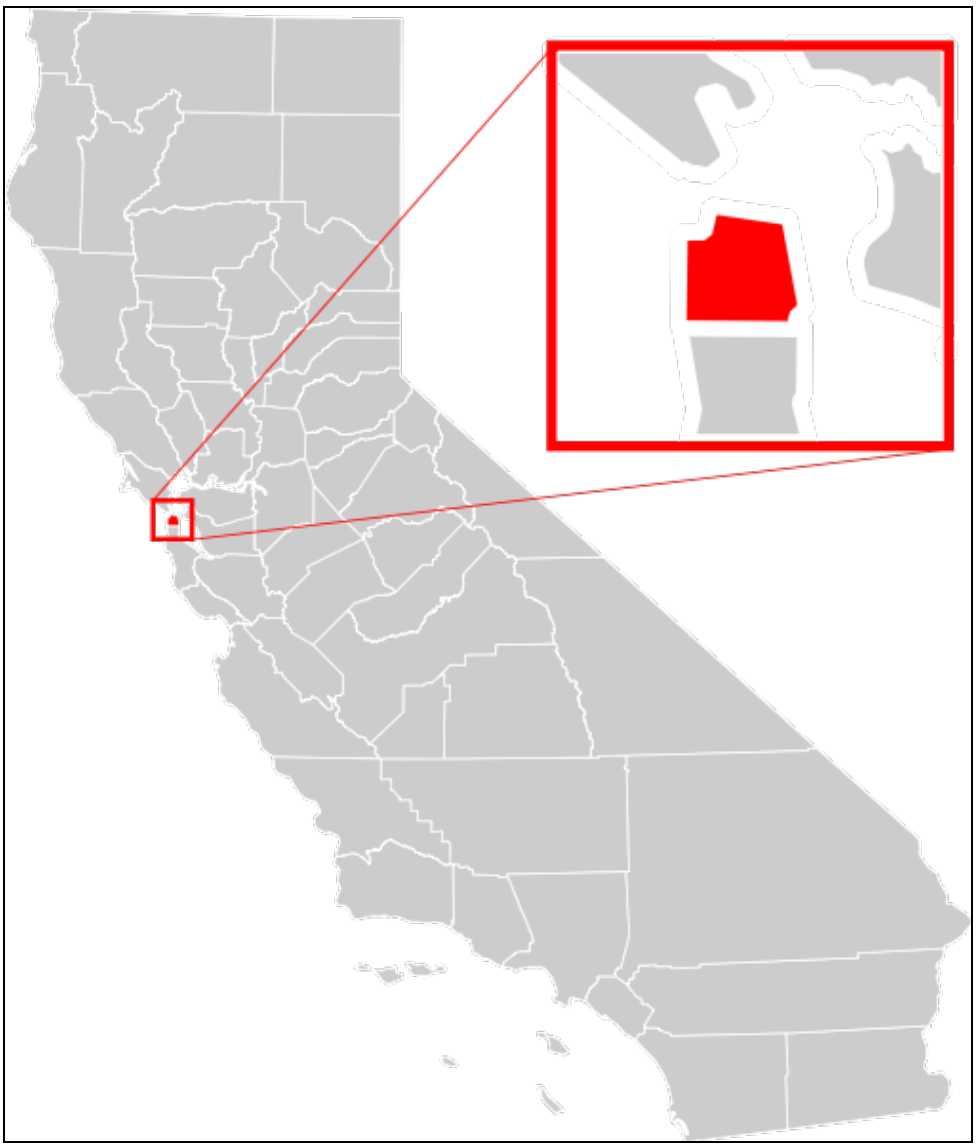

Figure 20. San Francisco County (Red) inset with the state of California (Gray) (From *Wikimedia Commons*, 2008).

THIS PAGE INTENTIONALLY LEFT BLANK

III. PROBLEM

A. OVERVIEW

The main problem addressed in this thesis was to evaluate combined classification techniques of LiDAR and multispectral data, maximizing accuracy and minimizing misclassification. The fusion techniques explored here attempt to preserve the grid and network created by human roads and buildings while still being able to spectrally classify the area of interest.

B. DATA SET AND COLLECTION METHODS

1. Golden Gate LiDAR Project

The LiDAR data used in this project comes from The Golden Gate LiDAR Project. The project collected LiDAR data, aerial photography, and hyperspectral imagery. At the time of this writing, the hyperspectral data were not available. The project collected data in Northern California and collected information on 835 square miles of Marin County, San Mateo County, Sonoma County, and San Francisco County. See Figure 21 for collection area. The flights were completed between April 23, 2010 and July 14, 2010 utilizing a Cessna 207 aircraft. The LiDAR system used was a Leica ALS60 MPiA (multi-pulse in air). The system collected multiple returns in X, Y, Z, as well as pulse intensity and full waveform data. Points were collected at a density of about 2 points per square meter with a 15% side lap in a 28 degree field of view. A network of ground control stations were used during the flights using a Trimble R7 with a Zephyr geodetic model 1 antenna. Flights were also coordinated to collect during the lowest tides possible. In order to achieve the best data collection, criteria included a low PDOP (Positional Dilution of Precision) of less than 2, a baseline no greater than 25 miles, a constant slope, and observation at moderate intensities (Hines, 2011).

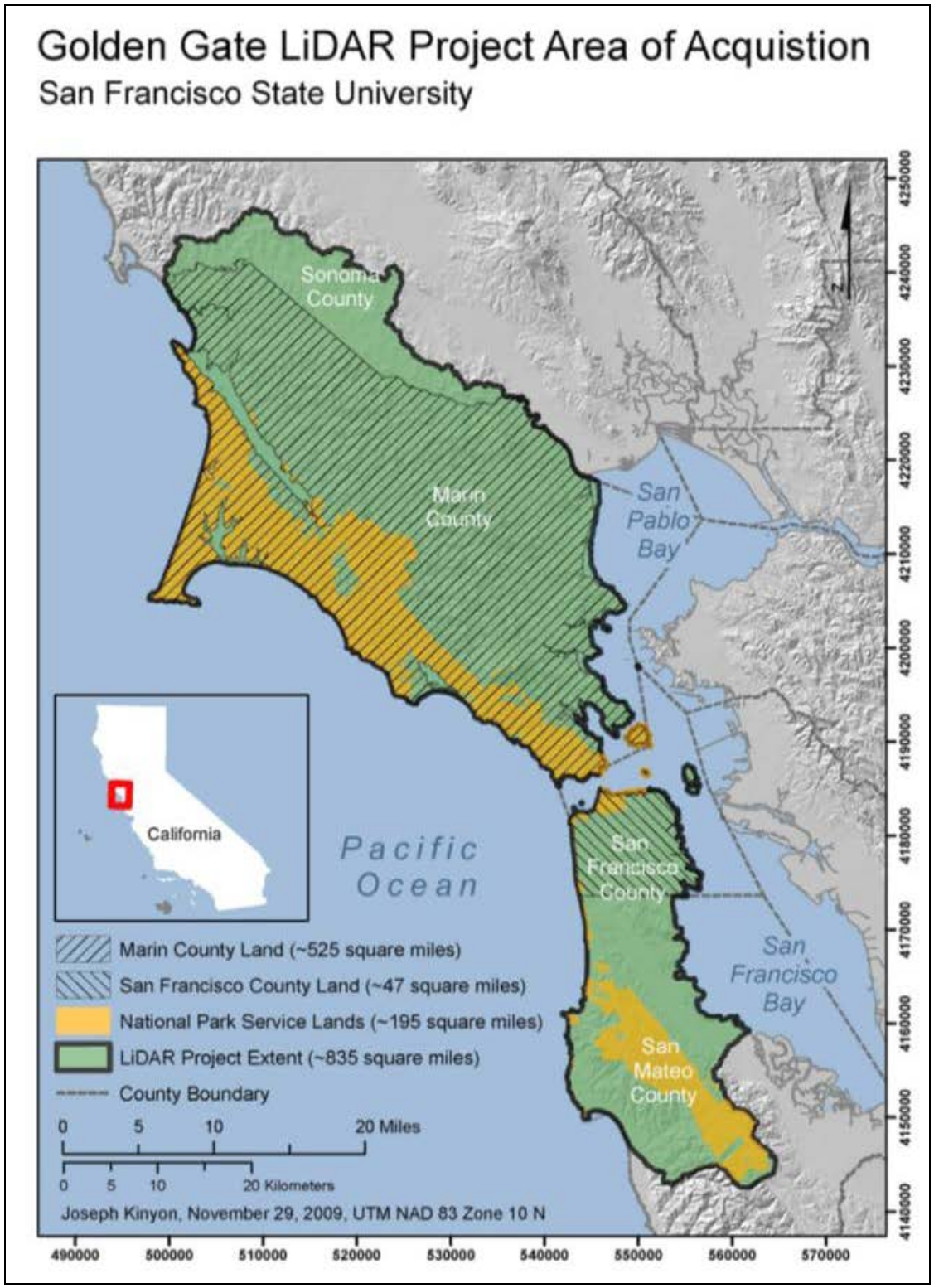

Figure 21. Golden Gate LiDAR Project acquisition area (From Hines, 2009)

The raw LiDAR data were initially processed by Earth Eye LLC, and further processed by The GGLP group at San Francisco State University. Calibration was achieved using information from GPS and IMU collects as well as attuned to sensor and flight line data. The points were auto-classified with algorithms that consider slope, angular, relationships, and distance which defined 95% of the project area. Further reclassification was done on more than 10% of the points with further manual inspection of the points. The resulting points were classified as follows:

- 1 - Processed, but unclassified
- 2 - Bare-earth, ground
- 4 - Vegetation, all above-ground objects including buildings, bridges, piers
- 7 - Noise
- 9 - Water

The LiDAR data was assessed at a vertical accuracy root mean square error of less than 9.25 cm. The delivered product is displayed in the UTM (Universal Transverse Mercator) coordinate system, with units in meters, in zone 10 north, with horizontal datum NAD83 (North American Datum of 1983), and vertical datum NGVD88 (North American Vertical Datum of 1988). Each tile is 1500 x 1500 meters and delivered as LAS (Laser File Format) v1.2 and v1.3 that included waveform. For this project, the LAS v1.2 tiles were utilized. See Figure 22 for a sample of the processed point cloud.

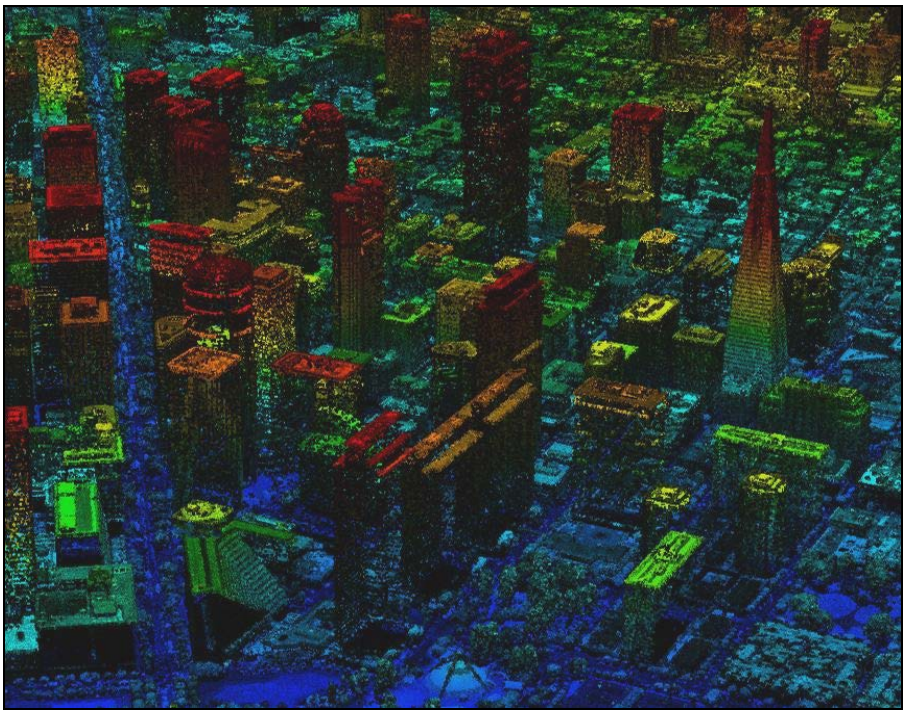

Figure 22. A sample of the GGLP point cloud over downtown San Francisco viewed in Quick Terrain Modeler

2. WorldView-2

The image used in this project was collected by WorldView-2 on November 8, 2011 at Zulu time 19:34:42 (11:34 AM, local Pacific Time). The image in centered on San Francisco County. In order to limit the amount of perceptual layover that is caused by the taller buildings, the image was chosen at a very close to nadir viewing angle of 15 degrees. This image was also chosen because it had very low cloud cover of about 1%. The sun elevation at the time of the image acquisition was 35.58 degrees, which does create longer shadows than a directly overhead sun. The image was cataloged under the name: 11NOV08193442-M2AS-052753574130_01_P002. The raw image was delivered in TIF format as DigitalGlobe's Standard 2A product type. The image had 2-meter square pixels and was projected in UTM, Zone 10 N with the WGS-84 (World Geodetic System of 1984) datum. See Figure 23 for an overview of the multispectral image.

Figure 23. The WorldView-2 multispectral image of San Francisco in true color

3. Subset based on LiDAR

The thesis focuses on the urban areas of San Francisco County. The chosen area consists of 25 LiDAR tiles in the northeast sector of San Francisco. The area was chosen because it included all of downtown, a portion of The Bay Bridge, coastal areas, piers, part of Golden Gate Park, commercial areas, and suburban areas. The area was selected as a good composition of typical urban features of larger metropolitan areas but is still manageable by a typical personal desktop computer. See Figure 24 for a map layout of the selected area.

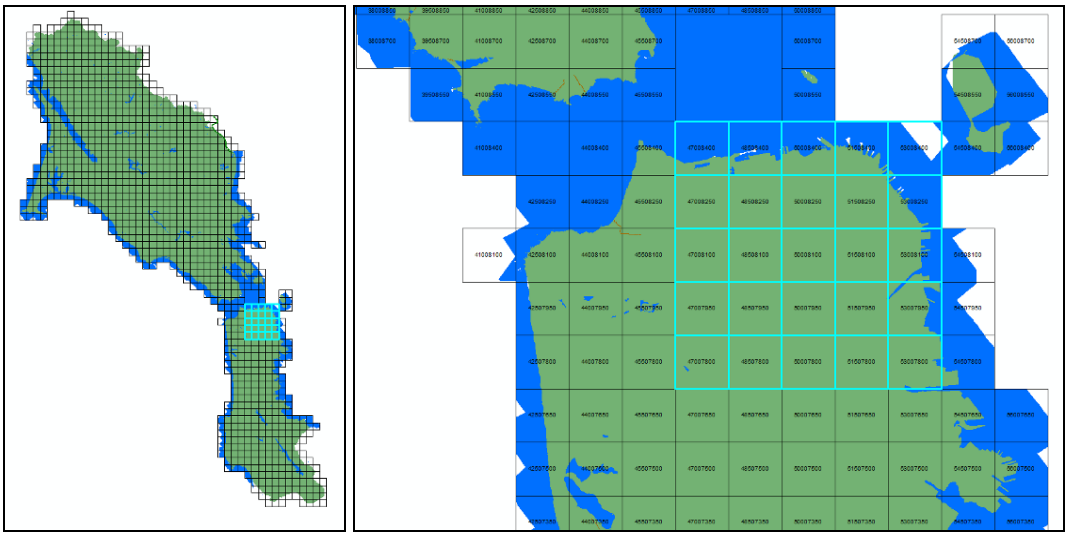

Figure 24. The area of interest as indicated by the cyan outlined tiles: left-full coverage region; right-San Francisco County study area

In order to perform fusion analytics between the multispectral and LiDAR sets, the information between the two must be aligned properly so as to not offset anything nor introduce noise into either image.

As the multispectral image is the basis of spectral classification, the masks and DEM created from the LiDAR data were matched and projected to the same UTM map projection and datum as the WorldView-2 image. Because it is more difficult to manipulate the actual points of the point cloud, the WorldView-2 image was orthorectified and cropped to match the LiDAR generated Digital Elevation Model. This is explained further in the Methods section of this thesis.

C. SOFTWARE USED

1. Quick Terrain Modeler 7.1.5

Quick Terrain Modeler (QTM) is a 3D visualization software package created by Applied Imagery and designed for use with LiDAR data. The software is used by many organizations within The Department of Defense including the U.S. Army AGC Buckeye program and The National Geospatial-Intelligence Agency's IEC platform. It has the ability to bring in LAS tiles and create point cloud or surface models. It utilizes proprietary file formats called QTA (point cloud), QTC (un-gridded point cloud), and

QTT (gridded surface) but has the capability to export models into a variety of other formats such as GeoTIFF, LAS, ASCII, and shapefile. It also has a multiplicity of tools that can perform analysis such as flood assessment, helicopter landing zones, and line of sight (Applied Imagery, 2012).

2. E3De 3.0

E3De (Environment for 3D Exploitation) is a LiDAR tool created by Exelis visual Information Solutions (VIS). E3De has the ability to process point cloud information and quickly extract and identify 3D features for fusing into traditional 2D imagery. Extractions include orthophoto, Digital Elevation Model, Digital Surface Model, buildings, power lines, and trees, among others. It also has the ability to manually refine generated features to better match reality. Products can be exported as topographic, raster, .csv, GeoTIFF, LAS, SHP, and ENVI image formats (Exelis Vis, 2012).

3. ENVI 4.8

ENVI (The Environment for Visualizing Images) is a powerful imagery analysis tool created by Exelis VIS. ENVI is a robust image processing and analysis system that can work with many sources of imagery from airborne and satellite systems like AVIRIS, WorldView, and RadarSat. It has the ability to process different types of data such as multispectral, hyperspectral, polarimetric, radar, and some LiDAR data. It has built in tools allowing for tasks such as change detection, registration, orthorectification, and classification. It can also work in many formats such as HDF, CDF, GeoTIFF, and NITF. The program is customizable, with many users creating their own custom code in order to perform more specific tasks not previously built into the software suite. This project used ENVI for applying LiDAR derived masks to spectral imagery and then classifying the image (Exelis VIS, 2012).

4. IDL 8.0.1

IDL (Interactive Data Language) is a programming language used for data analysis and commonly used for image processing. IDL is the programming backbone of ENVI and the language in which custom ENVI code is written. IDL has a dynamic

variable typing system that is useful avoiding recompilation and prototyping change variables and values (Exelis Vis, 2012). For this project, custom IDL code was written to merge separate classified images into one and generate a random sample of points for ground truth analysis.

IV. METHODS AND OBSERVATIONS

A. PROCESS OVERVIEW

The focus of this thesis was to create a robust technique for fusing LiDAR and spectral imagery for creation of a more accurate classified image than MSI alone. Essentially this technique used LiDAR to create a series of masks. The multispectral image was used to create a vegetation mask. Through a mixture of mask combinations and classification, this technique constrained pixels to meet a number of requirements before designation of seven general classes. A maximum likelihood classifier was run against each general class using a limited number of regions of interests. The resulting classified images were then combined into one. It was expected that this rule based classification technique would create a more accurate classified image than LiDAR or multispectral on their own.

B. POINT CLOUD PROCESSING

The basis for this technique required information from the LiDAR to be extracted and used in a raster form that can be transformed into a mask. The study area was defined by the selected number of tiles and E3De and Quick Terrain Modeler were used to extract particular sets of information.

1. E3De – DEM, DSM, Intensity, AGL

E3De has the ability to bulk process LAS tiles and generate a number of products based upon built-in algorithms from the software. The tiles were imported into E3De and the projection was set to match the WorldView data: UTM, datum WGS84, meter, and zone 10N. Using E3De's processing tools, a digital elevation model, digital surface model, and an orthophoto product were selected to be generated. The orthophoto product utilizes intensity values and creates a raster intensity image with values between 0 and 255. Each product was set to have a resolution of 2 meters, also to match the WorldView data. For the DEM, a setting called Filter Lower Points was set to Urban Area Filtering. Default settings were used for the other options and the process was then run on the data.

The resulting products were created in ENVI raster and elevation formats with Z values representing height above sea level in meters.

Another product, known as the AGL or above ground level was derived from the DEM and DSM. This image was used to give z values based on height above the surface value rather than height from a set sea level. In order to create this image, both the DEM and DSM were loaded into ENVI as bands. Band Math was then utilized to do a pixel by pixel subtraction of the Digital Elevation Model from the Digital Surface Model. The result is an AGL image with digital number values representing meters above the ground level. Figure 25 shows a representation of each of these images with darker pixels indicating lower values and lighter pixels indicating higher values.

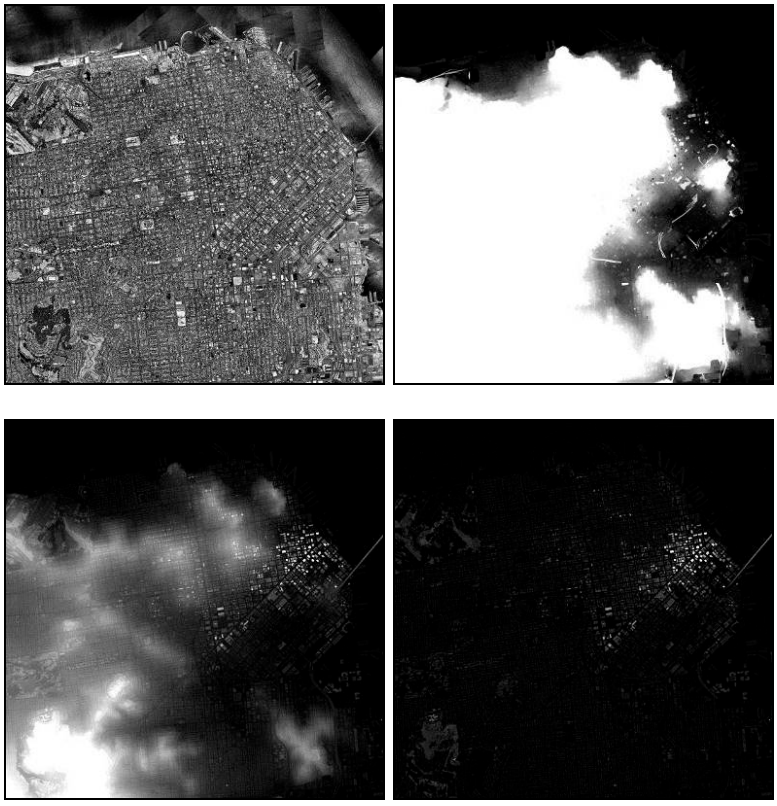

Figure 25. LiDAR E3De derived images: top left-intensity; top right-DEM; bottom left-DSM; bottom right-AGL

2. QTM – Classifications, Number of Returns

Two other types of LiDAR information were extracted from the LAS tiles, which include vendor-provided classification types and number of returns. Quick Terrain Modeler was utilized in order to create raster versions of these data. Single classification categories were loaded into QTM for both water and ground classifications. After the classification was loaded as a point cloud, it was then converted into a proprietary QTT surface model with simple interpolation smoothing and matched to the projection and resolution of the WorldView data. The QTT surface model was then exported as a GeoTIFF image, which can be utilized by ENVI for mask creation.

Quick Terrain Modeler also has the ability to remove features based on number of returns. All of the tiles were loaded into QTM and then analyzed utilizing the generate grid statistics tool. The number of returns variable was selected and metrics were calculated which could separate areas which received only one return or two or more returns. Points which only had one return were then removed from the loaded data via the filtering tool. The remaining points were exported as a GeoTIFF. This image was used to separate trees from grass. The multiple-return showed dense vegetation and also extracted building outlines. Figure 26 shows the two LiDAR classification images and the multiple-return image.

Figure 26. LiDAR QTM derived images: left-water class; center-ground class; right-multiple returns

C. MULTISPECTRAL PROCESSING

1. Conversion into Reflectance

The WorldView-2 imagery in this project was delivered by DigitalGlobe as a Standard Level 2A file. The image itself was in a raw state that displayed the collected intensities from the sensor. The image was first transformed into radiance. ENVI has a WorldView tool that allows for the process to be automated. The tool requires the *.IMD file, which includes metadata from the image that is used in the conversion, and the output is in floating point format to preserve data precision. The resulting spectrum from the radiance image resembles that of a solar spectrum. In order for reflectance conversion to run successfully, the radiance image was converted from an interleave type of band sequential (BSQ) to a band interleaved by line (BIL) type.

FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) atmospheric correction was used for this project. FLAASH is a widely used model developed by the Air Force Research Laboratory and its partner organizations. It removes atmospheric effects caused by aerosols and water vapor and creates an image in units of reflectance (Adler-Golden et at., 1999). ENVI has a FLAASH tool that requires the following inputs. Values listed which were acquired from the image metadata and regional characteristics of the scene:

- Scene Center: Sample: 4456, Line: 4009
- Scene Center: Latitude 37 44 36.84, Longitude -122 26 34.03
- Sensor Altitude: 770 km
- Ground Elevation: 0.0158 km
- Pixel Size: 2.0 m
- Flight Date: Nov 08 2011
- Flight Time: 19:34:42
- Atmospheric Model: Mid-Latitude Summer
- Aerosol Model: No Aerosol
- Aerosol Retrieval: None

The resulting FLAASH output image was a reflectance image that was spectrally corrected but not yet orthorectified and cropped to match the LiDAR data and masks.

Although visually, the conversions do not appear to make a significant change in the data, the spectral differences between the conversions are significant, and are displayed in Figure 27 of a sample of a grass vegetation spectrum from the images.

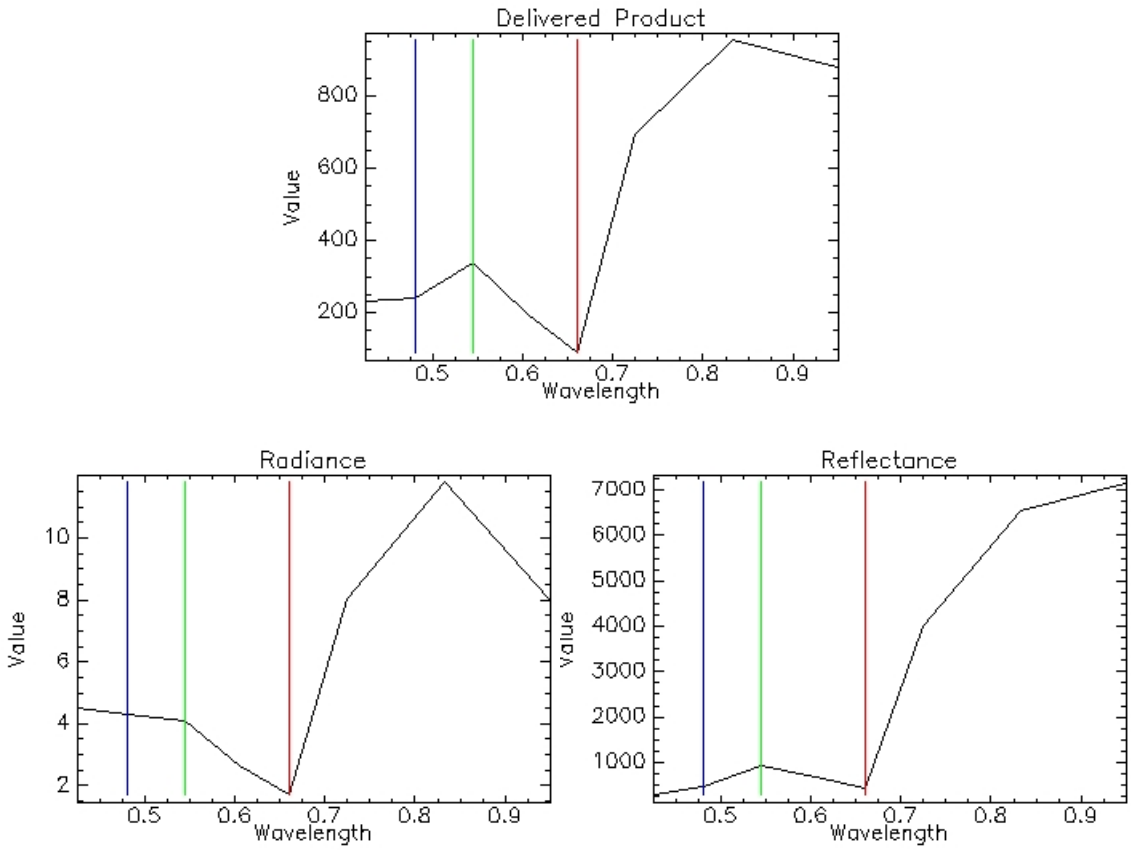

Figure 27. Spectral changes from raw WV-2 data to radiance and then reflectance

2. Orthorectification and Cropping to Subset

The ENVI orthorectification tool requires RPC coefficients and a Digital Elevation Model. The RPC coefficients were provided with the multispectral data as the *.RPB file. The DEM generated from the LiDAR was used in the processing. The setting to match an existing file was selected and the DEM was chosen. The result is an orthorectified reflectance cropped to match the LiDAR area of interest. Figure 28 shows the original data area and the cropped data area.

Note the no-data region in the northern part of the image. This region was not cropped so as to maintain the square LiDAR images, however when assessing the final classified images, this region was omitted.

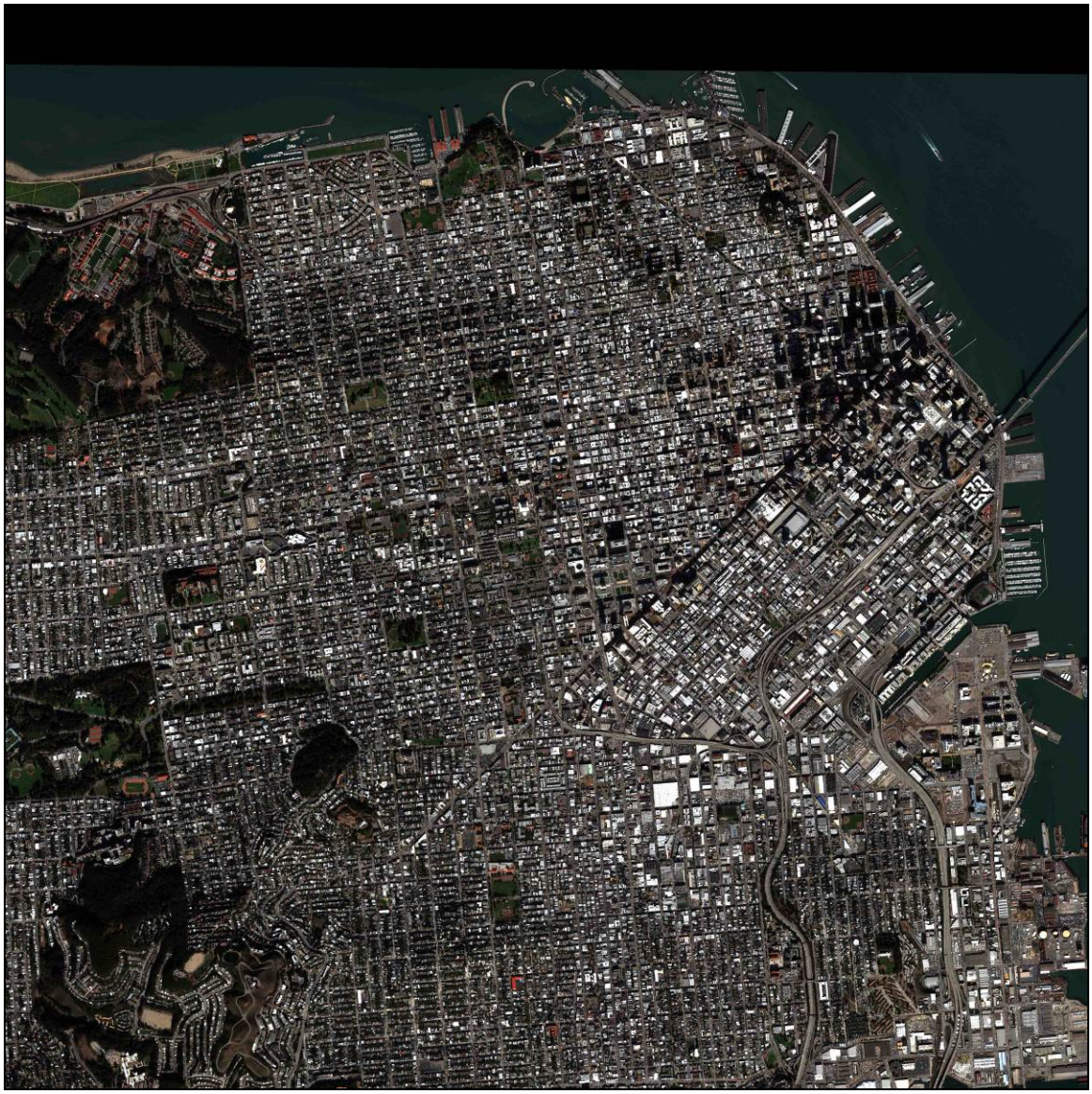

Figure 28. Orthorectified WorldView-2 image in true color

D. MASK CREATION

1. LiDAR-based Masks

The core of this fusion technique revolves around mask creation using the LiDAR data as the basis for the rules. Five masks were created from the LiDAR data which representing the water class, ground class, multiple returns, intensity, and above ground level.

ENVI has the ability to build masks based on the digital number values of an image. Generation of the water class, ground class, and multiple returns images was fairly straightforward as values greater than zero were determined to be features and anything else was not. The build mask tool allows these criteria to be entered and Figure 29 shows the three created masks. The mask's values are all now either zero, indicating the mask as off or one, indicating the mask as on.

Figure 29. Masks: left-water class mask; center-ground class mask; right-multiple returns mask

In a study on LiDAR intensity mentioned earlier from the University of Cambridge and University of Wales, research determined that most natural objects had LiDAR intensity returns of 50% or higher, whereas manmade materials were typically less than 50% (Antonarakis et al., 2008). The generated intensity image has values between 0 and 255. A histogram of the LiDAR intensity was created and is displayed in Figure 30. The histogram did indicate a natural inflection breakpoint between manmade

and natural near the value of 120, slightly less than 50%. This was utilized in order to differentiate natural and manmade surface features processed accordingly using build mask.

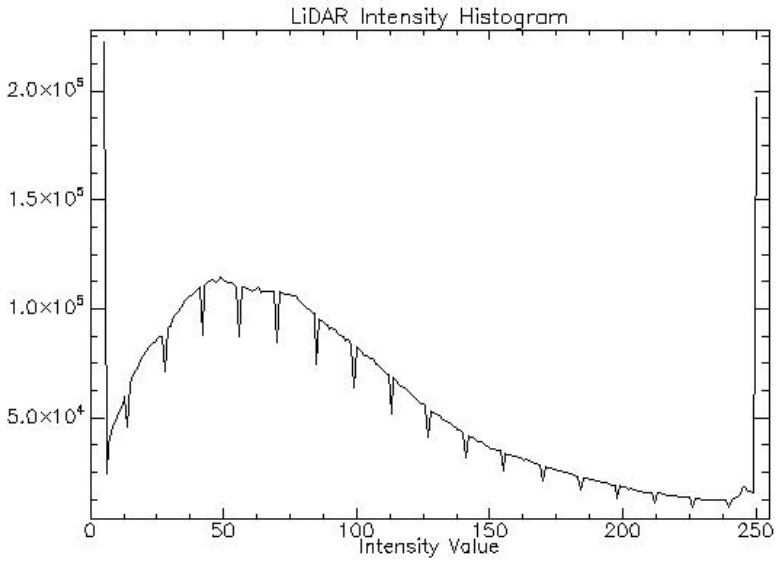

Figure 30. Histogram of LiDAR intensity values

In a similar manner, the AGL were used to differentiate regular buildings from skyscrapers. There is no set standard for what height distinguishes a building as a skyscraper, as it can be relative the rest of the skyline, but for the purposes of this project, the skyscraper threshold was set at fifty meters. Figure 31 shows the resulting Intensity and AGL masks.

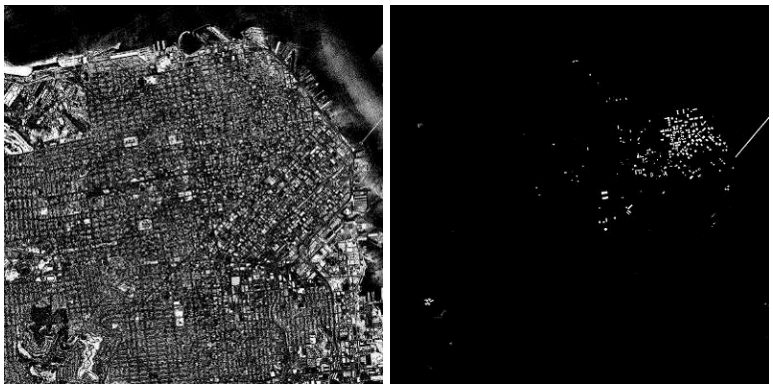

Figure 31. Masks: left-intensity mask (greater than 120); right-AGL mask (greater than fifty meters)

2. Spectrally-based Mask

The final mask is a vegetation mask created via the multispectral NDVI. ENVI has the ability to calculate this and some other vegetation indices. The resulting values of this algorithm lead to assignment of value between -1 and +1 to each pixel. As a standard rule, typical vegetation falls between 0.2 and 0.8. After analyzing the results of the WorldView values between 0.2 and 0.8, it became apparent that range was missing some vegetation, since it produced a value higher than 0.8. Readjusting the scale and analyzing results with NDVI values between 0.2 and 1.0 captured most of the vegetation. Using the build mask tool and setting the NDVI values between 0.2 and 1.0 created a mask band for vegetation. Figure 32 shows the progression from NDVI band to mask band.

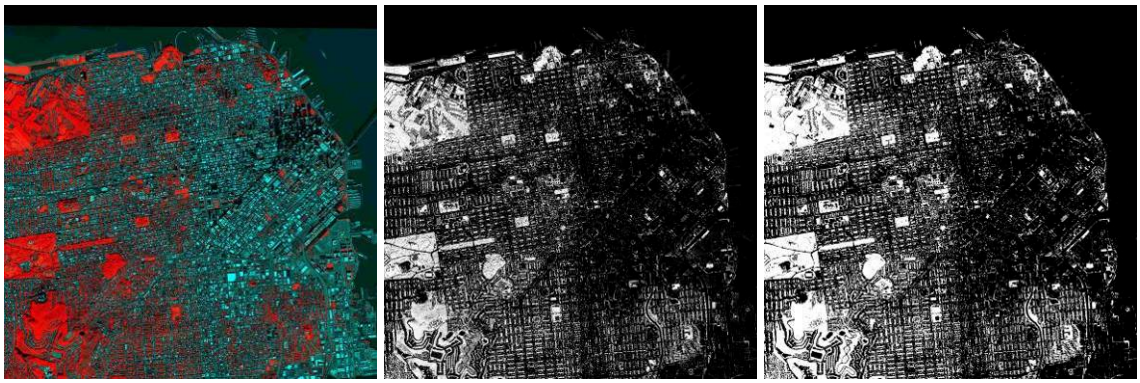

Figure 32. NDVI: left-NDVI false coloring as red band; center-NDVI displayed in grayscale; right-NDVI mask (greater than 0.2)

3. Fusing Masks

To begin the rule based classification process, the created masks were fused together by applying masks to other masks. This resulted in seven distinct classification sets based on LiDAR and NDVI. These classes included: water, tree, grass, earth, road, skyscraper, and building.

a. Water

The water mask was created solely on the original water class mask. All the areas in this region represent water.

b. Tree and Grass

The tree and grass masks first utilize areas that are considered not water class. Areas that have an NDVI value greater than 0.2 are then masked which indicate vegetation. The vegetation mask is further masked by the multiple returns mask. If the area also has multiple LiDAR returns, the resulting mask is the tree mask. If the area only has one LiDAR return, the resulting mask is the grass mask.

c. Earth and Road

The earth and road masks follow the process above for exclusion from the water class. The NDVI mask is then applied to ensure the NDVI value is less than 0.2 to indicate it is not vegetation. The next mask applied is the ground class mask which ensures the remainder is considered ground. It is then further masked by the intensity mask. If the intensity value of the area is greater than 120, the resulting mask is the earth mask. If the intensity value of the area is less than 120, the resulting mask is the road mask.

d. Skyscrapers and Buildings

The last set of masks was the skyscraper and building masks. Again, they initially follow the same procedure to determine that they are not in the water class. The NDVI mask was then applied to ensure a value of less than 0.2 indicating not vegetation, and in turn the ground class mask was applied this time to ensure the remainder was not considered ground. The above ground level mask was the last to be applied. If the AGL value is greater than fifty meters, the resulting mask is the skyscraper mask. If the AGL value was less than fifty meters, the resulting mask is the building mask.

Figure 33 shows the results of each of the seven fusion-derived masks representing the terminal nodes of the decision tree.

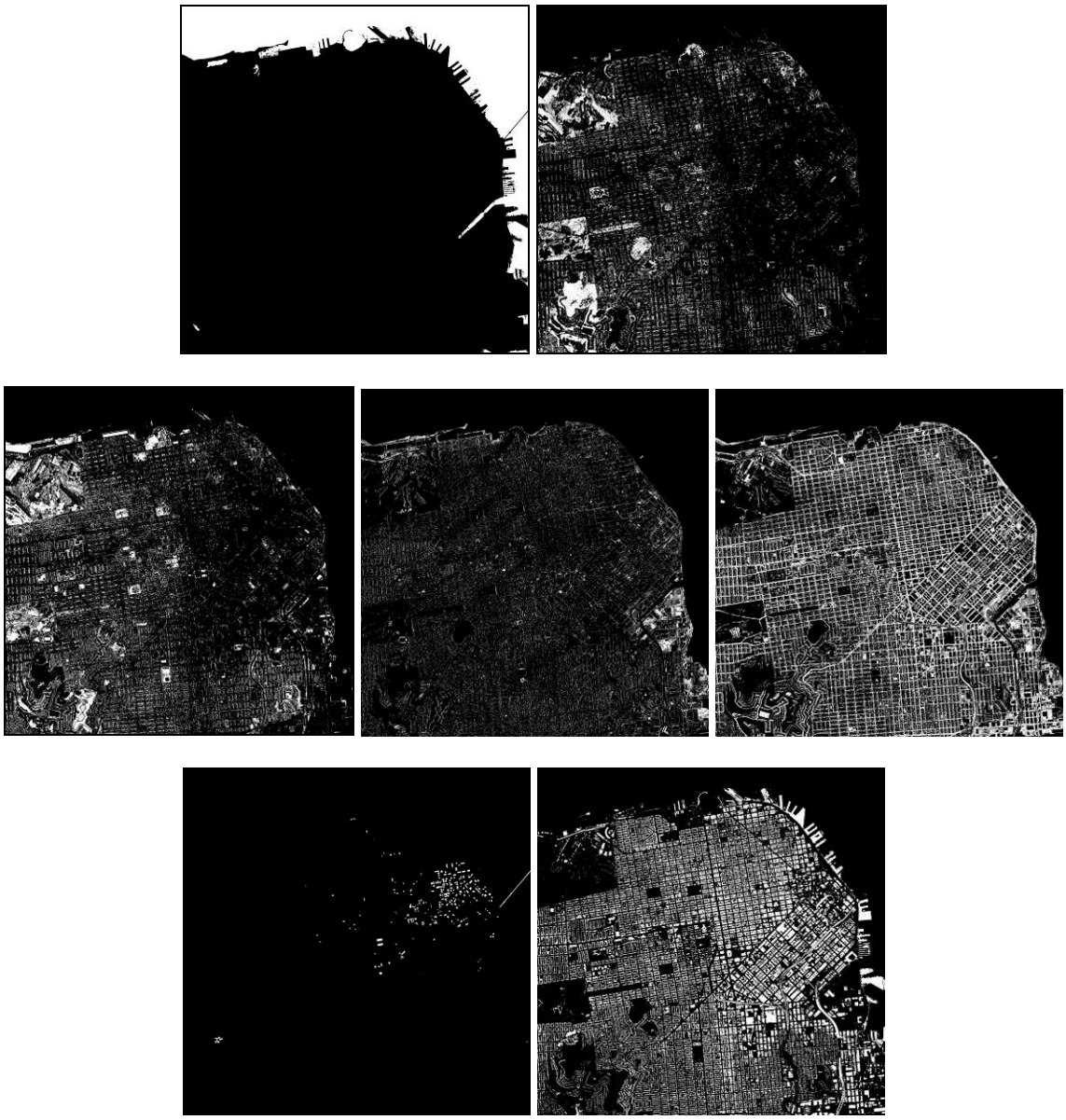

Figure 33. Fused masks: top left-water; top right-tree; middle left-grass; middle center-earth; middle right-road; bottom left-skyscraper; bottom right-building

E. REGIONS OF INTEREST AND CLASSIFICATION

1. Creating Regions of Interest

For this project, twelve regions of interest were created in order to run a classification tool against the images. The ROIs were created based on visible inspection of the true color imagery. Each classification was also designated to one of the seven

masks where that classification fell within that mask's parameters. The urban landscape and physical cues such as the road network and grid system were better preserved using this technique. The ROIs created along with their mask are as follows:

- **Water** – water
- **Tree** – tree1 (urban), tree2 (park)
- **Grass** – grass field, tennis court
- **Earth** – beach, soil
- **Roads** – pavement
- **Skyscraper** – skyscraper
- **Building** – commercial roof, residential roof, elevated pavement

Figure 34 shows the average spectra for each ROI. These were used as the training data for the maximum likelihood classifier. Note some of the similarities of the manmade spectra in multispectral data.

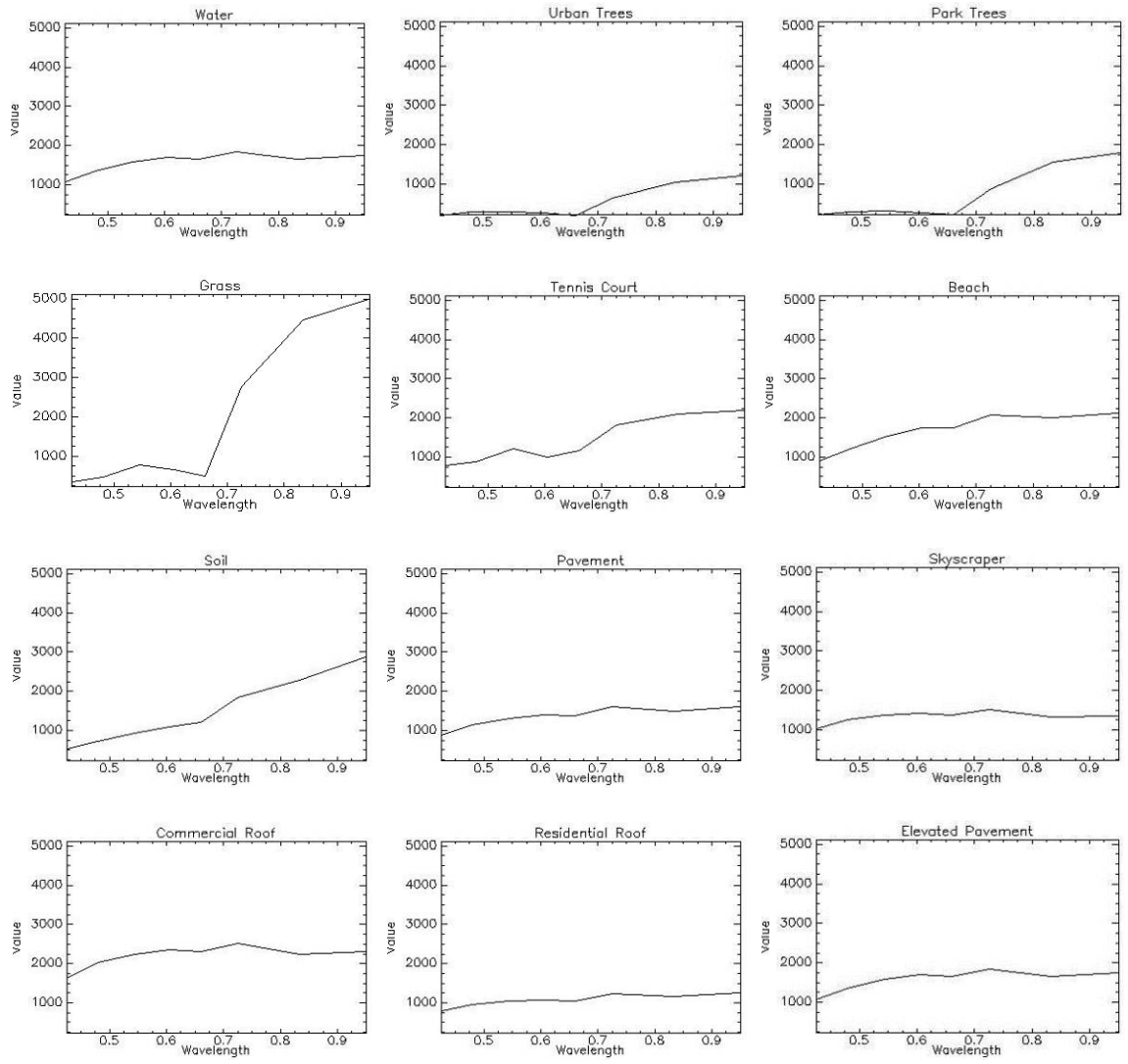

Figure 34. Training data: the average spectra for each region of interest

2. Classification

The Maximum Likelihood classifier was chosen to apply supervised classification with the created ROIs. Maximum likelihood is a classifier that assumes the statistics in each band are normally distributed and calculates the probability that each pixel belongs and assigns classification based upon its maximum likelihood. The WorldView reflectance image was selected as the input file and one mask was selected as the mask band. The corresponding ROIs were then selected for processing. In order to make the process as robust as possible, the probability threshold was set as ‘none’ and data scale

factor set at 1.00. Once the classified image was created, the file was saved as an ENVI data file. This process was repeated each time with the seven created masks utilizing each set of ROIs. Figure 35 shows each of the resulting seven classified images.

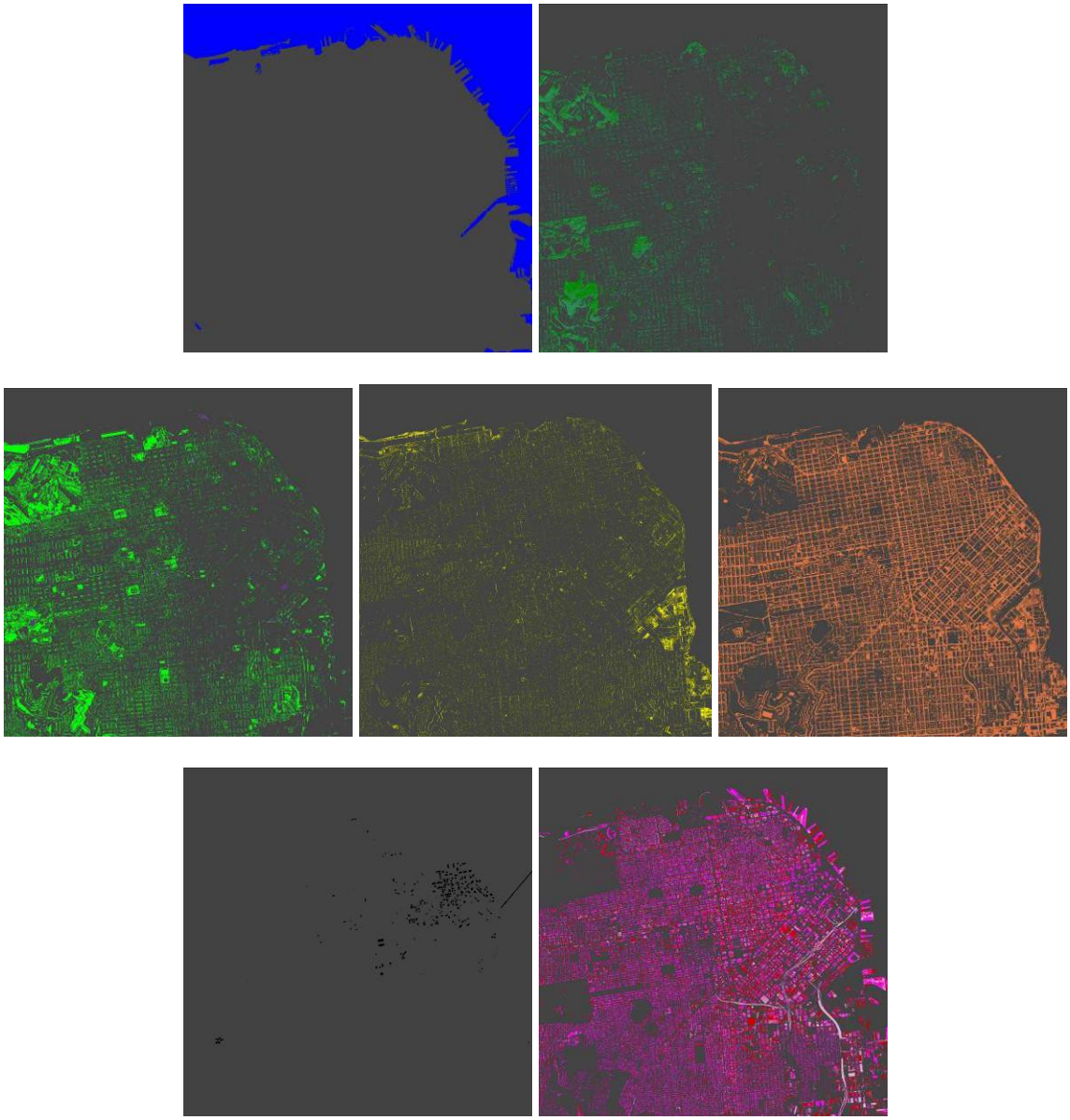

Figure 35. Masked classification results: top left-water; top right-tree; middle left-grass; middle center-earth; middle right-road; bottom left-skyscraper; bottom right-building

For comparison, the maximum likelihood classifier was run again, this time on the entire WorldView-2 image with all regions of interest, without utilizing the fused masks. Figure 36 shows the results of the classifier without fusion.

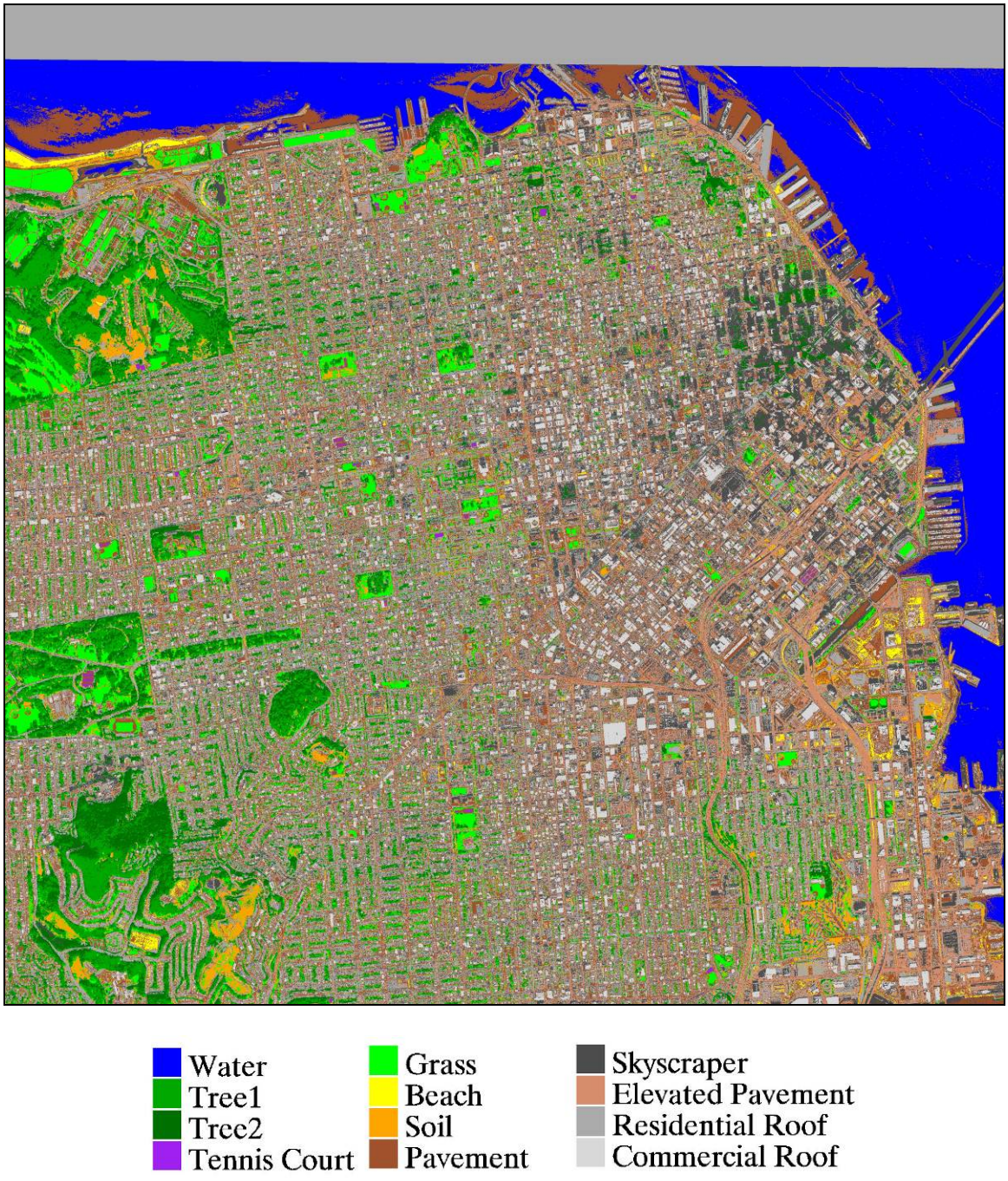

Figure 36. WorldView-2 classification results without fusion, multispectral only

3. Fusing the Classified Images

A composite image fusion of the seven masked classification images was performed using a custom IDL program. Arrays were created for each image. Each image was run sequentially with masked pixels in the first image being replaced with all values from the next array. This was repeated until each fused classified image had been scanned and a single coherent classification image remained with no pixels set as masked or unclassified.

The resulting image did not have an associated header file. In order to display the fused image correctly, geographic information was taken from the WorldView reflectance image. The classification values and colors were edited manually to correct discrepancies in class order and associated color.

The final fused classification, incorporating LiDAR fusion and limited spectral classifications is presented in Figure 37. The entire process is diagramed in the flowchart in Figure 38.

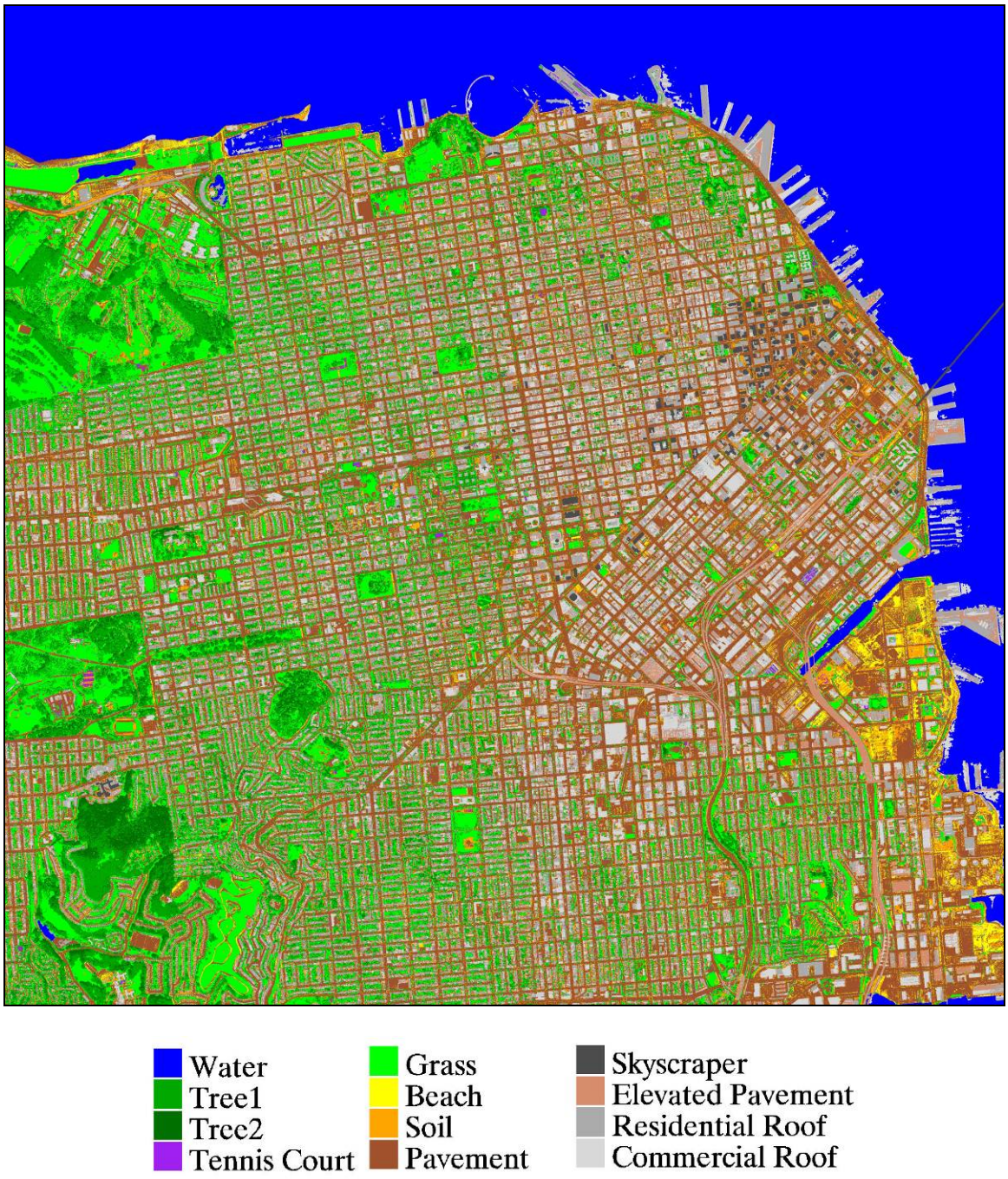

Figure 37. Final fused classification image

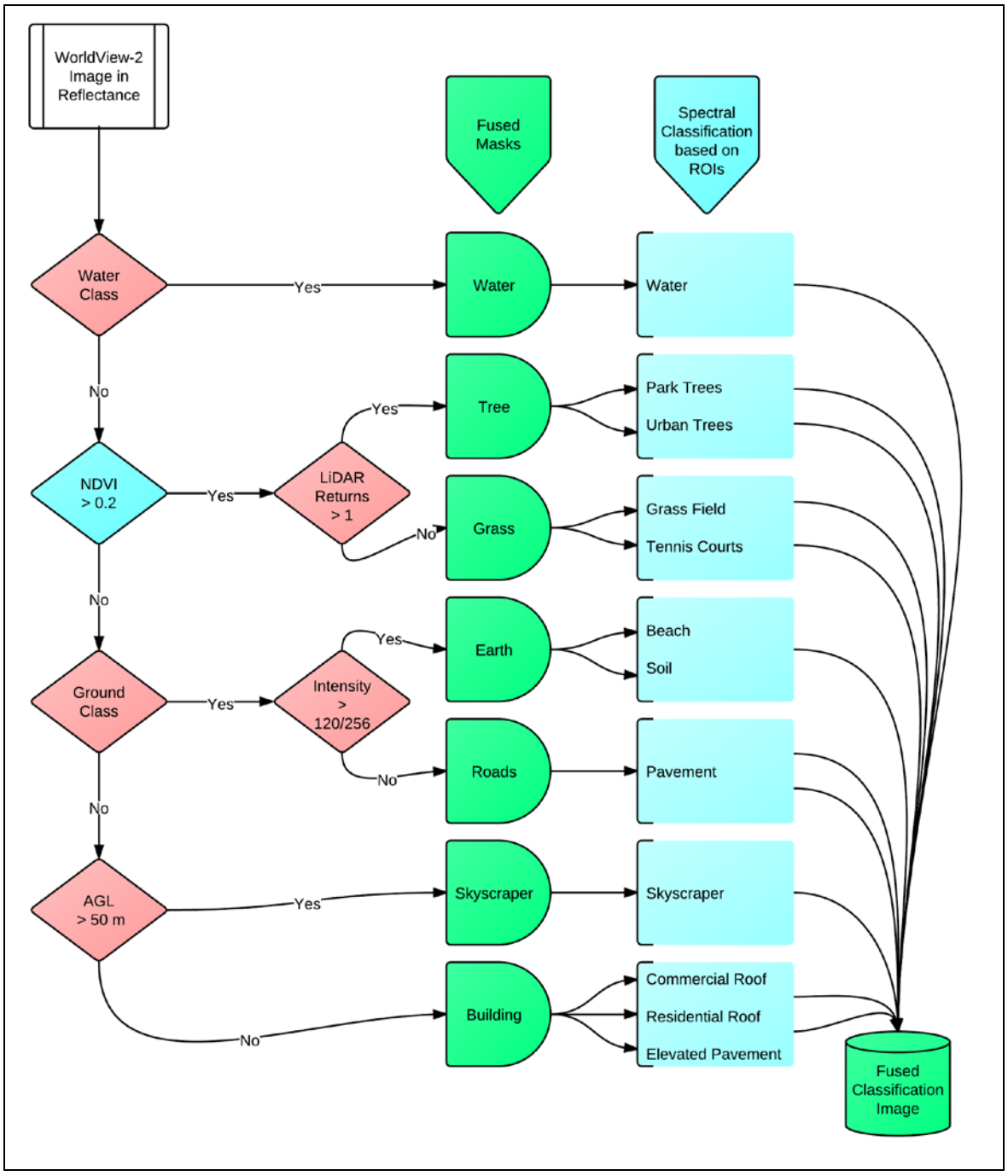

Figure 38. Flowchart of fusion classification technique decision tree

V. EVALUATION AND ANALYSIS

This chapter compares and evaluates the created products. The first section describes visual comparison between the MSI classification image and the fused classification image. A true color image is displayed next to them for reference. The next section analyses collected ground truth results and error matrices.

A. INITIAL VISUAL COMPARISON AND ANALYSIS

One of the most noticeable differences seen quickly in the classification results is how some of the water in the MSI class image was classified as pavement. Figure 39 is a sample of this near Gashouse Cove in the northern shore of San Francisco.

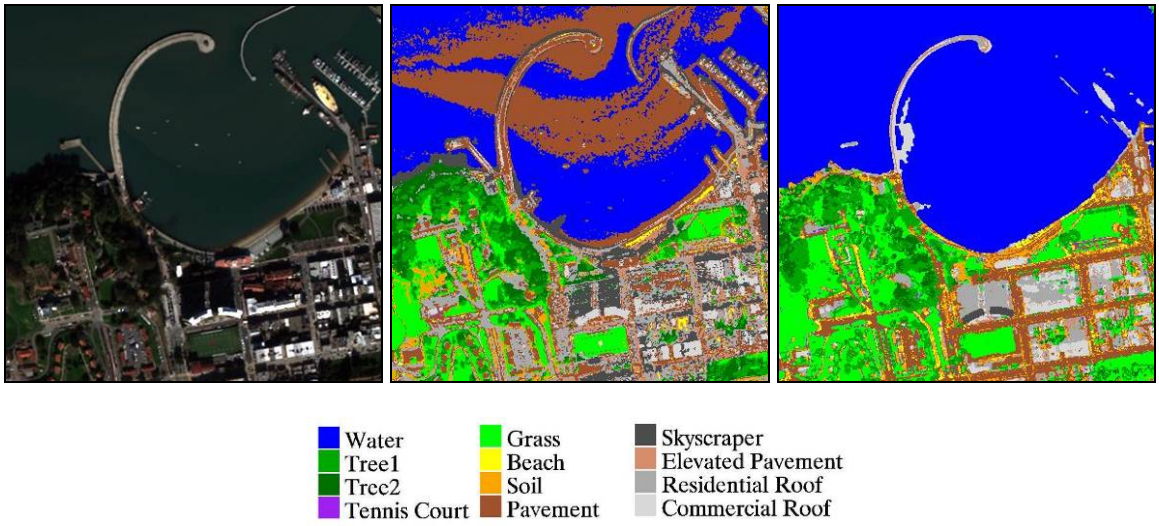

Figure 39. Northern shore near Gashouse Cove: left–true color; center-MSI classification; right-fused classification

From the true color image, it appears that there may have been sediment in the water that altered the spectra of those areas leading the classifier to predict pavement rather than water. The fused image does not suffer from this because a LiDAR based mask was applied to the water. There do seem to be errors in the fused image as part of the docks are missing and additional non-water areas are added. This is most likely due to errors in the LiDAR water classification or errors that occurred when the point cloud was converted into a raster format.

Another interesting variation is how vegetation in parks was classified. Figure 40 shows the northeast corner of Golden Gate Park.

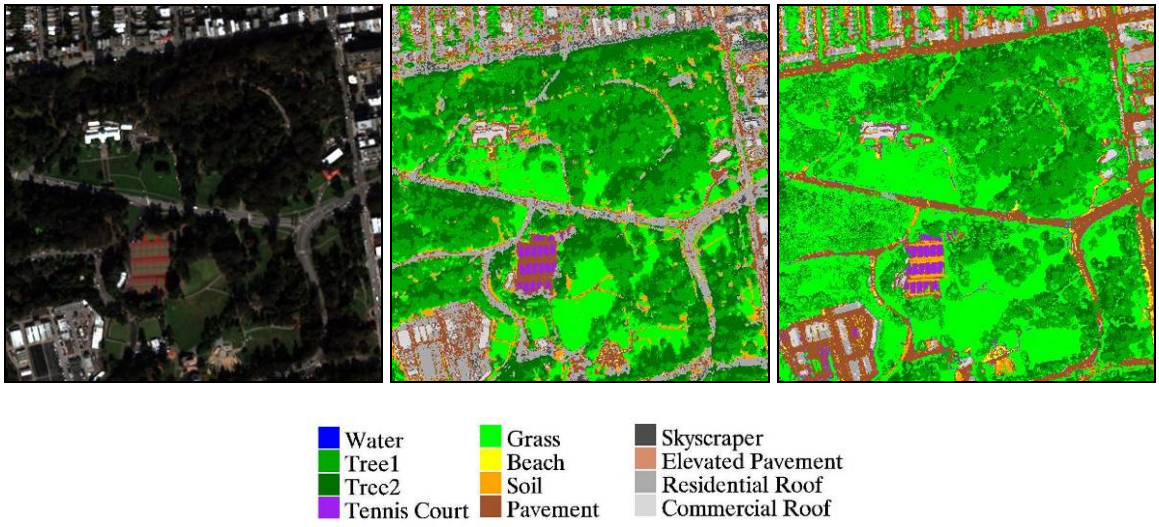

Figure 40. Northwest corner of Golden Gate Park: left–true color; center-MSI classification; right-fused classification

The treed areas in the MSI-only image match what the true color is displaying better than the fused results. The fused results display more sparsely laid out trees with more area classified as grass. The node used to differentiate trees from grass was the number of returns based on the LiDAR data. The theory behind this was that areas with multiple returns were more likely trees than grass. The results indicate that some treed areas also display a single return. Tree species and leaf thickness play a large role in this determination along with seasonal leaf-on and leaf-off status. In this example, the LiDAR data were collected in leaf-on spring and summer conditions. It seems that MSI alone may have proved to be more accurate in separating trees from grass than this fused result.

A large paved area near Pier 48 is displayed in Figure 41. This area shows a fairly large expanse of paved area with the large parking lot. Soil, beach, and elevated pavement are fairly mixed in the MSI image. Their spectra are very similar and are not varied enough to clearly differentiate them. The fused results show the same type of mixture occurring as well. The MSI-only image does indicate some areas that are skyscrapers. The true color image reveals that those areas are actually shadowed areas

whose spectra may appear similar to the region of interest created for the skyscraper. The fused result does do a slightly better job at preserving the road network geometry.

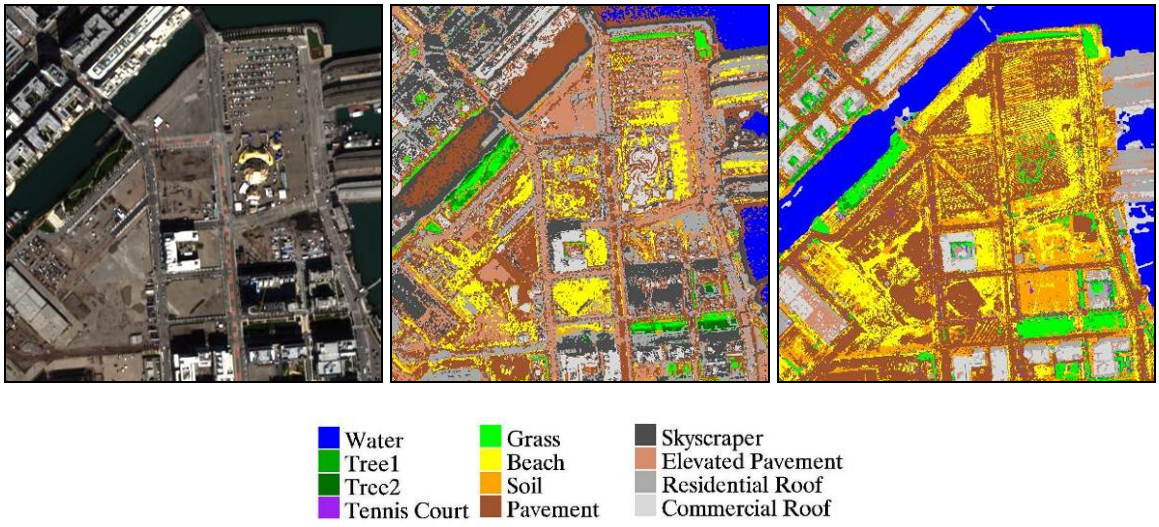

Figure 41. Dock area on eastern shore near Pier 48: left–true color; center-MSI classification; right-fused classification

The next area of analysis was deeper in the city and the road networks. Figure 42 shows part of the urban area near the junction of U.S. Highway 101 and Broadway Street, slightly north of downtown proper.

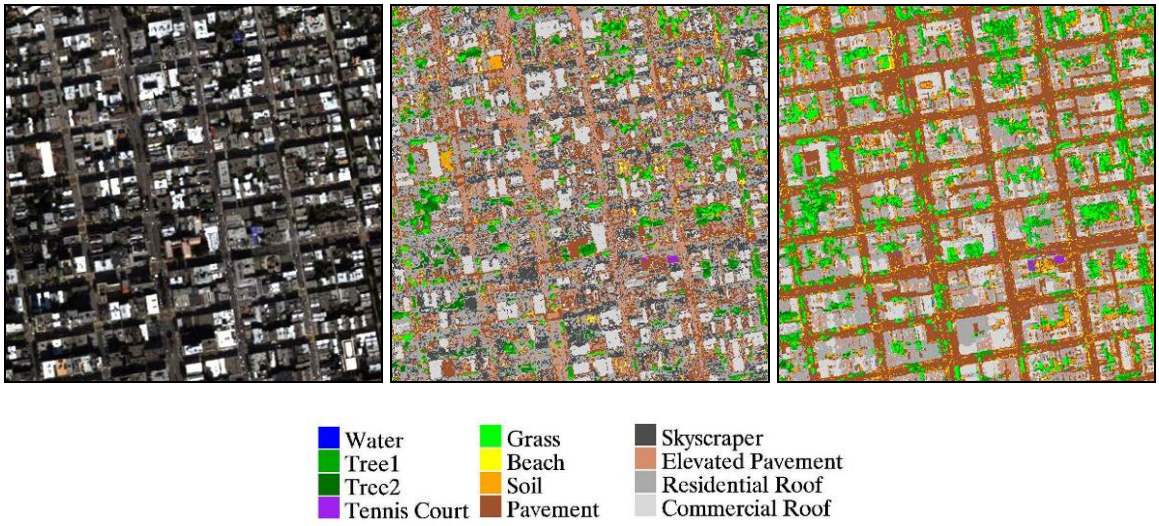

Figure 42. Urban area near U.S.-101 and Broadway: left–true color; center-MSI classification; right-fused classification

The misclassified pavement areas seen in the last example are present here as well. The MSI classification image does a decent job at distinguishing buildings and vegetation from the rest of the scene but much of the road network is lost in the rest of the classifications. The fused results do a good job at preserving the road network and building geometries. By incorporating the LiDAR information regarding ground and intensity as well as NDVI, the pavement network is kept crisp. Within the bounds of the roads, the fused image is able to distinguish some soil and vegetation in between roads on the medians. There is still some misclassification in the fused image with soil type classification sprinkled a bit amongst the roads. This could be attributed to an intensity threshold that may need to be adjusted and also the spectral similarities between the two classes.

The next set of images will be in the heart of downtown San Francisco. Figure 43 shows the results from that area.

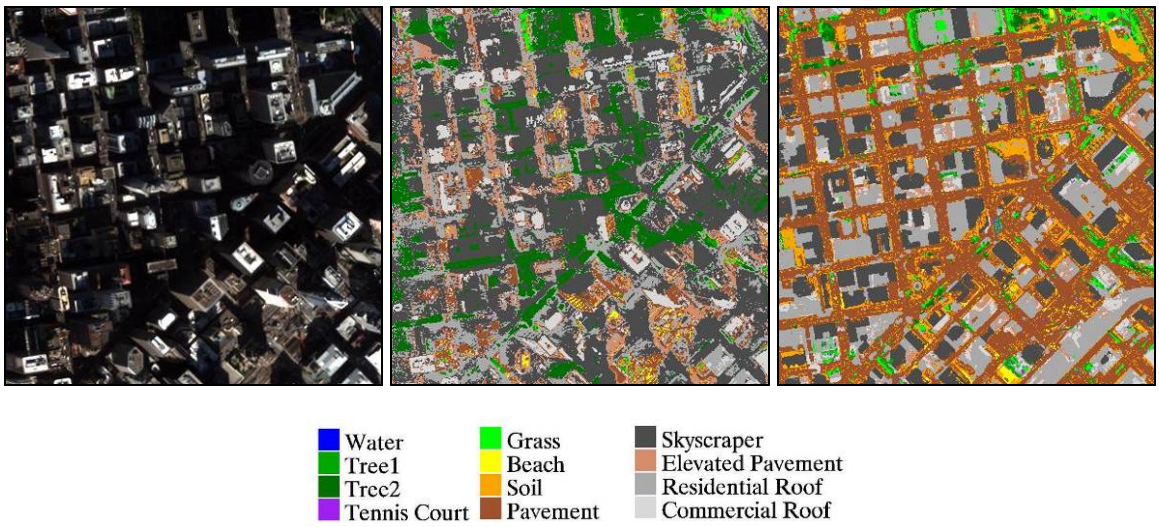

Figure 43. Downtown San Francisco: left–true color; center-MSI classification; right-fused classification

Due to the very tall buildings, there are many shadows cast in this region. Shadows tend to modify the normal spectrum of a material and make it appear like a different material. Because the WorldView-2 image is not exactly nadir, the layover effect is occurring in the image. This is an artifact where the objects with higher elevations appear in pixels

offset from their true ground locations. The LiDAR data do not suffer from this artifact due to the nature of the data, but practically all spectral sensors will show some slight layover of tall features if the image is not perfectly nadir. The MSI-only classification image is affected more by shadow, as almost all the shadowed areas are classified as one of the building types. Surprisingly, there are large areas in downtown that the MSI-only image classifies as types of trees, which may also be a result of the spectral modification caused by shadows. The fused image displays downtown quite well. The combination of LiDAR ground class and above ground levels performs well at distinguishing buildings from the road pavement and also dividing buildings and skyscrapers.

The last section that was visually inspected was the San Francisco end of the San Francisco-Oakland Bay Bridge. These images are displayed in Figure 44.

Figure 44. The San Francisco-Oakland Bay Bridge: left–true color; center-MSI classification; right-fused classification

The most noticeable feature of this image is that the Bay Bridge and its shadow are classified as two different objects in the MSI-only image. The bridge itself is classified as a mixture of elevated pavement, pavement, soil, and beach which is spectrally typical of the other non-building man made areas already analyzed. Although physically water, the bridge shadow on the water caused enough spectral dissimilarity to cause it to be classified as a skyscraper. The fusion image performs well in this scenario due to the

water classification from the LiDAR. The bridge is also classified as a skyscraper. Although a misclassification by name, the bridge does fit the rules set of not being water, not being vegetation, and being over fifty meters tall from the respective surface.

B. GROUND TRUTH AND ERROR MATRICES

In order to accurately evaluate the created products, ground truth was needed for the study area. A random sampling distribution was created throughout the study area using custom IDL code. At least 10 points were collected from each of the classification types. Ground truths were collected through a combination of on-site ground truthing and analysis of open source StreetView images and Earth imagery available from Google. 220 points were created and evaluated for this process. For each point, classification results were collected for the fused classification image, the multispectral classification image, and ground truth.

During ground truthing, it was determined that tree1 and tree2 classifications would best be combined for analysis, as tree species were difficult to determine based on resources available. Due to this, a combined tree class was used.

The northern area of the WorldView-2 image includes an area of no data. This area was beyond the limits of the multispectral image, but was necessary for the LiDAR processing and was obviously misclassified as residential roof in the MSI-only image. Any ground truth point created in the no-data region was omitted in order to provide a better analysis result of all images.

Each classification image was then compared to the ground truth and two error matrices were created.

1. Multispectral Classification Analysis

Table 1 is the error matrix created for the multispectral-only classification image.

		Ground Truth										
		Beach	Commercial Roof	Elevated Pavement	Grass	Pavement	Residential Roof	Skyscraper	Soil	Tennis Court	Trees	Water
Multispectral Classified Results	Beach	**0.43**	0.14	0.00	0.00	0.14	0.14	0.00	0.14	0.00	0.00	0.00
	Commercial	0.00	**0.44**	0.00	0.00	0.19	0.13	0.13	0.06	0.00	0.06	0.00
	Elevated	0.00	0.32	**0.37**	0.05	0.26	0.00	0.00	0.00	0.00	0.00	0.00
	Grass	0.00	0.00	0.00	**0.80**	0.10	0.10	0.00	0.00	0.00	0.00	0.00
	Pavement	0.00	0.22	0.00	0.03	**0.41**	0.11	0.03	0.03	0.03	0.11	0.05
	Residential	0.00	0.14	0.00	0.02	0.39	**0.15**	0.02	0.03	0.00	0.08	0.17
	Skyscraper	0.00	0.26	0.00	0.00	0.30	0.07	**0.26**	0.07	0.00	0.00	0.04
	Soil	0.00	0.13	0.00	0.00	0.13	0.13	0.00	**0.38**	0.00	0.25	0.00
	Tennis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00
	Trees	0.00	0.05	0.00	0.11	0.11	0.00	0.05	0.00	0.00	**0.63**	0.05
	Water	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**

Table 1. Error matrix of MSI-only results; overall accuracy was 45%

Tennis courts and water were found to be 100% accurate; each randomly sampled point with that classification was correctly classified as that material. Grass was well classified at 80% and trees were classified at 63% with some misclassifications. The rest of the

classifications were below 50% and had more of a mixture of error results. Minerals, soils, and manmade material tend to be spectrally similar and their distinctions may be less apparent.

2. Fused Classification Analysis

Table 2 is the error matrix for the fused classification results.

		Ground Truth										
		Beach	Commercial Roof	Elevated Pavement	Grass	Pavement	Residential Roof	Skyscraper	Soil	Tennis Court	Trees	Water
Fused Classified Results	Beach	**0.30**	0.00	0.00	0.10	0.40	0.00	0.00	0.20	0.00	0.00	0.00
	Commercial	0.00	**0.69**	0.00	0.00	0.06	0.13	0.00	0.06	0.00	0.00	0.06
	Elevated	0.00	0.42	**0.58**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Grass	0.00	0.10	0.00	**0.30**	0.13	0.10	0.00	0.10	0.00	0.23	0.03
	Pavement	0.00	0.12	0.00	0.02	**0.75**	0.07	0.00	0.02	0.00	0.03	0.00
	Residential	0.00	0.53	0.00	0.00	0.00	**0.41**	0.00	0.00	0.00	0.06	0.00
	Skyscraper	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00
	Soil	0.00	0.17	0.00	0.00	0.25	0.17	0.00	**0.25**	0.00	0.17	0.00
	Tennis	0.00	0.20	0.00	0.00	0.00	0.20	0.00	0.00	**0.60**	0.00	0.00
	Trees	0.00	0.00	0.00	0.12	0.12	0.06	0.00	0.00	0.00	**0.71**	0.00
	Water	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	**0.98**

Table 2. Error matrix of fused classification results; overall accuracy was 65%

The Fused Classification Results performed very well at maintaining city geometry when applying their classification. Skyscraper class and water class was rated at 100% and 98% respectively. In the mid-range, trees performed at 71%, tennis court at 60%, pavement at 75%, elevated pavement at 58% and commercial roof at 69%. Beach at 30% and soil at 25% performed low in comparison, most likely due to the spectral similarity as well as their geometric similarity which grouped them into the same terminal node.

Residential roof classification had an accuracy of 41%, which is also low. This may be due the study area having a greater number of commercial roofs in the area; many of the errors in residential roof were verified to be commercial roof at 42%. Another surprise result was that of grass at 30%. The MSI classification image had a much higher accuracy of 80%. When analyzing the fused results, 23% of the misclassifications were verified trees in the ground truth. Grass classification could also be the most susceptible to temporal changes. Grass can quickly change both spatially and spectrally if dug up for a construction project or even obscured by taller growth of other vegetation. It was also previously mentioned that this may have been caused by inaccuracies of the number of returns mask distinguishing trees from grass.

Overall the fused classification results had a total accuracy of 65% and the MSI-only classification had a total accuracy of 45%. This difference showed significant improvement in the classification results.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS

A. PRODUCT ASSESSMENT

This research fused airborne LiDAR data and WorldView-2 (WV-2) multispectral imagery (MSI) data to create an improved classification image of urban San Francisco, California.

A decision tree scenario was created by extracting features from the LiDAR, as well as NDVI from the multispectral data as raster mask decision tree nodes that resulted in seven general classes. Twelve regions of interest were created, then categorized and applied to the previous seven classes via the maximum likelihood classification and combined. This was compared to a multispectral classification image using the same ROIs.

The fused classification image did a better job of preserving urban geometries than MSI data alone and suffered less from shadow anomalies. Overall the fused LiDAR and MSI classification performed better with 65% accuracy than the MSI classification alone with 45% accuracy. The fused classification image performed well at maintaining the geometries of the city and representing ground features fairly accurately. When viewing the fused results, the image immediately appears more similar to that of a vector generated map.

The LiDAR and MSI fused classification image appears to be more representative of true reality than that of the multispectral-only classification image. There were some instances where the multispectral-only classification performed better such as differentiating trees from grass.

Adjustments should be made to node thresholds. Further refinements to the decision tree scheme could be made to improve final results.

B. PRODUCT LIMITATIONS AND FUTURE WORK

The product could be improved upon by acquiring different source data. The multispectral spectral resolution is not as high as that of a hyperspectral sensor. Using hyperspectral data, finer classifications could potentially be extracted such as soil types or tree species.

Temporal differences played a large factor in some of the discrepancies seen in the image classifications. The LiDAR data and the WorldView image were acquired with some time separation. Ideally, spectral imagery and LiDAR data for this type of project should be obtained during the same flight missions or near the same time. Without this time delay, temporal artifacts such as vegetation growth, urban construction, or the mobility of vehicles and boats would be reduced.

For future work, it would be interesting to see this technique applied to radar as well as other sources for nodes to be applied with spectral and LiDAR data. Another interesting idea would be to apply a more continuous model rather than a discrete binary model to each of the nodes. In this project, each node only had a yes or no option; it would be interesting to see how a number of bins could potentially lead to a more accurate classification.

Fusion method and techniques will continue to evolve as more data become available and software suites are adapted to utilize all collected information. It would be interesting to see this type of technique applied to multiple datasets in a single software program.

LIST OF REFERENCES

Anderson, G., Pukall, B., Allred, C., Jeong, L., Hoke, M., Chetwynd, J., Adler-Golden, S., Berk, A., Bernstein, L., Richtsmeier, S., Acharya, P., & Matthew, M. (1999). FLAASH and MODTRAN4: State-of-the-art atmospheric correction for hyperspectral data. *Aerospace Conference IEEE*. 4. 177–181.

Antonarakis, A., Richards, K., & Brasington, J. (2008). Object-based land cover classification using airborne lidar. *Remote Sensing of Environment*, 112, 2988–2998.

Applied Imagery. (2011). Quick terrain modeler. Retrieved from: http://www.appliedimagery.com/qtmodeler.php.

Arrington, M., Edwards, D., & Sengers, A. (2012). Validate and update of 3D urban features using multi-source fusion. *Geospatial InfoFusion II*. 8396, 83960K-1.

Coillie, F., Verbeke, L., & De Wulf, R. (2005). Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. *Remote Sensing of Environment*, 110, 476–487.

Cruchtley, S., & Crow, P. (2009). *The Light Fantastic: Using airborne laser scanning in archaeological survey*. Swindon: English Heritage.

Dial, G., & Grodecki, J. (2003). Applications of IKONOS imagery. *Proceedings of ASPRS 2003 Conference*. Achorage, Alaska.

DigitalGlobe. (2011). WorldView-2 data sheet. Retrieved from: http://www.digitalglobe.com/downloads/WorldView2-DS-WV2-Web.pdf.

Exelis VIS. (2012). E3De discover the next dimension of your data. Retrieved from: http://www.exelisvis.com/ProductsServices/E3De.aspx.

Exelis VIS. (2012). ENVI imagery becomes knowledge. Retrieved from: http://www.exelisvis.com/ProductsServices/ENVI.aspx.

Exelis VIS. (2012). IDL discover what's in your data. Retrieved from: http://www.exelisvis.com/ProductsServices/IDL.aspx.

Helt, M., (2005). Vegetation identification with lidar (Master's thesis). Retrieved from: http://www.nps.edu/faculty/olsen/Student_theses/05Sep_Helt.pdf.

Hines, E. (2011). Final report: Golden gate lidar project. San Francisco State University. Award #G10AC00122.

Holden, N., Horne, P., & Bewley, R.H. (2002). High-resolution digital airborne mapping and archaeology. *NATO Science Series I. Life and Behavioral Sciences*, 337, 173–180.

Ientilucci, E.J. (2012). Leveraging lidar data to aid in hyperspectral image target detection in the radiance domain. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, 8390, 839007–1.

Kanaev, A.V., Daniel, B.J., Neumann, J.G., Kim, A.M., & Lee, K.R. (2011). Object level HSI-LIDAR data fusion for automated detection of difficult targets. *Optics Express*, 19, 20916–20929.

Kim, A.M., Kruse, F., Olsen, R.C., & Clasen, C. (2012). Extraction of rooftops from lidar and multispectral imagery. *Imaging and Applied Optics Technical Digest*. Optical Society of America. RTu2E.1.

Kruse, F.A., & Perry, S.L. (2009). Improving multispectral mapping by spectral modeling with hyperspectral signatures. *Journal of Applied Remote Sensing*, 3, 033504.

Liu, K., Shi, W., & Zhang, Hua. (2010). A fuzzy topology-based maximum likelihood classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, 103–114.

Santos, P., & Negri, A. (1996). A comparison of the normalized difference vegetation index and rainfall for the Amazon and northeastern Brazil. *Journal of Applied Meteorology*. 36. 958–965.

Sohn, G., & Dowman, I. (2007). Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62, 43–63.

Stein, D., Beaven, S., Hoff, L., Winter, E., Schaum, A., & Stocker, A. (2002). Anomaly detection from hyperspectral imagery. *Signal Processing Magazine IEEE*, 19, 58–69.

Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., & Blair, J.B. (2011). Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. *Remote Sensing of Environment*, 115, 2917–2930.

Tucker, C. (1979). "Red and photographic infrared linear combinations for monitoring vegetation." *Remote Sensing of Environment*, 8, 127–150.

U.S. Cenus. (2010). San Francisco County, California QuickFacts. Retrieved from: http://quickfacts.census.gov/qfd/states/06/06075.html.

USGS. (2012). Landsat: A global land-imaging mission. Retrieved from: http://pubs.usgs.gov/fs/2012/3072/fs2012–3072.pdf.

Wikimedia Commons. (2008). San Francisco County Enlarged. Retrieved from: http://upload.wikimedia.org/wikipedia/commons/a/a1/California_county_map_%28San_Francisco_County_enlarged%29.svg.

Xu, M., Watanachaturaporn, P., Varshney, P., & Arora, M. (2005). Decision tree regression for soft classsification of remote sensing data. *Remote Sensing of Environment*, 97, 322–336.

Yuan, F., Sawaya, K., Loeffelholz, B., & Bauer, M. (2004). Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment*, 98, 317–328.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Richard C. Olsen
   Naval Postgraduate School
   Monterey, California

4. Fred A. Kruse
   Naval Postgraduate School
   Monterey, California

5. Dan C. Boger
   Naval Postgraduate School
   Monterey, California