

ANALYSIS OF THE STABILITY OF FEATURES AND SEPARABILITY IN  
AIRBORNE LASER SWATH MAPPING DATA

By

BRIAN JAMES LUZUM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

Copyright 2004

by

Brian James Luzum

To Phyllis.

## ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Clint Slatton, for providing direction to this research. He was able to narrow the focus of this research to an area that is interesting and relatively unexplored. His patience in answering my many questions and his constant encouragement has made this learning process both enjoyable and enlightening.

I would also like to thank my co-advisor, Dr. Ramesh Shrestha, for allowing me the freedom to pursue this research and encouraging me along the way. He has allowed me to learn a lot about airborne laser swath mapping (ALSM) systems and let me explore the aspects of post-processing the data that were most of interest to me.

The rest of my committee, Drs. William Carter, David Bloomquist, and Jasmeet Judge, have also been of considerable help. Dr. Carter answered many, many questions about the University of Florida (UF) ALSM system and the UF processing strategies. Dr. Bloomquist was always enjoyable to talk to and was very supportive of my work. Dr. Judge provided me with insight into the dissertation process, both technical and non-technical. The comments from all three have helped to improve this work.

I would like to thank Michael Sartori for explaining the processing strategies and helping me to learn how to produce top-quality ALSM data. The processed ALSM data used in this analysis were provided by Michael. I would also like to thank Mark, Tolee, Joong Yong, Devin, Jin, and Anders for the camaraderie and for taking the time to answer the many questions that I had.

I would like to thank Merri Sue Carter for facilitating the opportunity to come to UF. Once here, she also provided a lot of support which helped me through the arduous process of completing this program.

Finally, I would like to thank my wife, Phyllis, for helping me through these last couple of years. She has been incredibly supportive of the change in going from a comfortable job to a graduate student. She has also been an invaluable resource during the research process itself, willing to talk for hours about different ideas so that I could explore the many options available to me.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS .....	iv
LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
ABSTRACT .....	xii
CHAPTER	
1 INTRODUCTION .....	1
Background.....	1
Previous Research.....	3
2 ALSM BACKGROUND .....	6
Introduction.....	6
Equipment.....	7
Theory.....	9
Raw Data .....	13
Gridding Process.....	14
UF System Configuration.....	14
3 CLUSTER ANALYSIS AND PATTERN RECOGNITION .....	15
Introduction.....	15
Cluster Analysis.....	16
K-means Clustering.....	16
Metrics.....	18
Classical Cluster Validity .....	18
Dunn Index .....	19
Davies-Bouldin Index.....	19
Pattern Recognition .....	21
Nearest Neighbor.....	21
Maximum Likelihood.....	21
Rule-Based Systems .....	22

4	PHENOMENOLOGY .....	24
	Novel Set of ALSM Features .....	24
	Description of Data Sets .....	27
	Correlation Study.....	34
	Mean Elevation and Scan Angle .....	38
	Mean Intensity and Scan Angle.....	39
5	MEDIAN-BASED MEASURE.....	41
	Definition of Median-Based Measure .....	41
	Analysis of Measure .....	45
	Geographic Stability .....	47
	Temporal Stability .....	49
	Test Case.....	53
	Multidimensional Distance Measure .....	54
	Separation Provided by More Standard Features .....	56
	Summary of Standard Features.....	59
6	MUTUAL INFORMATION BASED MEASURE .....	63
	Analysis of Information Gain .....	70
	Comparison of Mutual Information to Median-Based Measure.....	72
	Effects of Grid Cell Size.....	74
	Stability.....	77
7	CONCLUSION.....	78
<b>APPENDIX</b>		
A	CORRELATION STUDY .....	83
B	COMPLETE LIST OF MEDIAN DISTANCES .....	90
C	RULE-BASED CLASSIFICATION SYSTEM.....	94
D	ALGORITHM USED FOR COMPUTING MUTUAL INFORMATION AND PARZEN WINDOWS.....	97
E	COMPLETE LIST OF MUTUAL INFORMATION MEASURES .....	100
F	GRID CELL SIZE ANALYSIS .....	104
	LIST OF REFERENCES.....	110
	BIOGRAPHICAL SKETCH .....	115

## LIST OF TABLES

<u>Table</u>	<u>page</u>
4-I. Complete list of feature initially studied.....	25
4-II. List of geometrically-related variables studied.....	27
4-III. Correlations of practical significance for Site 1. All correlations are with respect to the scan angle. NPS indicates that the correlation is not of practical significance. 36	36
4-IV. Correlations of practical significance for Site 2. All correlations are with respect to scan angle. NPS indicates that the correlation is not of practical significance .38	.38
5-I. One-dimensional separability $d_{f_i}$ between building and tree for each of the 11 features. ....	45
5-II. The four feature pairs with the greatest two-dimensional separability (in descending order) based on their average rank for all test areas. The numbers in parentheses are the ranks of the measure for that test area. ....	46
5-III. The distances from Site 3 for the same four feature pairs with the greatest two-dimensional separability (in descending order) as shown in Table 5-II. The numbers in parentheses are the ranks of the measure for Site 3. ....	50
6-I. Mutual information between the classes for each of the 11 features between building and tree classes. Information is being measured in bits. ....	70
6-II. The four feature pairs with the largest average ranking of mutual information between building and tree classes. The numbers in parentheses are the ranks of the measure for that test area and the information is being measured in bits. ....	72
6-III. The rankings of the four feature pairs (out of 55) with the largest average ranking of mutual information between building and tree classes. These are the same feature pairs presented in Table II. The first two columns of numbers give the rankings according to mutual information for Sites 1 and 2 while the fourth and fifth columns give the rankings using a median-based measure.....	73
6-IV. The mutual information of the four best features when using the buildings and tree classes. The analysis at Sites 1 and 2 is made with grid cell sizes of 3m, 5m, and 10m. The mutual information is measured in bits.....	75

6-V. The mutual information of the four best feature pairs when using the buildings and tree classes. The analysis at Sites 1 and 2 is made with grid cell sizes of 3m, 5m, and 10m. The mutual information is measured in bits. ....	76
A-I. Correlations for Site 1 (Flagler County, FL).....	83
A-II. Correlations for Site 2 (Mare Island, CA).....	86
B-I. A complete list of the median distance measures for all 55 feature pairs. ....	90
C-I. Rule-based logic for Site 1 using mean first stop intensity and standard deviation of the last stop intensity for Site 1. ....	94
C-II. Rule-based logic for Site 1 using the difference between mean first stop and low last stop elevation and mean first stop intensity for Site 1.....	94
C-III. Rule-based logic for Site 1 using the difference between mean stop elevations and standard deviation of first stop intensity for Site 1. ....	95
C-IV. Rule-based logic for Site 1 using the standard deviation of first stop elevation and difference between mean stop elevations for Site 1. ....	95
C-V. Rule-based logic for Site 1 using mean first stop intensity and standard deviation of the last stop intensity for Site 2. ....	95
C-VI. Rule-based logic for Site 1 using the difference between mean first stop and low last stop elevation and mean first stop intensity for Site 2.....	95
C-VII. Rule-based logic for Site 1 using the difference between mean stop elevations and standard deviation of first stop intensity for Site 2. ....	95
C-VIII. Rule-based logic for Site 1 using the standard deviation of first stop elevation and difference between mean stop elevations for Site 2. ....	96
D-I. A program listing of the algorithm for computing the two dimensional mutual information. ....	97
D-II. A program listing of the algorithm for computing the Parzen windows estimation of a two-dimensional probability distribution function.....	98
E-I. A complete list of the mutual information for all 55 feature pairs. The mutual information is given in bits.....	100
F-I. A complete list of the mutual information for all 11 features and 55 feature pairs. The mutual information is given in bits. ....	104

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1. Illustration of an ALSM system at work. Diagram by R. Shrestha.....	12
4-1. Illustration of the potential for scan angle to affect phenomenology. In this case, the measured first stop elevations could be significantly different depending on the angle at which the measurements are taken. Diagram by C. Slatton. ....	28
4-2. Schematic of the location of Site 1.....	29
4-3. Shaded relief (upper left) and intensity plot (upper right) of the Site 1 test area in FL. The $x$ and $y$ axes are measured in UTM (Zone 17) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the buildings shown in red and the trees shown in green.....	30
4-4. Schematic of the location of Site 1.....	31
4-5. Shaded relief (upper left) and intensity plot (upper right) of the Site 2 test area near Mare Island, CA. The $x$ and $y$ axes are measured in UTM (Zone 10) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the buildings shown in red and the trees shown in green. ....	31
4-6. Schematic of the location of Site 1.....	33
4-7. Shaded relief (left) and intensity plot (right) of the Site 3 test area in Napa Valley, CA from the May 2003 flight. The $x$ and $y$ axes are measured in UTM (Zone 10) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the trees shown in green. ....	33
4-8. Plots of the geometrically-related variables for Site 1. The number density plots of the first and last stops are in the upper left and right respectively. The scan angle and the gradient plots are in the lower left and right respectively. The scan angle is measured in degrees and the gradient is measured in meters.....	35
4-9. Plots of the geometrically-related variables for Site 2. The number density plots of the first and last stops are in the upper left and right respectively. The scan angle and the gradient plots are in the lower left and right respectively. The scan angle is measured in degrees and the gradient is measured in meters.....	37

5-1. Illustration of how mean, standard deviation, median and median absolute deviation change from normal distributions to bimodal distributions. The measure of location of the center of the samples (mean, median) is given by the solid lines while the measure of spread (standard deviation and mad) is given by the dashed lines. The blue lines are the mean and standard deviation of the blue distribution while the cyan lines are the median and mad of the blue distribution. The red lines are the mean and standard deviation of the red distribution while the magenta lines are the median and mad of the red distribution. ....	42
5-2. Scatter plots showing the separability using the feature pair mean first stop intensity and st. dev. of last stop intensity. This feature pair provides the largest average distance between classes for all test sites. Site 1 is on the left and Site 2 is on the right. ....	47
5-3. Plots of the vegetation training data on an intensity image (left) and the elevation of objects (right) from Site 2. Note that the elevations are measured in meters.....	48
5-4. Data used to test discrimination potential for temporally displaced data in Site 3. All tree points are from Site 3. Buildings points from Site 1 (FL Buildings) and Site 2 (MI Buildings) are inserted due to the lack of buildings in Site 3. Rule-based classifiers can be built along the lines described in Rule C or using cutoffs at a mean first stop intensity of 225 and a st. dev. of first stop intensity of 15. ....	51
5-5. Plots of the training area. The shaded relief is in the upper left, the intensity plot is in the upper right, a color IR DOQQ is in the lower left, while the classification plot is in the lower right. In the classification plot, the ground is shown in cyan, buildings are shown in green and vegetation is shown in red.....	53
5-6. Scatter plots showing the separability using the feature pair difference between mean first and last stop elevation and st. dev. of first stop intensity. Site 1 is on the left and Site 2 is on the right .....	57
5-7. Scatter plots showing the separability using the feature pair st. dev. of first stop elevation and difference between mean first and last stop elevation. Site 1 is on the left and Site 2 is on the right .....	58
5-8. Illustration showing how the number of pure pixels changes depending on the placement of the object on the grid. Diagram by A. Mamatras.....	61

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

ANALYSIS OF THE STABILITY OF FEATURES AND SEPARABILITY IN  
AIRBORNE LASER SWATH MAPPING DATA

By

Brian James Luzum

August 2004

Chair: K. Clint Slatton

Cochair: Ramesh L. Shrestha

Major Department: Civil and Coastal Engineering

A novel set of airborne laser swath mapping (ALSM) features are examined to determine their effectiveness in separating buildings from trees across geographically, topographically, and temporally diverse landscapes. The separability of these classes using the different features is quantified using a new median-based distance measure and an information-theoretic measure. For each of the test cases, it is possible to identify a feature space in which the distance between the two classes is large. This distance information is an indication of the separability between classes and is therefore indicative of the potential success likely when trying to classify ALSM data.

A total of eleven features were studied in this analysis. Some of the features, such as the standard deviation of the first stop elevations and the mean intensity of the first stop intensities are common in ALSM feature extraction. However, some of the features such as the standard deviation of last stop elevation and the standard deviation of both stop intensities have not been found in previous research.

A set of geometrically related variables are analyzed to determine their influence on the feature set. The scale dependence of the feature set is determined by analyzing how the mutual information changes as a function of grid cell size. This analysis provides new insights into the richness of simple two-return ALSM data and to the spatial and temporal stability of ALSM features when discriminating between classes.

Through this research a number of goals have been achieved. A comprehensive evaluation of both traditional and novel features has been made. A new median-based measure has been developed to quantify the separability of the classes. An information-theoretic based measure has been used to also quantify class separability. This concept, in many ways, is an extension of the median-based measure. Finally, a methodology has been developed which systematically measures and ranks the information contained in ALSM features with respect to a critical pair of classes: buildings and trees.

## CHAPTER 1 INTRODUCTION

### **Background**

The most common application of airborne laser swath mapping (ALSM) is to create detailed 2.5-dimensional maps from range observations. The map is considered to be 2.5-dimensional instead of 3-dimensional because the laser is not necessarily capable of detecting objects hidden under obstructing objects. However, that only makes partial use of the information that is acquired from a typical ALSM data collection. Even a casual inspection of ALSM elevation images or intensity images will show a wealth of detail. Objects such as buildings, roads, water, trees, and grass can be easily discerned. Closer inspection shows more variety in different kinds of vegetation, vehicles, and utility infrastructures.

Part of the reason that these objects show up clearly on the images is that human beings, through millennia of evolution, have become extremely efficient pattern recognition and object identification ‘machines.’ We have become very adept at identifying details that discriminate between objects of interest in our fields of vision. The goal of computer vision is to allow the computer to more closely ‘see’ the way that humans see. This is accomplished by programming the computer to calculate (and therefore observe) quantities that humans find useful in vision: brightness, contrast, texture, *etc.* To solve any computer vision problem, it is necessary to determine which of these quantities (features) is most useful in discriminating among the objects of interest.

Unfortunately in these problems, the first step can often be the most complicated. There are no theoretical guidelines to choosing the optimal feature space (Jain *et al.*, 1999). In fact, the way in which features are chosen and even the methods for evaluating the best feature set are still active areas of research. For an overview on recent research, see Duda *et al.* (2001) or Jain *et al.* (1999).

For this analysis, the problem of using an ALSM system as a remote sensing device will be constrained to attempting to discriminate between two important classes of objects, low rise buildings and trees. These two classes, in particular, were chosen because they are two of the most difficult to separate by computer (Maas, 1999a; Matikainen *et al.*, 2003). When looking at elevations or local gradients, these two classes often have similar values. In addition, these two classes are often closely intermingled spatially in urban and suburban environments, which can severely compromise the estimation of bare earth digital elevation model (DEMs). Measuring the separability of these two classes with respect to many ALSM data features provides insight into which features might best contribute to subsequent classification of ALSM data. The phenomenology of these new features can also be used to develop new adaptive filtering methods for estimating bare earth DEMs.

The separation of the two classes is measured using two methods. The first is a new median-based measure that is designed to use the non-Gaussian ALSM data that are often seen in data collections. The second is based on measuring the mutual information between the features and classes being studied. The large separations between the classes in a particular feature space are an indication that the features would make excellent candidates for use in classification.

To ensure that the results of the feature analysis are due to the interaction of the laser and the ground object only, a set of geometrically related variables is studied to determine their effect on the feature space. Also, the scale dependence of the results is tested by varying the size of the grid cells for the analysis.

### **Previous Research**

The idea of extracting features from airborne laser data is not new. Many of these ideas can be traced to the extensive research that has been done on extracting features from photogrammetric or orthophoto surveys. For instance Baltsavius *et al.* (1995), Haala (1995), Haala and Brenner (1998), and Fradkin *et al.* (1999) used various techniques to extract buildings from photo images.

The majority of the research on ALSM data has concentrated on building extraction. In this endeavor, there have been several different techniques employed. The geometrical properties of buildings (such as size, height and shape) were used to extract building footprints in Haithcoat *et al.* (2001). Krishnamoorthy *et al.* (2002) use curvature-based features and edge information to extract buildings from digital surface models. Hierarchical Bayesian nets were used by Brunn and Weidner (1998) to discriminate between buildings and vegetation in an urban environment. Snakes and dynamic programming optimization were used by R  ther *et al.* (2002) to model buildings, which allowed them to optimize the delineation of buildings as defined by the contours. Thuy and Tokunaga (2001) investigated the use of wavelets in segmenting airborne laser scanner data taking advantage of the multi-scale properties of the object space. Weed *et al.* (2002) used a lower envelope follower to create an estimate of the ground from raw ALSM and then used a gradient-based operator to extract building and vegetation. A separate category of research has evolved to take the extracted building

data and to reconstruct the buildings from that information. See for example Maas and Vosselmann (1999), Maas (1999b), Brenner (2000), Elaksher and Bethel (2001), Haithcoat *et al.* (2001), Fujii and Arikawa (2002) to name a few.

The use of airborne laser data in conjunction with other data sources has also been explored. Haala and Brenner (1999) used two data collection methods (laser data combined with multispectral imagery and laser data with ground plan information) to extract buildings from urban environments. Their discrimination methods centered on the use of textures and heights above ground level. Priestnall *et al.* (2000) used both topographic and spectral characteristics in conjunction with an artificial neural network to extract building and trees from an airborne laser data set. Both Fujii and Arikawa (2002) and McIntosh and Krupnik (2002) combined ALSM and photographic images in generating (urban) surface models. In addition, Park (2002) also used data fusion from ALSM and airborne digital photography to perform land cover classification using maximum likelihood, expert systems and Dempster-Schafer algorithms. The comparisons show that the Dempster-Schafer algorithm produces the best results.

There have been a few papers that have explored the possibility of using texture measures as a means of segmenting and classifying an airborne laser data set. One of the first uses of texture in feature extraction was the adoption of the texture transformation in Eckstein and Munkelt (1995). Hug (1997) looked not only at texture but also surface orientation and surface reflectance (intensities) to automatically detect buildings. Maas (1999a) looked into how variations in slopes at various points in an image could be used to determine the classification of objects. Elberink and Maas (2000) utilized the co-

occurrence matrix to segment and classify the laser scanner data. Height differences, surface and intensity textures, and shapes were used by Matikainen *et al.* (2003).

## CHAPTER 2 ALSM BACKGROUND

### **Introduction**

ALSM is a natural progression of the technology that started with the Geodimeter and the Tellurometer. The Geodimeter (a word abbreviated from geodetic distance meter) was created by Erik Bergstrand in the 1940s (Ewing and Mitchell, 1970). A Geodimeter computes the distance between the transmitter, to a reflector, and back again to a receiver by comparing the phase of the outgoing and incoming wave of light. The Tellurometer was developed in South Africa in the 1950s. It differed from the Geodimeter in that it used microwaves instead of light and it used both phase and pulse measuring to determine distance (Ewing and Mitchell, 1970).

In a similar vein, radar can also be used as a mapping tool. While it lacks the precise directional capability of the Geodimeter and the Tellurometer, radar does have the ability to measure both distance and intensity. The azimuthal component can be determined somewhat by beamwidth as well as intelligent discrimination of the return signal. Note, however, that radar can only provide two dimensional measurements of a three dimensional scene (range and azimuth). The accuracy can be approved by using synthetic aperture radar (SAR) to increase the aperture (and therefore resolution) of the mapping system. Interferometric SAR (InSAR or IfSAR) combines radar images from multiple perspectives to allow for a three dimensional view of a scene. However, none of these systems can produce data as accurate as that provided by ALSM.

Because the sensor is mounted in a fast moving craft and the high repetition laser fires at passive targets, ALSM is well suited to large-scale mapping projects. The device can be located in almost any kind of aircraft; helicopters, prop planes, and jets have all been used. However, ALSM has another advantage over other distance measuring devices in that it also can measure the intensity of the return. With the combined capability of measuring distances and intensities, ALSM can operate as an elementary remote sensing device.

### **Equipment**

Lasers produce coherent, polarized, collimated light that results in extremely small beam divergences, which make them ideal for measuring point-to-point distances. There is a great deal of variability in laser applications, energy production methods, and transmit energies. For the modest transmit energies that are used for laser ranging, a common type of laser used in these systems is a diode-pumped Neodymium Yttrium Aluminum Garnet (Nd:YAG) laser (Wehr and Lohr, 1999). These have become the preferred laser because they are compact, inexpensive and use small transmit energies making them ideal for a mobile platform. The most common wavelengths of light used are  $1.064\mu\text{m}$  (used in the UF system) and the frequency doubled  $0.530\mu\text{m}$ . The most common kinds of detectors for ALSM systems are avalanche photo diodes (APDs — used in the UF system) or photomultiplier tubes (PMTs).

To maximize the sampling of the ground, the laser should be fired as often as is practical. The number of transmitted pulses per second (pps) has increased as the technology has improved. The current output is typically on the order of tens of thousands of pps and can be as high as one hundred thousand pps depending on the type of ALSM system. Each pulse is roughly 10 ns long, which corresponds approximately to

3 m of distance (assuming the vacuum speed of light). Typical divergence of the beam for these laser systems is roughly 0.25 mrad (Huising and Gomes Pereira, 1998). At flying heights of 600m, the resulting spot size of each pulse is about 15cm in diameter while a flying height of 2000m would produce a spot size of 50cm in diameter. These types of laser ranging sensors are often referred to as “small footprint” lidars. This terminology evolved to distinguish them from larger footprint, e.g. 80 m, lidars that spread the beam out to penetrate porous media, such as vegetation canopies, more deeply. While the large footprint sensors can measure an effective optical reflectivity response as a function of depth for vegetation canopies, they are ill-suited for high-precision terrain mapping, which is the primary objective of the UF ALSM system. They also lead to ill-posed inverse problems because the resulting waveforms cannot uniquely map to a particular canopy configuration.

Since every pulse samples the environment, it is also desirable to ensure that the pulses are well distributed so that a wider area can be surveyed. This is typically done by inserting a movable mirror in the optical path of the system. There are several types of mirrors that have been used including oscillating mirror (used in the UF system), rotating optical wedge, nutating (Palmer scan) mirror, fiber scanner, and rotating polygon (Wehr and Lohr, 1999).

Baltsavius (1999a) gives a thorough overview of the major components found in ALSM systems, provides specifications for several commercially available systems, and also provides information on ALSM service providers. However, as with any rapidly evolving technology, the capabilities have increased in the five years since that that survey was published.

## Theory

The abundance of 3-dimensional information can only be used if the position and attitude of the laser is known in 3-dimensional space. The positioning of the aircraft (and hence the laser which is rigidly attached) is accomplished by using an onboard global positioning system (GPS) receiver. The GPS constellation consists of 24 satellites that are continuously broadcasting a coded signal on carrier waves containing satellite ephemerides and timing information. By combining this information from at least four satellites, a receiver can determine its position. The accuracy of the position will be affected by the exact observing technique used, ephemeris errors, atmospheric delays, and receiver delays. See El-Rabbany (2002) or Kaplan (1996) for an overview of the GPS signal and position estimation via GPS.

To meet the high accuracy standards for ALSM, it is necessary to use double difference carrier phase observations. This requires that a base station be set up, preferably in the vicinity of area to be surveyed. This base station should be set up either at a point with known accurate coordinates or the receiver should be operated long enough so that accurate coordinates can be determined from the observations. As the airplane is surveying, it will be making simultaneous observations with its on-board GPS receiver. Because the base station and the aircraft receiver are geographically close, many of the error sources will be highly correlated between the two receivers. Therefore by differencing the observation equations between the two receivers, many of the errors will effectively cancel. Using these differenced data from the base station, an accurate trajectory of the airplane can be computed. See any introductory GPS book (such as Hofmann-Wellenhof, 2002) for a more detailed discussion of the procedure.

The attitude of the aircraft can be determined with an inertial measurement unit (IMU). A typical IMU consists of solid-state accelerometers and fiber optic gyroscopes so the IMU can measure the acceleration and angular velocity. The IMU is firmly mounted to the aircraft so that the instrument can measure changes in motion. The accelerations and velocities can be integrated to obtain estimates of the position of the aircraft but these estimates will be degraded by integration errors. The resulting random walk causes significant errors in the trajectory of the aircraft. The way to ameliorate this effect is to shorten the time over which the integration is performed by re-fixing the position of the aircraft. For more details on inertial systems see Jekeli (2000) or Rogers (2000).

Kalman filters provide an optimal combination, in a minimum mean square error sense, of the accelerations, velocities, and positions to give an accurate measure of the location and attitude of the airborne sensor (Levy, 1997). These filters work by taking the inputs and propagating (predicting) them forward through the covariance information. The filter then combines the predicted information with the observations while accounting for the possibility of a bad observation. In addition to propagating the state variables, the filter must also reassess the covariance information. It is important to note that the complementary behavior of the errors allows the filter to determine the correction to each component's error (Maune, 2001). The Kalman filter has the advantage of moving forward using only the information from the most recent data cycle (Levy, 1997). Because this algorithm is more efficient than an algorithm that needs multiple past data cycles, it is capable of being used in real time to produce navigation data. However, in

order to perform well, it is essential to model the errors correctly (Levy, 1997). For a more complete explanation of Kalman filters see Brown and Hwang (1996).

Through post-processing information from the onboard GPS receiver and IMU, a blended navigation solution can be created which allows for the estimation of the platform location and orientation of the laser at every instant that it is fired. Combining this information with scanning mirror angles and the estimates of the range to the ground enables the creation of a map of the surface, often called a digital surface model (DSM), of the area surveyed. The subject of integrated navigation systems is necessarily complex and lengthy. A modern treatment of the topic can be found in Jekeli (2000) or Rogers (2000).

In order to get the most accurate results from an airborne laser system, the system needs to be calibrated carefully. There are a series of parameters that describe the orientation of the sensor system with respect to the reference frame of the aircraft. It is necessary, for instance, to know the offset in position from the GPS antenna and the reference point on the sensor system. In addition, corrections in roll, pitch and yaw must be accounted for. Usually, nominal values for these corrections are provided but it is preferable to check these values by surveying an area with known coordinates (Maune, 2001). The frequency of the calibration is determined by the accuracy requirements of the survey. At a minimum, calibration should be performed whenever any changes are made to the ALSM configuration but should be performed more often to improve the accuracy of the ALSM results. Since errors in the calibration values have distinct signatures, comparison of an ALSM survey to a traditional survey will indicate whether any uncompensated errors still exist in the calibration values.

There are also calibration parameters associated with the optical system. These parameters are called range bias and range walk (Vaughn *et al.*, 1996). Range bias is caused by an improper timing error in the optical system causing all ranges to be systematically offset. Range walk, on the other hand, is a phenomenon caused by the variable (could be linear or nonlinear) response of the sensor across the range of the intensity values seen in the laser returns. If the variation of this response is improperly modeled, the range of the return signal will appear to be systematically too long or too short. In other words, measured range is not only a function of the range, it will also be a function of the intensity of the return. One way to handle this problem is to determine a set of corrections based on the return intensity. This look-up table should greatly reduce the effect of range walk and should be used in any ALSM data collection where accuracy is a concern. Such look-up tables must generally be supplied by the ALSM manufacturer.

For a complete description of the geometry governing the airborne laser scanning, see Baltsavias (1999b).

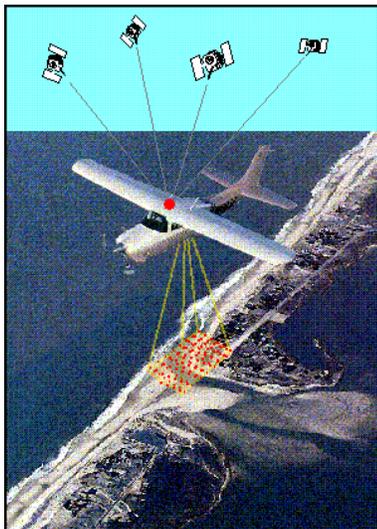


Figure 2-1. Illustration of an ALSM system at work. Diagram by R. Shrestha.

### **Raw Data**

Depending on the ALSM system, different observables are available in the output data. For instance, in addition to timing the first return of the round-trip of a laser pulse, on some systems, it is possible to also measure the time of the last return of a pulse (Axelsson, 1999). For the UF system, the first return triggers a range gate and the last return is determined by looking backwards from the end of the prescribed range to find the “first” return in the reverse direction. The typical assumption is that the first return would be reflected from the first surface encountered by the laser while the last return would be reflected from the last surface encountered (often the ground). By having information on both returns, it is possible to measure the height between the first and last surface for soft targets (*i.e.* targets without solid surfaces). This situation sometimes occurs when dealing with vegetation. However, it should be noted that for small footprint systems, even canopies can appear opaque for a given laser pulse.

Another kind of information that is sometimes available from ALSM systems is the intensity of the returned pulse (Wehr and Lohr, 1999). Since the fired pulse is constant in intensity, the intensity of the return pulse is a function of spreading losses over the path length, absorption by the atmosphere, and the absorption and/or reflection by the target. For the modest altitudes and distances covered in a single ALSM acquisition, the atmospheric conditions are generally considered to be constant. Therefore the effects of the reflectance of the ground objects will dominate the differences in the intensity of the returns. In short, the intensity can be used to discriminate between different kinds of objects on the ground (Hug, 1997; Hug and Wehr, 1997).

### **Gridding Process**

Due to the fact that the ALSM system is placed on an aircraft moving in three dimensions for both position and attitude, the way in which the laser pulses hit the ground will be unevenly spaced. In order to obtain data that are on evenly spaced nodes, the data are usually gridded. For this analysis, the initial features are computed for all data falling inside of a 3m by 3m grid cell. The size of the grid cell is determined by the density of the laser returns on the ground. The effects of using larger, 5m and 10m, grid cells are examined and the results are presented in Section 7.

### **UF System Configuration**

This analysis was performed using data that were collected by the University of Florida's Optech 1233 ALSM system. The laser system was firing at 33,333 Hz and the scan angle was set to 20°. The scan rate is typically set to 28Hz. The spot dispersal created by the oscillating mirror creates a saw-toothed pattern on the ground. The system was flown between ~600m and ~1000m above ground level for the surveys used in this analysis. Flying at this altitude, the swath width will range from ~400m to ~700m. The GPS data were collected using an Ashtech Z-12 receiver on board the plane. The attitude of the aircraft was determined with a Litton LN200A inertial measurement unit. The ALSM data were processed using Realm 3.2c, the GPS data were processed using the kinematic and rapid static (KARS) software (Mader, 1992), and the trajectory was computed using PosPac 4.02 (Sartori, personal communication, 2004).

## CHAPTER 3 CLUSTER ANALYSIS AND PATTERN RECOGNITION

### **Introduction**

Computer vision has historically been used to identify or observe objects based on passive optical intensity images (either single channel gray scale or 3-channel RGB images). The idea is to have the computer calculate measures that are useful in identifying the objects in question either on individual picture elements (pixels) or on groups of pixels. Comparing these measures between objects of interest or against undesirable objects can allow for a quantitative way to discriminate between classes of objects — data segmentation

Over the last two decades, the field of computer vision has blossomed with a wide variety of ways to measure image properties. Some of the many methods that have been used include thresholding, binary image morphology, region properties, filtering, edge detection, analysis of spatial frequencies, and texture measures, to name just a few. Each of these procedures has areas in which it excels and others in which it performs poorly. It is necessary, usually through trial and error, to determine which procedure will work well at discriminating the classes of objects that are of interest to a particular study.

Typically the procedures are evaluated using training data sets. Pixels with known classification are chosen as training data. Various procedures are run on these subsets of data and the measurements of image properties are recorded. Different measures are compared both inside of each class and between classes of objects to find the measure that will most effectively differentiate between the classes. Once the best procedures

have been determined, they are used on an image of unknown classification. If classification of an image is the desirable result, then the classifications can then be compared to the truth to determine the classification accuracy. “Truth” usually takes the form of independent information, such as *in situ* measurements of the imaged scene, or simply a portioned subset of user-classified data that was not used in the training of the classifier.

### **Cluster Analysis**

Data segmentation is the process of partitioning the data set into homogeneous groups of data. Cluster analysis is the subsequent evaluation of those groups in feature space to assess how distinct they are under the clustering method employed. In general, the idea is that when considering a particular property (feature), the data inside a cluster should be more similar to each other than they are to data points from other clusters (Jain *et al.*, 1999). There are several ways in which clustering can be accomplished including k-means, fuzzy clustering and neural networks. To give an idea of how clustering works k-means clustering will be described below. For details on other kinds of clustering, see Jain *et al.* (1999).

#### **K-means Clustering**

K-means clustering is a general algorithm that is used to partition a data set (image in this case) into  $k$  distinguishable regions in feature space (although desirable, the regions need not be disjoint). Each region is defined by its mean value along the  $n$  different axes or bases that define the feature space, where  $n$  can be as small as one (as in an intensity image), three in an RGB image, or considerably larger in more complicated feature spaces.

The algorithm for k-means clustering is iterative. Choose the number of clusters,  $k$ , that the data will be partitioned into. This number can be estimated from information about the data set being analyzed or can be estimated using the Dunn Index or Davies-Bouldin Index (see the section on Cluster Validity). An initial estimate is made of the mean value for the  $k$  clusters usually by randomly choosing the mean feature values from  $k$  points. Each of the data points is then assigned to one of the  $k$  clusters based on the minimum distance between the data point and the mean for the cluster in feature space. After each point is assigned, new cluster means are calculated. This iterative process is usually terminated when the difference in cluster means is sufficiently small. Note that due to the way this procedure is defined, each of the  $k$  groups will always have at least one member after classification assuming that the number of objects classified ( $N$ ) is larger than  $k$ .

This algorithm is straightforward to implement and is simple to understand and so it has a certain level of popularity. It is also an unsupervised classification scheme which is an additional benefit. Unsupervised classification, as the name implies, does not require any intervention on the user's part in order to obtain a classification while supervised classification does.

There are two main problems associated with this scheme. The first is that different initial parameters can lead to different classifications of the unknown objects. It is common to run a number of partitions and choose the one with the smallest overall residual between data points and cluster means. Also, determining the number of clusters is often problematic. In some cases, the number of clusters is intuitive. For instance, in handwriting recognition, the number of clusters should be equal to the number of

alphanumeric characters being recognized. However, in most cases, the number of clusters is not known before the analysis has begun.

### Metrics

Metrics are used to measure distances between features. Let  $D$  be a metric and let  $x$ ,  $y$ , and  $z$  be three vectors. Note that  $D$  is a binary operation. A metric is defined to have four properties (Duda *et al.*, 2001):

$$\text{Nonnegativity:} \quad D(x, y) \geq 0;$$

$$\text{Reflexivity:} \quad D(x, y) = 0 \text{ iff } x = y;$$

$$\text{Symmetry:} \quad D(x, y) = D(y, x);$$

$$\text{Triangle Inequality:} \quad D(x, y) + D(y, z) \geq D(x, z).$$

The most general metric is the Minkowski metric which is given by (Duda *et al.*, 2001).

$$L_n(x, y) = \left( \sum_{k=1}^d |x_k - y_k|^n \right)^{1/n} \quad (3-1)$$

where  $d$  is the number of points and  $L_n$  is often referred to as a norm. The value of  $n$  defines the type of norm used in measuring the distances. For instance,  $n = 1$  is often referred to as the Manhattan or city-block metric because distances between points are measured along each coordinate much like would be the case if traveling around Manhattan. The standard Euclidean norm is given by  $n = 2$ .

### Classical Cluster Validity

While it can be relatively easy to form clusters with a given data set, it is more problematic to determine whether clusters represent distinct entities. There is a considerable body of work devoted to determining the cluster validity (Halkidi *et al.*,

2001). In general, the idea is to compare the intercluster spacing to the intracluster spread. The methods are often categorized as using external criteria, internal criteria, and relative criteria. Many of these criteria assume that the data are either Gaussian and/or have no outliers.

### Dunn Index

The Dunn Index (DI) was created in attempt to quantify the separation between compact, well-separated clusters (Dunn, 1973, 1974). The DI for two clusters  $c_i$  and  $c_j$  is defined as (Halkidi *et al.*, 2001)

$$DI = \min_{i=1 \dots nc} \left\{ \min_{j=i+1 \dots nc} \left( \frac{d(c_i, c_j)}{\max_{k=1 \dots nc} diam(c_k)} \right) \right\} \quad (3-2)$$

where

$$d(c_i, c_j) = \min_{x \in c_i, y \in c_j} d(x, y) \quad (3-3)$$

and the cluster diameter is defined as

$$diam(C) = \max_{x, y \in C} d(x, y). \quad (3-4)$$

For compact, well-separated clusters, the Dunn Index should be large. Typically the number of clusters is varied until the Dunn Index reaches a maximum. This number is then considered to be the optimal number of clusters.

### Davies-Bouldin Index

The Davies-Bouldin Index (DBI) is also designed to measure compact, well-separated clusters (Davies and Bouldin, 1979). The DBI is defined as (Halkidi *et al.*, 2001)

$$DBI = \frac{1}{nc} \sum_{i=1}^{nc} R_i \quad (3-5)$$

where

$$R_i = \max_{i=1 \dots nc, i \neq j} R_{ij}, \quad i = 1, \dots, nc. \quad (3-6)$$

$$R_{ij} = \frac{s_i + s_j}{d_{ij,n}}. \quad (3-7)$$

In this case,  $d_{ij,n}$  is the Minkowski n-norm, and  $s_i$  and  $s_j$  are measures of the dispersion or spread of clusters  $i$  and  $j$  respectively and  $nc$  is the number of columns. The dispersion  $s_i$  is calculated from (Rajan and Slatton, 2004)

$$s_i = \left( \frac{1}{|c_i|} \sum_{x \in c_i} \|x - \mu_{c_i}\|_2^q \right)^{\frac{1}{q}}. \quad (3-8)$$

For compact, well-separated clusters, the Davies-Bouldin Index should be small.

Typically the number of clusters is varied until the Davies-Bouldin Index reaches a minimum. This number is then considered to be the optimal number of clusters.

It should be noted that both the Dunn Index and the Davies-Bouldin Index are defined using min/max measures. In other words, it will measure the extremes in intracluster spread and intercluster spacing. These indices are not well suited for data with large outliers (Pal and Biswas, 1997). Although alternative indices have been suggested (Bezdek and Pal, 1998; Pal and Biswas, 1997), these are not guaranteed to work on any arbitrary data set and the computability can be problematic for large data sets (such as ALSM creates).

## **Pattern Recognition**

There are many potential ways in which patterns can be recognized by a computer. The one aspect that all of these methods have in common is that they provide a way (schema) that provides rules for the computer to follow to make a decision on determining the classification of a pixel. There are two general classes of classification schemes: unsupervised and supervised. The k-means algorithm, discussed previously, can be used as a classification scheme. The nearest neighbor, maximum likelihood, and expert system methods discussed below are some of the more basic schemes.

### **Nearest Neighbor**

The nearest neighbor algorithm is the simplest classification method. The algorithm, as the name implies, searches through the training data in feature space and finds the known training data that is closest to unknown point. The unknown point is assigned the classification of this closest training data point. After classification, this data point is not included in the training data.

While this method is extremely straightforward to implement, the results are far from optimal. Clearly when using data with significant overlap (such as data with large outliers), the resulting classification will be somewhat haphazard. While this problem can be overcome, to some extent, by using the classification from the  $m$  nearest neighbors, the classification results are often unspectacular.

### **Maximum Likelihood**

While k-means clustering looks at which class most closely matches an object, maximum likelihood takes the concept further by calculating directly the probability that an object belongs to a given class. To perform this calculation, information about the distribution of data in each class is necessary. Typically the means and standard

deviations for each class are computed from the training data. Some *a priori* estimate of the probability of each class,  $P(\omega_j)$  as well as the class (or state) conditional probability density,  $p(x | \omega_j)$ , is needed. The conditional density is the probability of  $x$  given that  $\omega_j$  is true.

The posterior probability is then calculated using Bayes' Theorem (Duda *et al.*, 2001)

$$P(\omega_j | x) = \frac{p(x | \omega_j)P(\omega_j)}{p(x)} \quad (3-9)$$

where the probability,  $p(x)$ , is calculated from

$$p(x) = \sum_{j=1}^n p(x | \omega_j)P(\omega_j) \quad (3-10)$$

where  $n$  is the number of classes,  $x$  is the observed feature and  $\omega$  is the  $n$  states of nature.

Assuming that all incorrect classification costs the same, the classification decision is made based on which class has the greatest conditional probability.

Note that in most cases, this method implicitly assumes that the probability distributions can easily be parameterized. In short, this method usually assumes that the data have a Gaussian distribution. By calculating the mean and standard deviation, the probability distribution is well determined. If the data do not have a Gaussian distribution, this process becomes much more complicated.

### **Rule-Based Systems**

In an effort to more closely match the way in which humans analyze images, rule-based systems have been used in pattern recognition. Rule-based systems use a series of

rules, much the way that people do, in order to draw conclusions. This algorithm is easy to implement if the rules are easy to discern. It is particularly applicable to instances where the classes can be described by relationships rather than by instances (Duda *et al.*, 2001).

Rule-based systems are usually implemented using a series of if-then statements which are used to define the relationships between the various classes. A series of predicates are created which when evaluated, assign a true/false value to each input data point. Based on this output, each data point can be classified.

Rule-based systems suffer from a few drawbacks. There is no innate probability associated with the rules. Because of this, it can be challenging to use rules when there is a low signal to noise ratio and therefore a large corresponding Bayesian error (Duda *et al.*, 2001). In addition, there is no guaranteed way of determining an optimal set of rules for pattern recognition (Duda *et al.*, 2001).

## CHAPTER 4 PHENOMENOLOGY

### **Novel Set of ALSM Features**

A feature is a quantitative or qualitative property of an object. When performing classification of data, it can be extremely challenging to determine the optimal feature space from a large number of potential features since there are no theoretical guidelines in making the optimal choice (Jain *et al.*, 1999). For an overview of recent research in this area, see, for example, Duda *et al.* (2001) or Jain *et al.* (1999).

For this work, a set of features is calculated by binning the data into 3m by 3m grid cells and then performing the calculations on all measurements within a cell. The grid cell size is chosen based on the ALSM data density. To give an example of how the features are computed, each ALSM data point is initially placed in the appropriate grid cell. Once all data points have been assigned to a grid cell, a series of features (means, standard deviations, *etc.*) are computed from all data inside of a particular grid cell. This process is repeated over the entire data set.

It is customary to talk about elevations and intensities from the first and last stop when discussing the features that can be observed by a typical ALSM system. However, there are several other features that an ALSM system is capable of measuring. In an effort to determine the potential separability of classes in ALSM data, a more complete set of features is chosen for study. These features initially included are presented in Table 4-I. To compensate for the different altitudes and path lengths at which the data

were acquired, the intensities are corrected so that they are equivalent to the intensity values that would have been observed if the range were 600m (GEM, 2004).

Table 4-I. Complete list of feature initially studied.

<b>Feature</b>
Number First Stop Elevation
Number Last Stop Elevation
Mean First Stop Elevation
Mean Last Stop Elevation
Standard Deviation of First Stop Elevation
Standard Deviation of Last Stop Elevation
Difference in Mean Elevation between Stops
Difference between First Stop Elevation and Ground Estimate
Difference between Last Stop Elevation and Ground Estimate
Difference between First Stop Elevation and Low Last Stop Elevation
Mean First Stop Intensity
Mean Last Stop Intensity
Standard Deviation of First Stop Intensity
Standard Deviation of Last Stop Intensity
Difference in Mean Intensities between Stops
Scan Angle
Maximum Gradient

Many of these features are self-explanatory, but a few of them deserve additional explanation. For this analysis, the ground is estimated by choosing the lowest elevation (minimum last stop) within a 3 grid cell by 3 grid cell area (81m<sup>2</sup> area). This procedure will work well for reasonably flat terrain but will not provide a very good estimate of the ground over areas of sharp topographic relief. If more accurate estimates of the ground are needed, better procedures could be used or more accurate external information could be incorporated. The scan angle was computed by taking the computed trajectory of the aircraft and dropping a perpendicular line for each data point to this trajectory. By calculating the length of the perpendicular line and using an estimate of the height above ground level of the aircraft, the scan angle can be estimated. The maximum gradient of the ground is computed using the differences in elevation of the ground in a three by

three pixel area surrounding each pixel. For each pixel, elevation differences are computed between the central pixel and the eight surrounding pixels. The maximum absolute elevation difference is chosen to estimate the maximum gradient at each location.

After initial analyses, it became apparent that mean first and last stop elevations are of marginal use. The metrics used in this analysis made it possible to rank the metrics in terms of their effectiveness in separating classes using specific features or feature pairs. Thus it is not necessary to carry poorly performing features throughout the analysis. While these two elevation features can be used as a proxy for heights of objects in a pixel containing relatively flat terrain, over terrain with significant topographical relief, mean elevations inside a pixel are more influenced by the ground elevation than the object height. As a result, these two features are excluded from further analyses of the separation of classes.

Four of the features, listed in Table 4-II, are separated into a special class of geometrically-related variables. These variables have little predictive power in terms of classification, but they could significantly influence the feature measurements because they could alter the geometry of the laser pulse/ground object interaction. Any change in this relationship could affect either the elevation or the intensity features. The effects of these geometric variables on the other features are considered in the section Correlation Study. Each of the remaining features provides a unique perspective on the data set. By using a wider variety of features, a better attempt at spanning the feature space can be made. This results in a more systematic analysis of the feature space.

Table 4-II. List of geometrically-related variables studied.

<b>Feature</b>
Number First Stop Elevation
Number Last Stop Elevation
Scan Angle
Maximum Gradient

One category of features that has been excluded from this analysis is textures.

Textures describe the amplitude and direction of the local variation of image intensity.

Examples of texture measures include energy, entropy, contrast, homogeneity, correlation (Shapiro and Stockman, 2001), Laws' Texture Energy Measures (Laws, 1980), and Gabor filters (Jain *et al.*, 1999; Jain and Farrokhnia 1991). The reason that textures have not been utilized in this analysis is that textures are computed over a number of adjoining pixels (a sliding window). For instance, Laws' textures are typically computed over a 5 pixel by 5 pixel area. Given the precise (sub-meter) planimetric capability of ALSM, it seemed incongruous to smooth the measurements out to an area that could be 15 meters (5 pixels times 3 meters per pixel) on a side. It is possible that smaller grid cells could be used if a higher density of laser points were possible. However, this would require data from multiple strips of the current ALSM system which would, in turn, complicate the correlation study.

### **Description of Data Sets**

It is necessary to assess the impact of certain sensor parameters on the tree and building signatures before ranking the features. In particular, I wished to determine if the scan angle at which the objects are observed significantly influenced the ability to discriminate between the classes. The scan angle will influence the geometry of the interaction between the laser pulse and the terrain and landcover and therefore could in principle significantly influence the return signal. For an illustration of this, see Figure 4-

1 below. In order to facilitate this study, data are taken from a single flight strip in all cases. This allowed the computation of scan angles for each data point.

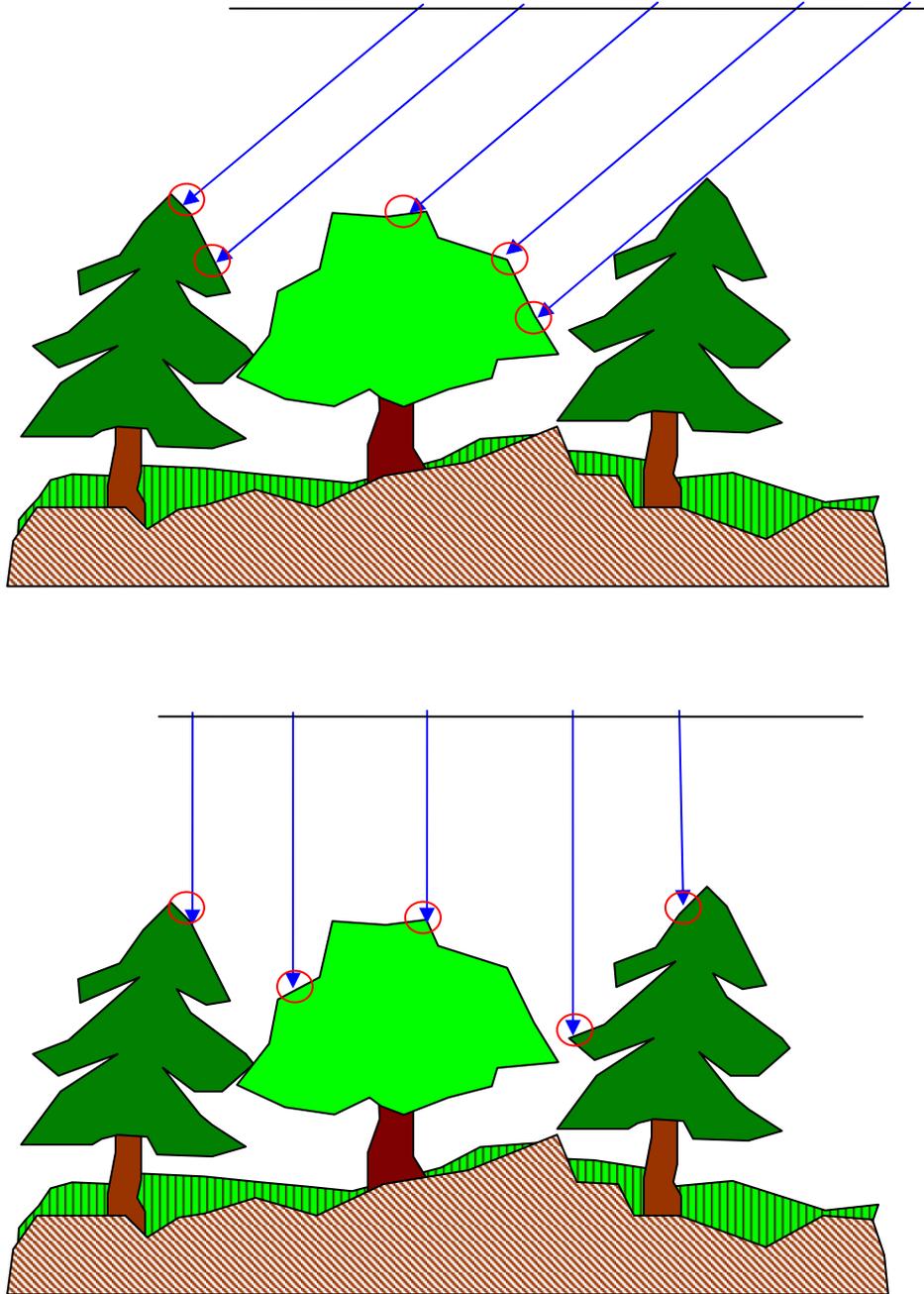


Figure 4-1. Illustration of the potential for scan angle to affect phenomenology. In this case, the measured first stop elevations could be significantly different depending on the angle at which the measurements are taken. Diagram by C. Slatton.

The UF ALSM system underwent a number of hardware and software upgrades in Spring, 2003. Included in these upgrades was a fix to the system that allowed the intensity to be determined correctly for both the first and the last stop. In order to take advantage of these improvements, the data chosen for this analysis were from data collections after this Spring, 2003 upgrade. This choice somewhat limited the number of data sets that could be used.

The swath width of usable data is bounded by the sensor altitude and scan angle settings. In order to test the robustness of the new features and separability measures, a variety of landscapes is chosen. A local data set near Gamble Rogers Memorial State Recreation Area in Flagler County, FL is the initial data set (hereafter referred to as Site 1 — see Figure 4-2). Site 1 is barrier island with sandy beaches, covered in palmetto under story, moderately sized live oaks and sparse but tall palm trees. The data were acquired in May, 2003 for a project to map coastal areas in northeastern Florida. A shaded relief digital surface map (DSM) and a return intensity plot of the area are shown in Figure 4-3. In addition, a color infrared digital ortho quad quarter (DOQQ) and a shaded relief of the training area are shown in Figure 4-3. The training data consist of 1011 coordinate pairs for the buildings and 5152 coordinate pairs for vegetation.



Figure 4-2. Schematic of the location of Site 1.

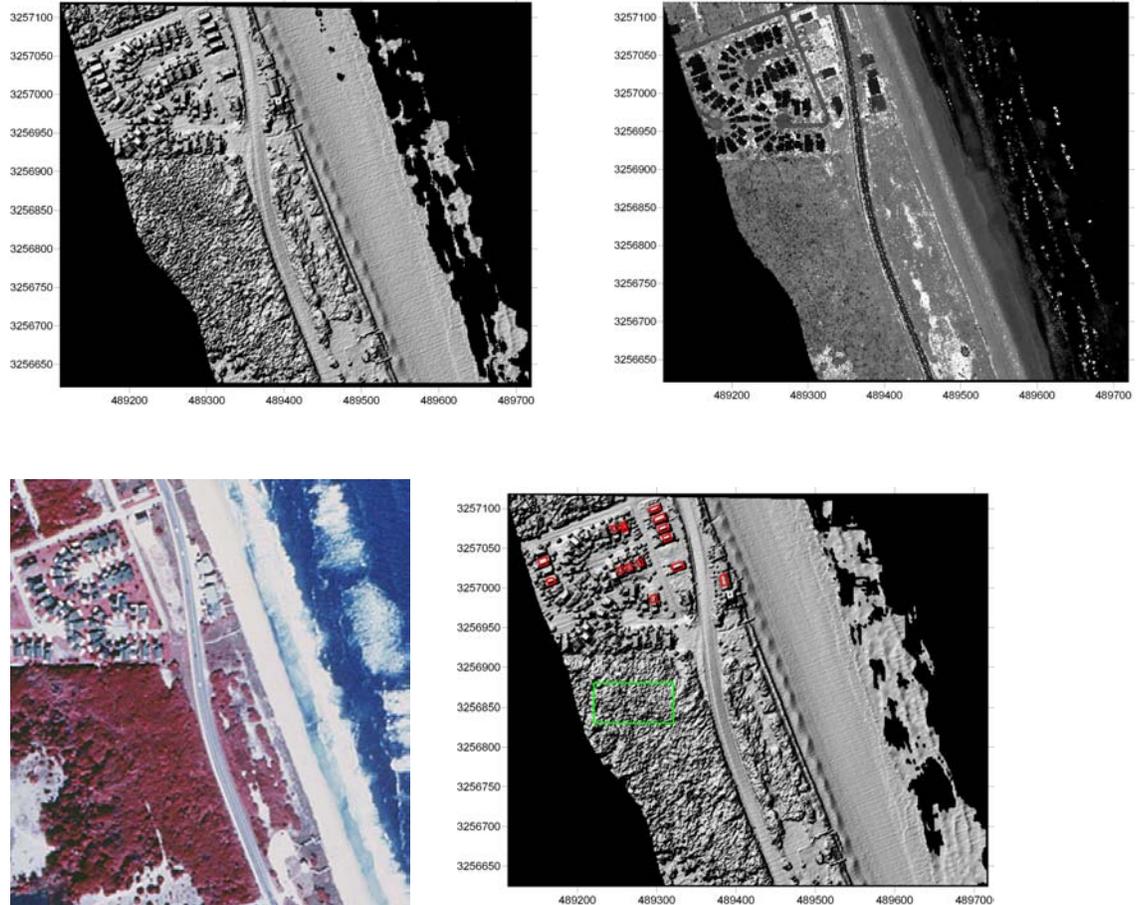


Figure 4-3. Shaded relief (upper left) and intensity plot (upper right) of the Site 1 test area in FL. The  $x$  and  $y$  axes are measured in UTM (Zone 17) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the buildings shown in red and the trees shown in green.

The Florida site is chosen as the initial data set because its geographical proximity made it accessible for ground observations and because it exhibited very little topographic relief so that modulation of the features by topography is not a significant factor. A second test area is chosen that exhibited greater topographic relief and vastly different vegetation to test the robustness of the separability. An area near Mare Island, CA (near San Francisco — see Figure 4-4) is chosen (hereafter referred to as Site 2). Site 2 is in the foothills of a coastal mountain range and with live oaks and taller ponderosa pine trees. The data for this area were acquired over a number of days in May, 2003,

although the data used here are from a single flight line. This area, shown in Figure 4-5, exhibited roughly 100m of topographic relief over a one-half kilometer distance. The training data consist of 1490 coordinate pairs for the buildings and 4797 coordinate pairs for vegetation.



Figure 4-4. Schematic of the location of Site 1.

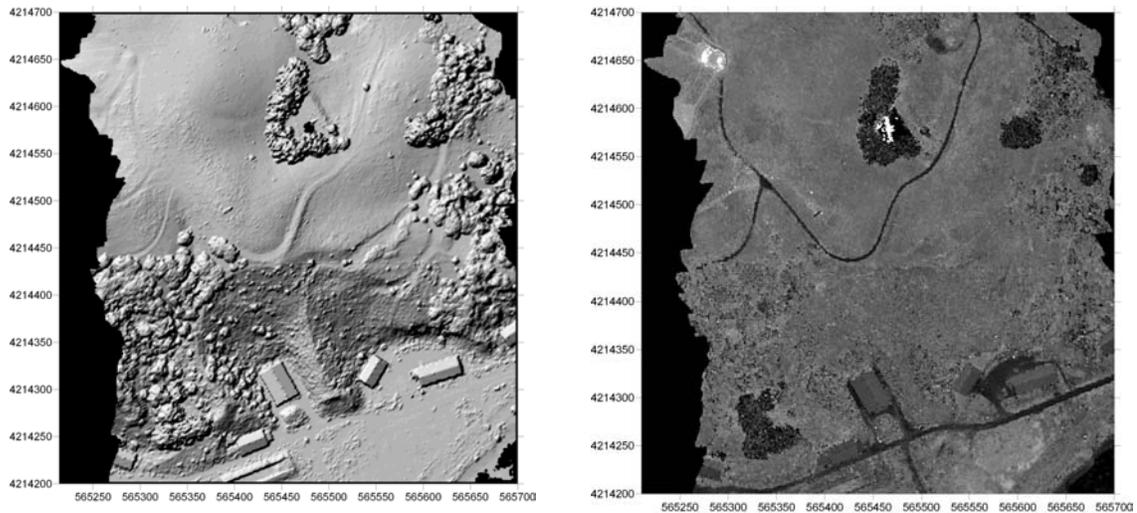
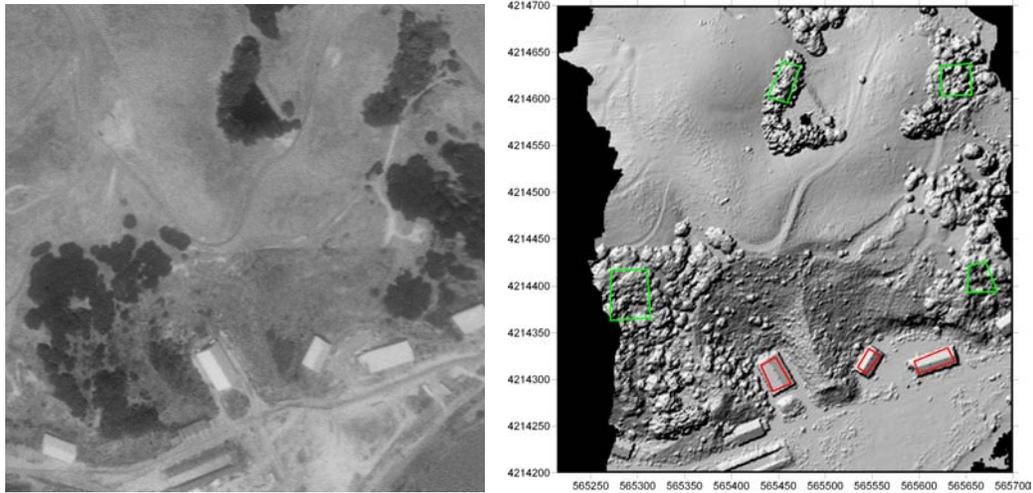


Figure 4-5. Shaded relief (upper left) and intensity plot (upper right) of the Site 2 test area near Mare Island, CA. The  $x$  and  $y$  axes are measured in UTM (Zone 10) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the buildings shown in red and the trees shown in green.

Figure 4-5—Continued.



While Sites 1 and 2 provide a variety of topographic relief and vegetation, as well as geographic separation, there is still the possibility that any set of results could be time sensitive. I wanted to test whether or not the features could provide good separability for data acquired at different times of the year. For instance, it is reasonable to assume that intensity measures could be adversely affected by seasonal changes. Tree canopies can exhibit significant changes in structural and spectral characteristics between summer and winter. This in turn bears on the question of when a site should be mapped for optimal data segmentation. In an effort to determine the effects of temporal displacement, two surveys of the same area taken roughly 6 months apart (May and November, 2003) are chosen from an area near the northern part of Napa Valley, CA (hereafter referred to as Site 3 — see Figure 4-6), as shown in Figure 4-7. This is a forested area with a deciduous coniferous mix. The training data consist of 14235 coordinate pairs for vegetation.



Figure 4-6. Schematic of the location of Site 1.

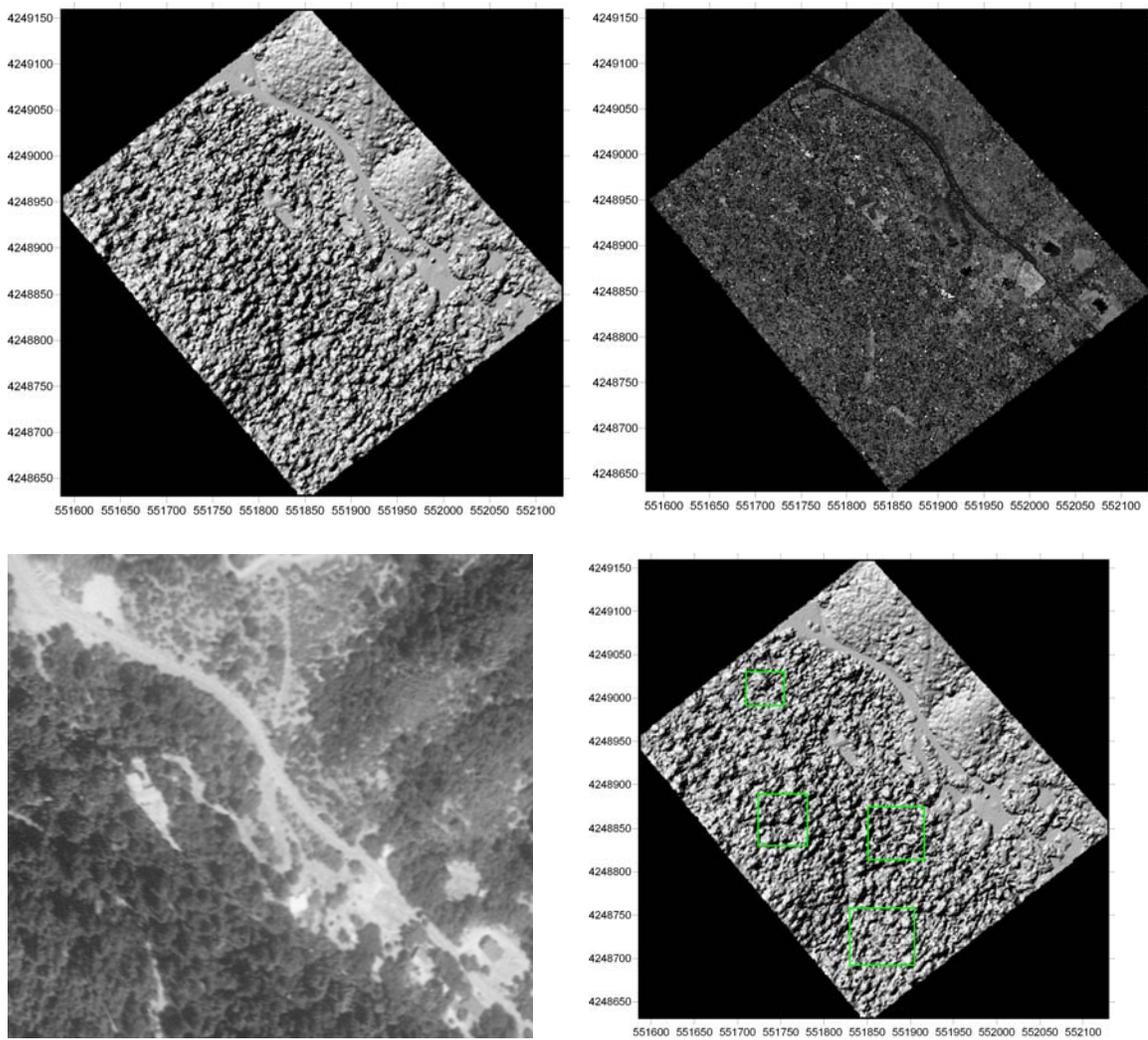


Figure 4-7. Shaded relief (left) and intensity plot (right) of the Site 3 test area in Napa Valley, CA from the May 2003 flight. The  $x$  and  $y$  axes are measured in UTM (Zone 10) meters. A DOQQ of the area is provided on the lower left. Training data is shown in the lower right with the trees shown in green.

Training data are chosen from within each of the test areas to determine the ability of these new features and distance measures to discriminate between two classes of data (buildings and trees). Using a shaded relief and intensity plot of the ALSM data as well as digital ortho quad quarter (DOQQ) imagery from each area, pixels that contained buildings and trees could be determined. Coordinates of building pixels and tree pixels are recorded into separate files. These coordinates are used as the training data for this analysis. Because there are no buildings in the rural Site 3 location, the training data there are for trees only and cover the same area for the two temporally displaced acquisitions.

### **Correlation Study**

It is conceivable that due to ALSM acquisition geometry (variable scan angle), some feature clusters could be affected by the changes in the number of points in each grid (first and last stop), the scan angle at which the data are acquired, or the gradient of the local slope of the ground. Analyses of these variables are performed for Site 1 (Flagler County, FL) and Site 2 (Mare Island, CA). Plots of these geometrically-related variables for Site 1 are shown in Figure 4-8.

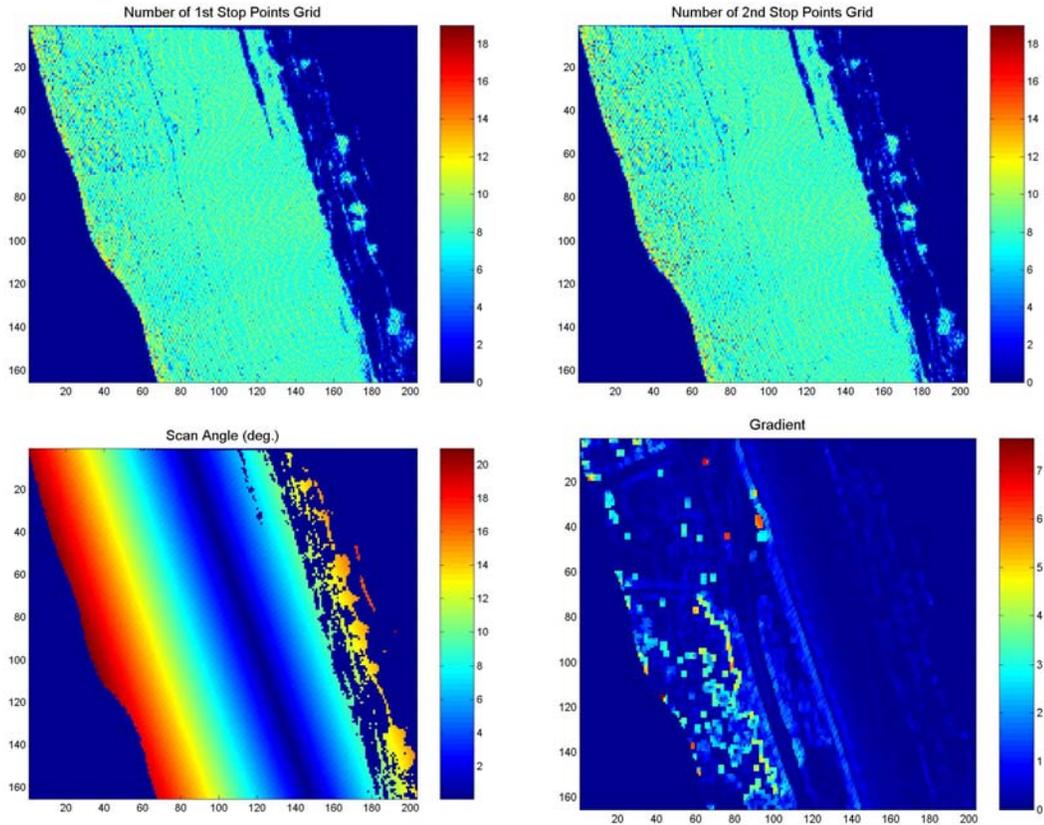


Figure 4-8. Plots of the geometrically-related variables for Site 1. The number density plots of the first and last stops are in the upper left and right respectively. The scan angle and the gradient plots are in the lower left and right respectively. The scan angle is measured in degrees and the gradient is measured in meters.

There are two interesting features in the plots of the number of points. There are a large number of points at the edge of the swath, which is a result of the kind of scanner being used in the ALSM system. Sawtooth scanners lead to longer dwell times at the edges of the swath. Also there appears to be two small-magnitude periodicities in the data: a low frequency oscillation along the trajectory and a higher frequency oscillation orthogonal to the trajectory. The low frequency oscillation is most likely due to variations in the aircraft trajectory or attitude while the higher frequency oscillation is thought to be due to the discrete increments at which the scanner mirror angle is sampled.

To test the relationship between each of the thirteen features and these four geometrically-related variables, correlations and significance are computed. The most significant results for Site 1 are listed in Table 4-III. In analyzing these results, it is important to keep in mind the difference between statistical significance and practical significance. This is often a problem of statistical analyses when using large samples (thousands of points in this case) (see, *e.g.*, Devore (1982)). Many of the correlations are statistically significant. However, most of the correlations are small in absolute value and have marginal practical significance. The variables listed in Tables 4-III and 4-IV have been chosen because they have a correlation greater than 0.5 in absolute value, which indicates that they may have a strong practical significance as well as a statistical significance. Note that there are no correlations for point density and maximum gradient of the ground because the correlations for the variables are not of practical significance. For a complete list of the correlations, see Appendix A.

Although Table 4-III indicates a significant correlation of mean elevations with scan angle, this is not believed to be due to the classic “frown” artifact known to affect the cross-swath elevations of uncalibrated ALSM. The UF-ALSM is rigorously calibrated before major surveys. Also, the trajectory of the aircraft in this acquisition is closely aligned with the maximum ground elevation of the barrier island. Barrier islands usually display their highest elevations as a ridge just behind the beach paralleling shore.

Table 4-III. Correlations of practical significance for Site 1. All correlations are with respect to the scan angle. NPS indicates that the correlation is not of practical significance.

<b>Feature</b>	<b>Building</b>	<b>Tree</b>
Mean First Stop Elevation	-0.86	-0.80
Mean Last Stop Elevation	-0.86	-0.75
Difference between Mean First Stop and Ground Elevation	-0.54	NPS
Difference between Mean Last Stop and Ground Elevation	-0.53	NPS

A similar analysis is performed on Site 2. The plots of the geometrically-related variables are given in Figure 4-9. Once again there are more points toward the edges and there is a slightly more pronounced high frequency pattern in the number of first and last stop points although this time the pattern is along the trajectory. The correlations of practical significance between the 13 features and the 4 geometrically-related variables for Site 2 are shown in Table 4-IV.

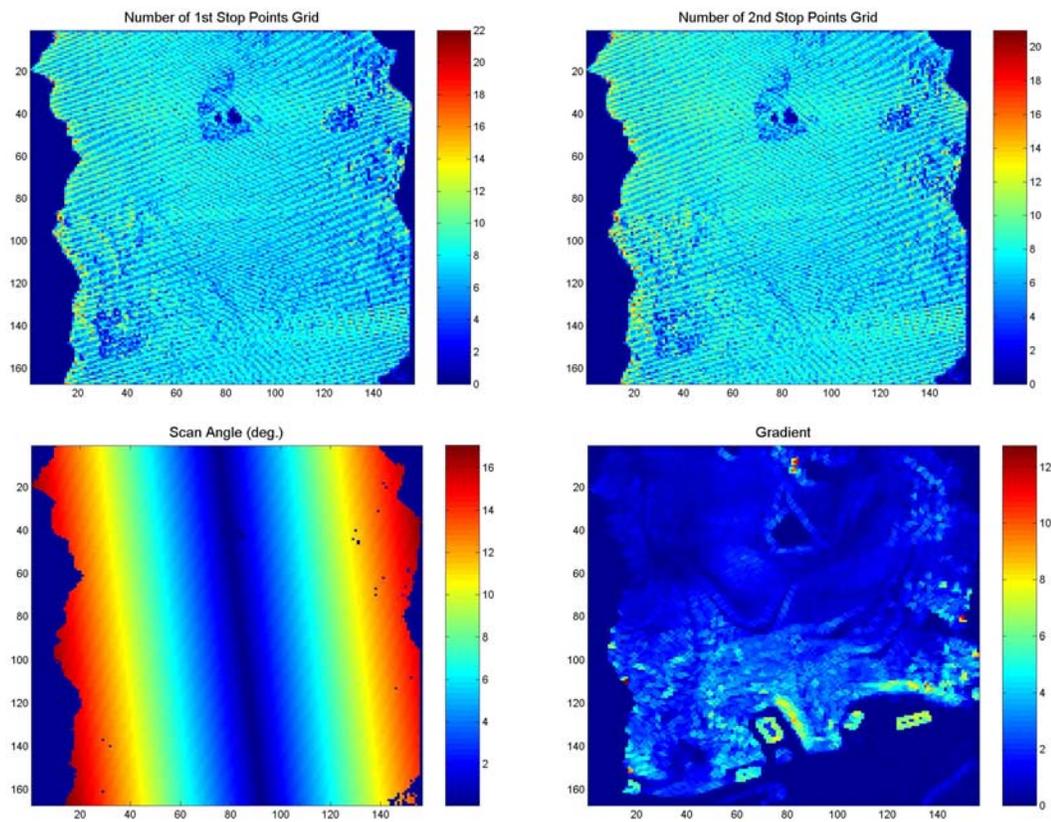


Figure 4-9. Plots of the geometrically-related variables for Site 2. The number density plots of the first and last stops are in the upper left and right respectively. The scan angle and the gradient plots are in the lower left and right respectively. The scan angle is measured in degrees and the gradient is measured in meters.

Table 4-IV. Correlations of practical significance for Site 2. All correlations are with respect to the scan angle. NPS indicates that the correlation is not of practical significance.

<b>Feature</b>	<b>Building</b>	<b>Tree</b>
Mean First Stop Elevation	NPS	-0.56
St. Dev. of Last Stop Elevation	NPS	-0.67
Difference in Stop Elevation	NPS	-0.68
Difference between Mean First Stop and Ground Elevation	NPS	-0.77
Difference between Mean Last Stop and Ground Elevation	NPS	-0.60
Difference between Mean First Stop and Low Last Stop Elevation	NPS	-0.70
Mean First Stop Intensity	NPS	0.69
Mean Last Stop Intensity	NPS	0.67

In general, looking at the plots of the number density for both Site 1 and 2, the number of points is surprisingly uniform. The largest variation appears to be related to known operational constraints of the ALSM system. This probably explains why there are so few practically significant results from these two variables.

While there are a few interesting relationships, the strongest correlations are between mean stop elevations and scan angle. These features are particularly important because mean elevations are used in the calculation of several of the other features. Understanding how features react to different geometries will make it easier to understand the results for other features. The reasons for these relationships are relatively straightforward as will be shown below.

### **Mean Elevation and Scan Angle**

The ALSM system is continually collecting data but the scan angle at which the data are collected is changing. When the ALSM system is observing objects at nadir ( $0^\circ$  scan angle), the first return is likely to be from the top of the object. As the ALSM system collects data from a greater scan angle, there is an increased likelihood that the first return will not be from the top of an object but will be from the side of the object instead. The same sort of relationship will hold for the last stop return as well, although

there will be lower limit to this effect, namely the ground. This relationship explains the negative correlation between mean elevations and scan angle. This relationship will hold for all objects with vertical extent above the ground and with sufficient inter-object separation to allow obliquely incident laser illumination to reach below the top levels. Thus it should hold for buildings and trees for most site locations.

Many of the elevation related features (differences in elevations, for instance) also have a negative correlation. The reasoning for this is much the same as for mean elevation. Since the first stop elevation (and also last stop elevation) will be lower, as scan angle increases, the difference in elevations will likely be smaller as the scan angle increases. This explains much of the variation seen in the elevations as a function of scan angle.

In general these relationships caused changes to the elevation-based features that are on the order of a few meters. This is entirely consistent given the correlation coefficients in Tables 4-III and 4-IV. For instance, with correlations for the elevation-based variables ranging from  $\sim 0.5$ – $0.8$  and the test data ranging over  $\sim 10^\circ$ , then change in the features should be on the order of 5–8 meters.

### **Mean Intensity and Scan Angle**

For Site 2, there is a significant correlation between mean intensity and scan angle for the tree class. However, this appears to be strongly influenced by the choice of training data. The test data do not have a uniform distribution of intensities. In particular, as will be shown in the Geographical Stability Section, there is particularly dark (low intensity) vegetation in the test data that helps to skew the results. When these data are removed from the analysis, the correlations drop noticeably and are no longer of practical significance.

While there are significant correlations between the scan angle and some of these features, it is important to note that the maximum scan angle points did not disproportionately occupy the overlap regions in the scatter plots for either elevation or intensity based features. This is probably due to the limited scan angle range of  $\pm 20^\circ$ . It is also an indication that the correlations of these features with scan angle actually have minimal influence on the separability of the classes. The correlations of all features versus the maximum gradient are found to be less than 0.5 and so are not listed in the tables.

## CHAPTER 5 MEDIAN-BASED MEASURE

### **Definition of Median-Based Measure**

Distance between two classes is one of the most standard measures in determining separability. Typically, separability is a function of both the intercluster spacing and the intracluster scatter. The distance is usually measured with a standard metric which should have four properties (Duda *et al.*, 2001): nonnegativity, reflexivity, symmetry, and the triangle inequality. The most commonly used metrics are some form of the Minkowski metric (Duda *et al.*, 2001)

$$L_k(\vec{a}, \vec{b}) = \left( \sum_{i=1}^d |a_i - b_i|^k \right)^{1/k} \quad (5-1)$$

where  $a_i$  and  $b_i$  are elements (coordinates) of vectors  $\vec{a}$  and  $\vec{b}$  and  $d$  is the number of data points. These metrics work well with Gaussian data, but unfortunately, ALSM data are not generally Gaussian. The ALSM data in feature space are often multimodal and rarely follow a normal distribution.

The mean and standard deviation, the two primary measures of location and spread of data, can be badly affected by multimodal data and therefore can be a poor estimator of the separation between two samples. An example of this is shown in Figure 5-1. In the upper figure, with two normal distributions that are well separated, the means and median are the same which the standard deviations and median absolute deviation (mad) are similar. In the lower figure, the distributions are still well separated. However, the separation, as indicated by the means and standard deviations of the two distributions,

show little separation while the median-based measures show more separation between the two distributions.

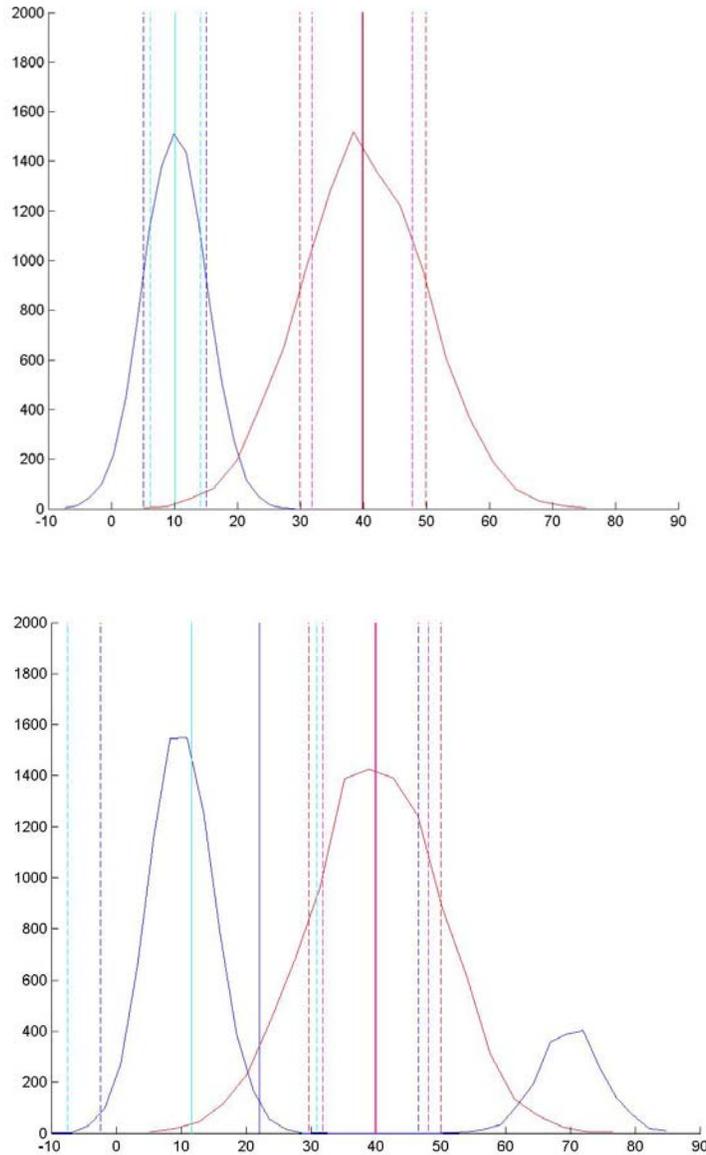


Figure 5-1. Illustration of how mean, standard deviation, median and median absolute deviation change from normal distributions to bimodal distributions. The measure of location of the center of the samples (mean, median) is given by the solid lines while the measure of spread (standard deviation and mad) is given by the dashed lines. The blue lines are the mean and standard deviation of the blue distribution while the cyan lines are the median and mad of the blue distribution. The red lines are the mean and standard deviation of the red distribution while the magenta lines are the median and mad of the red distribution.

A novel measure is created that is not reliant on the normality of the distribution of the data (Luzum *et al.*, 2004a, 2004b). The measure is similar to determining the significance of normally distributed data but uses the median, which is a statistic that is relatively unaffected by the outliers that are found in ALSM data:

$$d_{f_i} = \frac{\left| \text{median}(f_i^{\text{building}}) - \text{median}(f_i^{\text{tree}}) \right|}{\sqrt{(\text{mad}(f_i^{\text{building}}))^2 + (\text{mad}(f_i^{\text{tree}}))^2}} \quad (5-2)$$

where

$f_i^{\text{building}}$  is the measure of a feature  $i$  for all building pixels,

$f_i^{\text{tree}}$  is the measure of a feature  $i$  for all tree pixels,

$i$  is a number from 1 to 11 (to include all 11 features), and

$\text{mad}$  is the median absolute deviation, which is given by

$$\text{mad}(x) = \text{median}|x_i - \text{median}(x)|. \quad (5-3)$$

The distance  $d_{f_i}$  is a measure of the separation of the cluster centers in feature space relative to the spread of each of the clusters using measures that do not assume a Gaussian data distribution. This measure is unitless which is necessary when comparing features with different units. The minimum distance is zero but there is no upper bound on the maximum distance. Note, however, that this median-based measure makes no calculation of the amount of overlap between the two densities. It only calculates a distance between two classes with respect to the spread of their distributions. From empirical evidence, distances in excess of 3 indicate that the classes should be well separated and should therefore lead to successful data segmentation.

To determine the separability of features in a two-dimensional space, the following extension to the distance formula is created:

$$d_{ij} = \max(d_{f_i}, d_{f_j}) + \min(d_{f_i}, d_{f_j}) \frac{\left(2 - \sum_{k=1}^2 |\rho_k|\right)}{2} \quad (5-4)$$

where  $d_{f_i}$  is the one-dimensional distance between the two classes, tree and building, for feature  $i$  as defined above. Similarly,  $d_{f_j}$  is the distance for feature  $j$ , and  $\rho_k$  is the correlation between these two features for a single class  $k$ . For this research,  $k$  will run from 1 to 2, which indicate the tree and building classes. Note that this distance is computed between two features at a time so that for the 11 features chosen for this study, there will be 55 total distances computed.

This two-dimensional distance measure is an extension of the standard Manhattan (city-block) metric,  $L_1$ . It is unlikely that two features would be independent (orthogonal), so this measure has been modified to account for the potential correlation that exists between the two features. The measure has two desirable properties that lead to an intuitive notion of two-dimensional separability:

- If the data are completely redundant ( $\rho = 1$ ), then the second term is zero and so the distance is equal to the maximum of  $d_{f_i}$  and  $d_{f_j}$ . If no new information is provided with a second feature (*i.e.* the data are completely correlated), then the distance does not increase;
- If the data are completely uncorrelated ( $\rho = 0$ ), then the distance is the sum of the two one-dimensional distances.

In strict terms, this distance measure is not a metric since it does not meet the reflexivity property. The reason for this is that the distance measure is defined so that the distance between a class and itself cannot be computed directly; it can only be computed indirectly. However, these distance measures have the advantage of not being influenced by outliers. When dealing with ALSM data, this quality cannot be overemphasized.

Non-Gaussian data are prevalent in ALSM, making measures which assume a Gaussian distribution of the data, such as the use of the Dunn Index (Dunn, 1973, 1974) or the

Davies-Bouldin Index (Davies-Bouldin, 1979), of limited utility (Bezdek and Pal, 1998). The novel distance measures in (2) and (4) are not hindered by these problems.

### Analysis of Measure

The results of the one-dimensional distance measure between the building and tree classes for Sites 1 and 2 over the 11 features are provided in Table 5-I. Note that in this analysis, elevation is measured in meters while intensities can only be recorded as unitless numbers by the ALSM system.

Table 5-I. One-dimensional separability  $d_{f_i}$  between building and tree for each of the 11 features.

Feature	Site 1	Site 2
St. Dev. of first stop elevation	0.253	1.059
St. Dev. of last stop elevation	1.276	1.253
Difference between mean first and mean last elevation	1.015	1.001
Difference between mean first elevation and ground	0.372	1.754
Difference between mean last elevation and ground	0.050	1.447
Difference between mean first and low last elevation	2.609	1.074
Mean first stop intensity	3.621	0.794
Mean last stop intensity	2.988	0.260
St. Dev. of first stop intensity	2.784	2.609
St. Dev. of last stop intensity	3.211	2.483
Difference between mean first and mean last intensities	1.000	0.571

This distance measure works as expected in estimating the separation of the two classes. For features with large distance measures, graphically, the separability is good; for features with a small distance measure, the separability is poor. Table 5-I also shows some encouraging results. For example, the standard deviation of the first and last stop intensities appear to provide good separation between the two different classes for data in both the Florida and California test data. The four best results of the two-dimensional

distance are provided in Table 5-II. For a complete table of the median-based distances for all 55 feature pairs, see Appendix B.

Table 5-II. The four feature pairs with the greatest two-dimensional separability (in descending order) based on their average rank for all test areas. The numbers in parentheses are the ranks of the measure for that test area.

<b>Feature</b>	<b>Site 1</b>	<b>Site 2</b>
Mean First Stop Intensity and St. Dev. Last Stop Intensity	5.485 (2)	3.072 (10)
Mean First Stop Intensity and St. Dev. First Stop Intensity	4.391 (7)	2.926 (11)
Difference between Mean First and Low Last Stop Elevation and St. Dev. First Stop Intensity	4.819 (3)	3.262 (7)
St. Dev. First Stop Intensity and St. Dev. Last Stop Intensity	3.919 (12)	3.316 (6)

Scatter plots for the feature pair mean first stop intensity and st. dev. of last stop intensity for Sites 1 and 2 are shown in Figure 5-2. This feature pair has the best average ranking according to the median-based measure. While mean intensities have been used previously in classification (see, for example, Hug (1997)), standard deviation of intensities is not a common feature. Figure 5-2 indicates that good feature pairs provide considerable separation between the two classes. Also note in the plots that the data are distinctly non-Gaussian. To ascertain how these numerical separation distances relate to classification accuracies, confusion matrix values are given in the Geographic Stability and Temporal Stability sections.

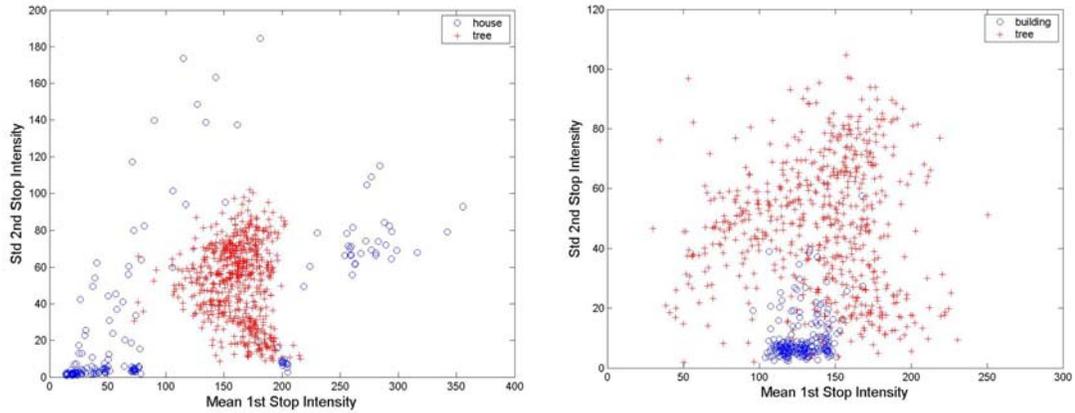


Figure 5-2. Scatter plots showing the separability using the feature pair mean first stop intensity and st. dev. of last stop intensity. This feature pair provides the largest average distance between classes for all test sites. Site 1 is on the left and Site 2 is on the right.

### Geographic Stability

In general, the median-based measure shows reasonable consistency across data from different geographic locations. The agreement between Sites 1 and 2 is not perfect but is still good. The primary reason that Site 2 does not agree more with Site 1 is that the effectiveness of two sets of features changes. For the Mare Island test area, the difference in mean first elevation and ground becomes noticeably more effective and mean intensities become less effective. This may seem counterintuitive, but the reason for these changes can be seen in a more detailed analysis of the test area. Plots indicating the vegetation training data and the elevation of objects in the test area are shown Figure 5-3.

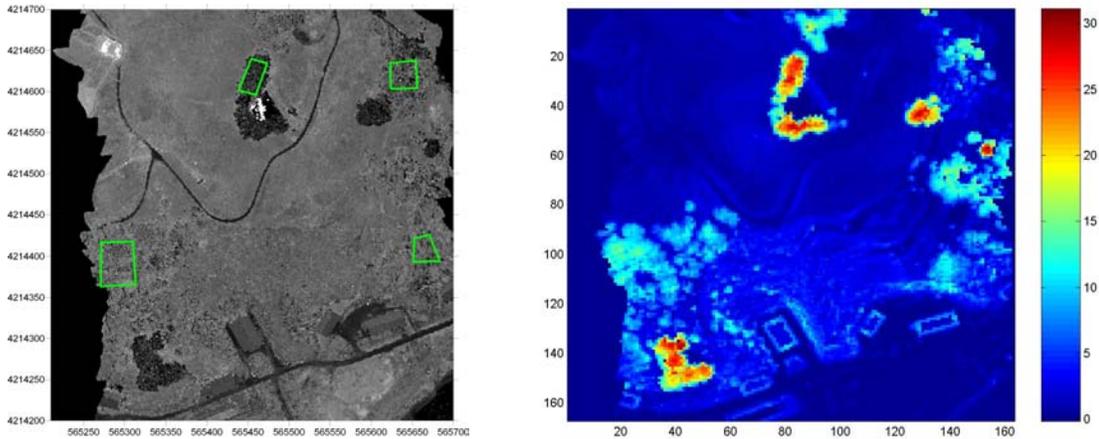


Figure 5-3. Plots of the vegetation training data on an intensity image (left) and the elevation of objects (right) from Site 2. Note that the elevations are measured in meters.

From Figure 5-3, it is clear that the vegetation data from one of the training data sets (the upper central area, hereafter called Area 1) are taken from a very different kind of vegetation than in the other three. The vegetation in Area 1 has noticeably darker intensity returns than the other three areas. In addition, the vegetation in Area 1 is roughly 30m high while in the other areas, the vegetation is roughly 10m high. These facts provide a strong indication that very different kinds of vegetation are included in this data set. This heterogeneity of vegetation data both in intensity and height impacted the distance between the two classes in the fashion mentioned above. Running a test case without Area 1 reduces the distance measures for the elevation-based features and increases the measures for the intensity-based features.

While this is unfortunate, it is encouraging to note that the mean intensities still provide some of the better features based on the rank of the distances. In spite of these disadvantages, the feature pair mean first stop intensity and st. dev. of last stop intensity is still one of the ten best measures for Site 2. This indicates that although different vegetation types may exist in an area, ALSM is still able to differentiate between building

and tree classes. Note also, that it would be possible to use two tree classes that represent two archetypes of trees (*e.g.* shorter deciduous *vs.* taller coniferous) to obviate this confusion, while still keeping requirements for *a priori* knowledge to a minimum.

It is important to note that although the ranking of the features differs somewhat depending on the site, it is possible for the test cases used in this analysis to find a feature space that led to large distances between the two classes. These large distances are associated with larger separations between classes and are therefore an indicator of success in classification. For a simple rule-based classifier (see Appendix C for more details) using the feature pair mean first stop intensity and st. dev. of the last stop intensity, it is possible to achieve 97% classification accuracy for the buildings and 98% for the vegetation for Site 1. For the same feature pair, it is possible to achieve 95% classification for the buildings and 96% for the vegetation for Site 2. Note that the better classification accuracy is achieved for the site with the greater distance for the mean first stop intensity and st. dev. of the last stop intensity feature pair.

### **Temporal Stability**

In order to realize a set of ALSM features that can be used to robustly segment trees and buildings across data sets, that feature set must provide separability for different terrain and landcover types (demonstrated in the previous section) as well as for different seasonal conditions. To evaluate the seasonal effects on the features, we examined Site 3 in May and November.

For surface mapping, the ideal case is for the data to remain temporally invariant so that mapping could take place during any season with consistent results. However, terrain conditions and landcover are not static. Vegetation, in particular, is temporally variant. Leaf on/leaf off conditions should change both elevation-based features and

intensity-based features since the effective laser cross section and the dominant spectral characteristics of the tree foliage will change. The analysis of the temporally displaced data confirms that the median-based measure can discern differences between data taken from the same area at different times of the year. In particular, the intensity-based measures provide the greatest discrimination between the temporally displaced data sets which is not surprising. Since much of the reflectance in the near infrared of the laser (1064 nm) is provided by the chlorophyll in the leaves (Maas, 1999), seasonal variations are to be expected. The two-dimensional median-based distances for the same training areas in Site 3 observed at two times of the year are shown in Table 5-III.

Table 5-III. The distances from Site 3 for the same four feature pairs with the greatest two-dimensional separability (in descending order) as shown in Table 5-II. The numbers in parentheses are the ranks of the measure for Site 3.

<b>Feature</b>	<b>Site 3</b>
Mean First Stop Intensity and St. Dev. Last Stop Intensity	1.333 (1)
Mean First Stop Intensity and St. Dev. First Stop Intensity	1.229 (4)
Difference between Mean First and Low Last Stop Elevation and St. Dev. First Stop Intensity	1.027 (12)
St. Dev. First Stop Intensity and St. Dev. Last Stop Intensity	0.941 (17)

A critical question is whether temporally displaced data (like vegetation is expected to be) will preclude discrimination with other classes (such as buildings). Since the building data is extremely sparse in Site 3, feature measurements from buildings in Sites 1 and 2 are used along with feature measurements from trees in Site 3 in order to simulate the potential for building/tree separation. Figure 5-4 shows a scatter plot using the features with one of the largest two-dimensional separations, namely mean first stop intensity and st. dev. of first stop intensity.

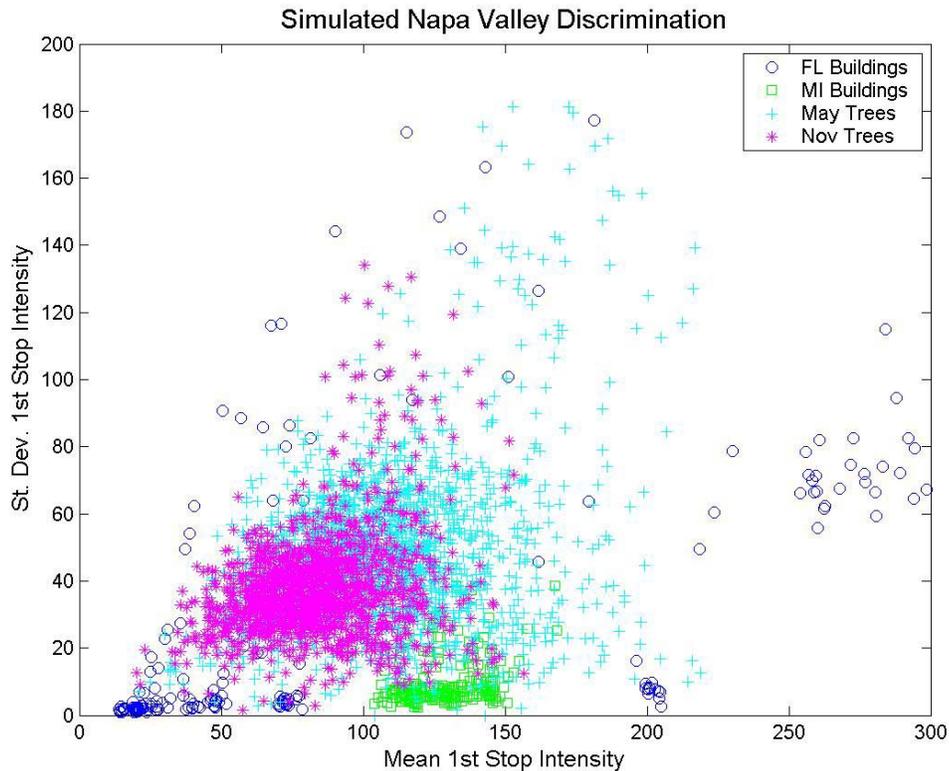


Figure 5-4. Data used to test discrimination potential for temporally displaced data in Site 3. All tree points are from Site 3. Buildings points from Site 1 (FL Buildings) and Site 2 (MI Buildings) are inserted due to the lack of buildings in Site 3. Rule-based classifiers can be built along the lines described in Rule C or using cutoffs at a mean first stop intensity of 225 and a st. dev. of first stop intensity of 15.

To test the separability of the simulated buildings in the vegetated area in Site 3, distances using the median-based measure are computed. Initially, feature measurements from the Flagler (FL) and Mare Island (MI) building data sets are combined to form one, heterogeneous building data set. The distance between this combined building data set and the May vegetation is 1.392 while the distance between the combined building data set and the November vegetation is 2.801. If only the Site 1 (FL) buildings are used in the simulation, the distances are 2.563 and 2.291, respectively. If only the Site 2 buildings are used, the distances are 2.355 and 4.681, respectively.

These numbers are comparable to the numbers shown in Table 5-II indicating that good separation is still possible even though the distribution of data in feature space is not temporally invariant. To test this, a simple rule-based classifier was constructed for the heterogeneous building data set and each of the two time variant vegetation data sets. The classification accuracies for the heterogeneous building and May vegetation data sets are 90% and 88%, for buildings and trees respectively and for the heterogeneous building and November data sets there are 93% and 95% accuracies for the buildings and trees. Note that the greater classification accuracy is found in the case where the median-based distance is greater.

It is interesting to note that this separation holds in spite of the fact that the building data taken from Site 1 and Site 2 are noticeably different. These differences, mostly in the mean first stop intensity, are due to the fact that the buildings from Site 1 and Site 2 have different roofing materials. This situation is not unexpected. For this particular simulation, using the Site 1 simulated buildings, whether the data were collected in May or November would make little difference. However, when using the Site 2 simulated buildings, better separation is achieved using data taken in November rather than the data taken in May. The classification accuracies are consistent with these distance measures. This indicates that even though separation is possible, the amount of separation and hence classification accuracy may be temporally variant. Again it is important to note that even in this less than optimal case, it is possible to find feature pairs that provided good distances between the different classes. The large distances indicated feature pairs with good classification potential.

### Test Case

As a test of these findings, a classification using the rules derived from the training data in Site 1 (see Appendix C) were applied to an area north of Site 1. The rules involve finding all non-ground points and then determining which of those had mean first stop intensities between 100 and 200 and standard deviation of last stop intensities between 8 and 103. All objects fitting these criteria were classified as trees while all other non-ground objects were classified as buildings. A view of the area and the classification results are shown in Figure 5-5.

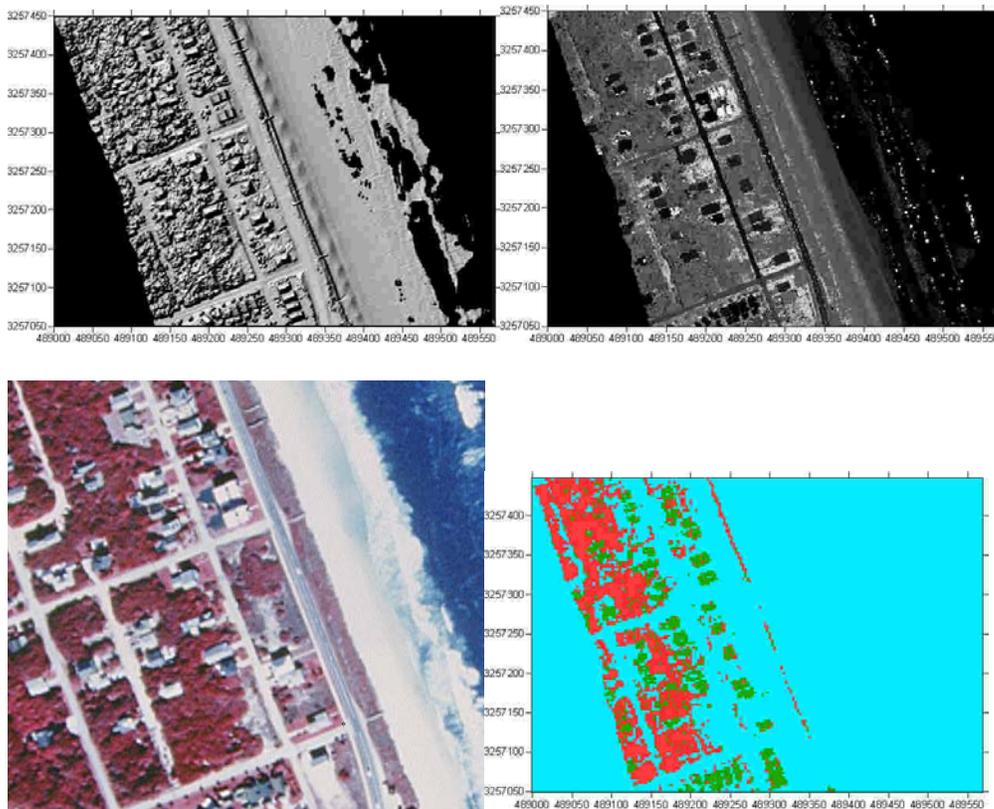


Figure 5-5. Plots of the training area. The shaded relief is in the upper left, the intensity plot is in the upper right, a color IR DOQQ is in the lower left, while the classification plot is in the lower right. In the classification plot, the ground is shown in cyan, buildings are shown in green and vegetation is shown in red.

The classifier appears to have worked reasonably well. In particular, it has performed extremely well on the buildings finding even those houses which are co-

mingled with the vegetation. It has also done reasonable well in determining vegetation, by classifying most of the vegetation in the suburban area correctly. Most of the cleared areas and the roads have been correctly identified as ground. The one area where the classifier appears to have failed is on the back side of the dunes which it classifies as trees. This classification results from the fact that the dunes appear to be above ground level and also appear bright with significant intensity variations. While this area may have some vegetation on it, it is not filled with trees.

### Multidimensional Distance Measure

Recall that the two dimensional distance measure is given by

$$d_{ij} = \max(d_i, d_j) + \min(d_i, d_j) \frac{\left(2 - \sum_{k=1}^2 |\rho_k|\right)}{2}. \quad (5-5)$$

This concept can be applied to higher dimensions. The math necessary to formulate the multidimensional distance is taken from the principle of inclusion-exclusion used in Set Theory. In two dimensions, this theory states that

$$|A \cup B| = |A| + |B| - |A \cap B| \quad (5-6)$$

where  $||$  denotes the cardinality of the set. In short, the size of the combined set made up of the elements of  $A$  and  $B$  will be the size of  $A$  plus the size of  $B$  minus the number of elements that  $A$  and  $B$  share. If this last term were not included, the points in both  $A$  and  $B$  would be counted twice.

The correspondence between this principle and the novel distance measure is straightforward. The new distance measure is trying to quantify the separability between two classes but must correct for information that is common to both one-dimensional distances. Whatever information is shared between the two one-dimensional distances

(correlation) does not provide additional separation and must be accounted for so that it is not counted twice in the distance calculation.

The multidimensional extension of the principle of inclusion-exclusion is given by (Leon-Garcia, 1994)

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{1 \leq i \leq n} |A_i| - \sum_{1 \leq i_1 < i_2 \leq n} |A_{i_1} \cap A_{i_2}| + \sum_{1 \leq i_1 < i_2 < i_3 \leq n} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| - \dots$$

$$(-1)^{n-1} \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_p}|. \quad (5-7)$$

This concept of compensating for what has been overcounted or undercounted can be applied to the multidimensional distance measure.

Without loss of generality, assume that there are three distances for features 1, 2, and 3 such that  $d_1 > d_2 > d_3$ . The extension of the distance measure to three dimensions is given by

$$d_{123} = d_1 + d_2 \left( 1 - \frac{\sum \rho_{12}}{2} \right) + d_3 \left( 1 - \frac{\sum \rho_{23} + \sum \rho_{13} - \frac{(\sum \rho_{13} + \sum \rho_{23}) \sum \rho_{12}}{4}}{2} \right). \quad (5-8)$$

In this equation,  $\rho_{ij}$  is the correlation between features  $i$  and  $j$  for a single class.

The first two terms of the equation are the same as in the two-dimensional distance formula. The third term is the extension to three dimensions. In form, the three dimensional extension is similar to the two-dimensional correlation correction. The first two correlation sums in the numerator are the attempt to remove the information from the third feature distance measurement that is redundant to distances from features 1 and 2. However, those summations will remove twice (doubly count) the correlation

information between features 1 and 2. To compensate for this, the third term adds back in this information.

While it is possible to expand the distance measures to multiple dimensions, it may not always be desirable to do so. As the dimensionality increases, it becomes more difficult to visualize the data. There is also a significant concern with diminishing returns at a cost of increased computational complexity. Due to the nature of the distance equations, the increase in separation between two and three dimensions will be smaller than the increase in separation between one and two dimensions. At the same time, the computational cost increases with each added dimension. The cost of computing the correlations increases as  $nN$  where  $n$  is the number of dimensions and  $N$  is the number of points needing to be correlated. For ALSM,  $N$  can be quite large. The correlation matrix necessary for these computations grows as  $n^2$ . Given these factors, two-dimensional distances are reasonable in most cases.

### **Separation Provided by More Standard Features**

There are several features that have been used in the past to try to distinguish between classes of objects. Two that have been used with varying degrees of success are differences in elevations between stops and standard deviations of first stop elevations. It seems intuitive that these features would provide information on whether a grid cell is a building or a tree since buildings would be expected to be a hard target (objects with a solid surface) while trees are expected to be a soft target (objects with a non-solid surface). It is informative to look at these features and compare the results to some of the other features that have larger median-based distance measures and therefore would, *a priori*, seem to be better features to use for classification.

One of the best feature pairs is mean first intensity and standard deviation of last stop intensity. This feature pair provides the largest average distance between classes for all test sites when using the median-based measure and the second best when using the information-based measure. A scatter plot of these two features for the Flagler County, FL (Site 1) data set and the Mare Island, CA (Site 2) data set are shown in Figure 5-1.

A more traditional feature pair is difference between mean first and last stop elevation and standard deviation first stop intensity. This feature pair takes advantage of the determination of whether a data return is from a hard target or a soft target and uses the intensities of the return to discriminate between the two classes. A scatter plot of these two features for the Flagler County, FL data set and the Mare Island, CA data set are shown in Figure 5-6.

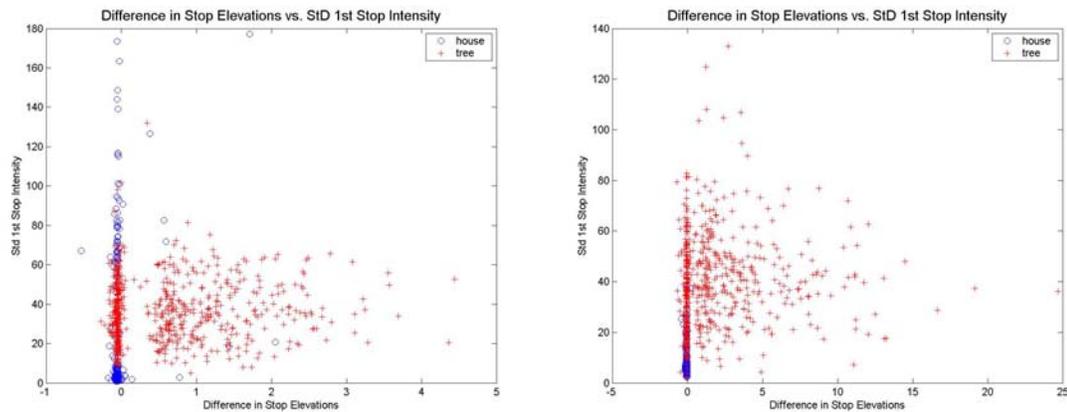


Figure 5-6. Scatter plots showing the separability using the feature pair difference between mean first and last stop elevation and st. dev. of first stop intensity. Site 1 is on the left and Site 2 is on the right

Notice that there is a considerable amount of overlap between the two classes using this feature pair. In particular, this appears to be caused by the large percentage of trees which do not appear to have true last stops. Even though nearly all buildings have a

small difference in stop elevations, several of the trees apparently have canopies thick enough to prevent a last stop and so are also intermingled with the building data.

Based on the plot, classification accuracy would be expected to be lower. Using a rule-based classifier, for the Florida data set, accuracies of 99% was achieved for the buildings and 59% was achieved for the trees. For the California data set, accuracies of 99% was achieved for the buildings and 85% was achieved for the trees.

Another traditional feature pair is standard deviation first stop elevation and difference between mean first and mean last stop elevation. This feature pair uses two common elevation-based features used for the discrimination of the two classes. A scatter plot of these two features for the Flagler County, FL data set and the Mare Island, CA data set are shown in Figure 5-7.

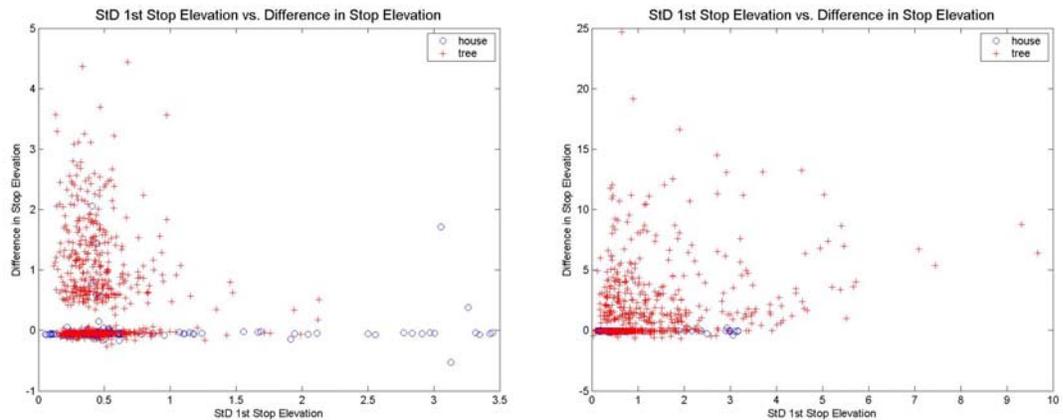


Figure 5-7. Scatter plots showing the separability using the feature pair st. dev. of first stop elevation and difference between mean first and last stop elevation. Site 1 is on the left and Site 2 is on the right

Again, there is a considerable amount of overlap between the two classes using this feature pair. In particular, this appears to be caused by the large percentage of trees which do not appear to have true last stops. This causes both the difference in stop elevations as well as the standard deviation of the first stop elevations to be small for the

vegetation. These characteristics, which are indicative of a hard target, overlap with the characteristics expected from the buildings.

Based on the plot, classification accuracy is again expected to be lower. Using a simple rule-based classifier, for the Florida data set, accuracies of 94% was achieved for the buildings and 58% was achieved for the trees. For the California data set, accuracies of 100% was achieved for the buildings and 59% was achieved for the trees.

As mentioned in the section on Definition of Median-Based Measure, the median-based measure cannot directly estimate the amount of overlap between two classes in feature space. It indirectly estimates this by looking at statistics calculated on the two classes. The advantage of this is that in many cases, this indirect calculation appears to be reasonably close to what would be expected by a robust calculation and this method is computationally efficient. This point, however, does raise the question as to whether direct calculations of the overlap might be more exact and therefore lead to more accurate estimates of separability between classes. This is the concept behind relative entropy (also known as the Kullback-Leibler distance) which is given by (Duda *et al.*, 2001)

$$D_{KL}(p(x), q(x)) = \sum_x q(x) \ln \frac{q(x)}{p(x)} \quad (5-9)$$

where  $p(x)$  and  $q(x)$  are two probability distributions. This idea will be developed more fully in Chapter 6 through the use of mutual information.

### Summary of Standard Features

For this analysis, the rule-based that was used to discriminate between building and trees used the idea that buildings are hard targets and therefore should exhibit little change in elevations between stops and also should have small standard deviations in the first stop elevations. Using this idea, the classifier correctly identifies nearly all building

cells since most buildings follow the ascribed rules. It also shows some proficiency at identifying the grid cells with a difference in mean stop elevations with the tree class.

Where the classification breaks down is with the considerable number of tree cells (~30–40%) which do not have true last returns. This lack of last return causes the tree cells to appear as hard targets and so the cells are assigned the classification of ‘building.’ This fact causes nearly all of the misclassification seen in this analysis.

Even though this lack of true last return is unfortunate, it is not entirely unexpected. The canopies of trees and the optical path of the laser pulse will not always combine in such a way that good last returns will be received from laser pulses fired into vegetation. It is also important to remember that a true last stop will not be received if the difference between the stops is less than 2.5m. Any last return that is less than 2.5m will not even be detected because of this hardware limitation. Therefore, the current ALSM system will be insensitive to many of the small scale variations that exist in the tree crown which could theoretically be used to differentiate between buildings and trees. To sense thick trees as a diffuse medium while using a small footprint (as we have), a laser shot must enter a gap in the canopy. Thus the notion of “trees are soft targets” only holds for large footprints or for comparing the spread in first stop returns over a large grid cell. There is an inherent scaling dependence.

In a more realistic situation, the problem in classification will not necessarily be isolated to the trees. For now, assume that all buildings roofs are flat and have no objects protruding from them. Even in this simplified case, the number of pure building grid cells can be as large as 100% (unlikely) down to as little as roughly 25% depending on the size and shape of the buildings, grid cell size, and the circumstances of the data

collection (see Figure 5-8). Though it is not a serious concern for these particular data because of the way the training data were chosen and the size of the grid cell, in most real-world applications, a significant proportion of buildings cells will not be hard targets. Sloped roofs and objects like chimneys will only further reduce the number of pure planar building pixels.

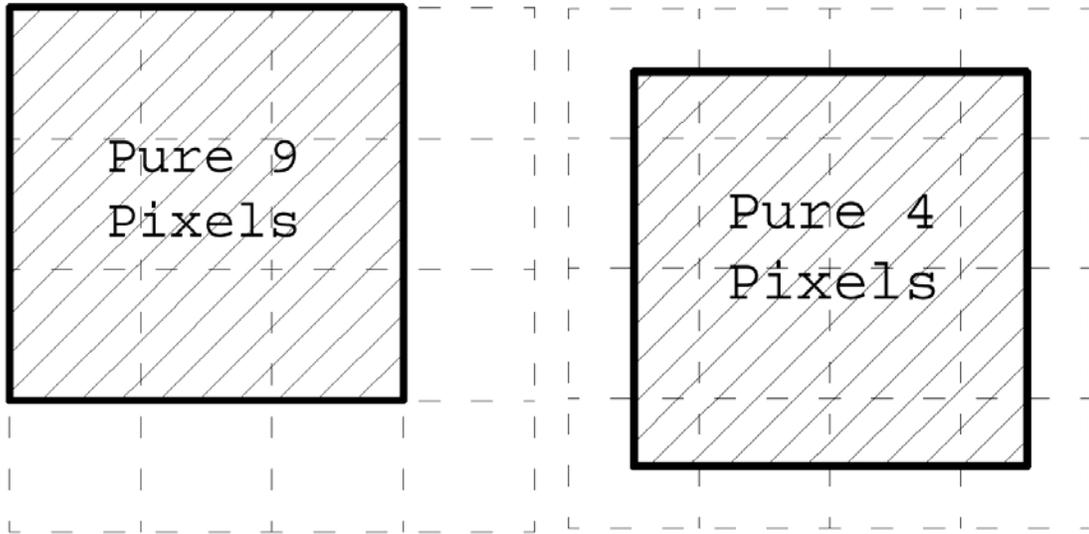


Figure 5-8. Illustration showing how the number of pure pixels changes depending on the placement of the object on the grid. Diagram by A. Mamas.

The traditional features have some merit. Note that the building classification accuracy of the mean first and last stop elevation and st. dev. of first stop intensity and st. dev. of first stop elevation and difference between mean first and last stop elevation feature pairs is actually better than in the first (best) feature pair. These feature pairs are good at identifying building but at the cost of misclassifying trees as buildings. This lack of discrimination hurt these feature pairs' classification ability.

It is interesting to note that the difference in mean stops appears to work better in California than in Florida. One potential reason for this is that the vegetation in Florida is expected to be thicker because the climate is more tropical while the vegetation in

California should be less dense because the climate is more moderate. The sparser vegetation in the California site may allow for more true last stops which will improve this features' capability at discriminating between buildings and trees.

## CHAPTER 6 MUTUAL INFORMATION BASED MEASURE

The concept of information theory was first presented in Shannon's seminal work (Shannon, 1948). Information theory is a codified method of determining how much is known about the answers to questions. Often, information is determined through the use of entropies which are a measure of the inherent uncertainty. Mutual information can be computed for two or more random variables and it is a measure of how much the uncertainty in one variable is reduced by knowledge of the other variable(s). See, Cover and Thomas (1991) for a more thorough description of mutual information. Mutual information can be computed as (Duda *et al.*, 2001)

$$I(p; q) = H(p) - H(p | q) = \sum_{x,y} r(x, y) \log_2 \frac{r(x, y)}{p(x)q(y)} \quad (6-1)$$

where  $p$  and  $q$  are random variables,  $H(p)$  is the entropy of variable  $p$ ,  $H(p | q)$  is the conditional entropy of  $p$  given  $q$ ,  $r(x, y)$  is the joint probability distribution of finding values  $x$  and  $y$ ,  $p(x)$  is the probability distribution function (pdf) of  $p$ , and  $q(y)$  is the pdf of  $q$ .

The use of mutual information for this analysis is based on the formulation presented in Grall-Maës and Beuseroy (2002) and was first implemented in Luzum *et al.* (2004c). Let  $X$  be a random variable describing one or more features observed and let  $C$  be a random variable that describes the class of the object. Then the mutual information can be written

$$I(C; X) = H(C) - H(C | X) \quad (6-2)$$

where  $H(C)$  is the entropy of class  $C$  and  $H(C|X)$  is the conditional entropy of class  $C$  given  $X$ . For discretely sampled variables, the entropies can be written as

$$H(C) = -\sum_C P_C \log_2 P_C \quad (6-3)$$

and

$$H(C|X) = -\sum_C P_C \sum_x p_C(x) \log_2 \frac{P_C p_C(x)}{p(x)} \Delta x \quad (6-4)$$

where  $P_C$  is the estimated pdf of class  $C$ ,  $p_C(x)$  is the estimated conditional pdf of class  $C$  given  $x$ ,  $p(x)$  is the total pdf of  $x$ , and  $\Delta x$  is the increment between the samples of  $X$ . There is no theoretical way to choose  $\Delta x$ , so a value should be chosen empirically so that the separation between distributions (if it exists), is preserved. As an order of magnitude estimate,  $\Delta x$  should be no more than one-tenth of the range of  $x$ .

For this analysis  $P_C$  can be well determined from the training data. Given that the number of points that are building training data and the number of points that are tree training data are known *a priori*, the ratio  $P_C$  for each class  $C$  can be computed directly. However, in general,  $P_C$  must be estimated by sampling the data and using the ratios determined from the sampling as the *a priori* estimates.

Both  $p_C(x)$  and  $p(x)$  are estimated from the training data. To compute  $p_C(x)$ , the pdf for each class will need to be computed separately. These data are distinctly non-Gaussian and so a method must be used which makes no assumption about the distribution of data such as a Parzen-window approach. A pdf can be estimated using (Duda *et al.*, 2001)

$$p(\vec{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{V} \varphi_{x,h} \left( \frac{\vec{x} - x_i}{h} \right) \quad (6-5)$$

where  $\vec{x}$  are the grid nodes at which the pdf will be estimated,  $x_i$  are the coordinates of the data in feature space,  $n$  is the number of data points in feature space,  $h$  is the length of the edge of the hypercube,  $V$  is the volume of the hypercube and  $\varphi_{x,h}(x_i)$  is a Parzen window function that is 1 if the data point falls within the hypercube and 0 otherwise.

The Parzen windows approach is straightforward in concept. In practice,  $\vec{x}$  is chosen so that they cover the part of feature space where analysis will take place. For instance, feature space with only isolated points is often uninteresting because it contains outliers and is not an area where segmentation of data is an issue and so it is often safe to not estimate the pdf over this area. The reason for minimizing the expanse is that the computational complexity increases as the size of  $\vec{x}$  increases.

In addition to  $\vec{x}$ , the parameter  $h$  must also be chosen. This parameter determines the size of the hypercube and the choice in  $h$  can influence the shape of the estimated pdf. If  $h$  is too large, the estimated pdf tends to be overly smooth while if  $h$  is too small, the pdf to be too spiky (Duda *et al.*, 2001). Because there is no procedure for determining the optimal value of  $h$ , it is necessary to determine this parameter empirically. The best  $h$  to use is one that preserves the details of the distribution of data without creating artifacts that appear to be caused by the sampling of feature space. As an order of magnitude estimate,  $h$  should be no more than one-tenth of the range of  $\vec{x}$ .

Once the input parameters have been chosen has been chosen, the Parzen window procedure is applied to the data in feature space. For every point  $x'$  in  $\vec{x}$ , the window function determines how many data points from data set lie within the hypercube

centered at  $x'$ . By normalizing this count for the volume of the hypercube, it is possible to estimate the pdf by dividing this normalized count by the total number of points in the feature space data set.

The features used in this analysis differ in scale by roughly an order of magnitude. A number of different parameter combinations were tested to ensure that the probability distribution functions are reasonable. None of the combinations was entirely acceptable until the Parzen window approach was modified to use a “hyperbox” instead of a hypercube. This allowed for the length of each edge to be based on the span of the data. This modification allowed for more realistic probability distributions particularly in feature pairs where one feature was elevation-based and the other was intensity-based. In practice, the edge of the “hyperbox” was estimated by scaling  $\Delta x$  since this increment was chosen so that the number of bins in across the feature space would be roughly constant.

The edge length of this hyperbox was determined empirically by repeated testing. The edge ( $h$ ) used in the analysis is  $h = 10 * \Delta x$  where  $\Delta x$  is the differential width used in the calculation of the mutual information. This number is particularly appropriate because the differential width was chosen so that the number of bins along each axis would be approximately the same for each of the 11 features studied.

Throughout the trials using different parameters for the Parzen window approach, the rank of the best features showed minimal change. The results from the parameter set which created the best distribution functions also seems to produce the most intuitive mutual information results. Throughout the trials using different parameter sets, Site 1 showed minimal change. Site 2 showed greater variability over the parameter sets. From

the scatter plots in Figure 5-2, the separation between the two classes in Site 1 and 2 appears to be comparable therefore the mutual information would be expected to be very similar. The final results appear to produce mutual information results which are rather consistent between Sites 1 and 2.

Using different parameter sets, Site 3 showed the largest change. From the scatter plots in Figures 5-2 and 5-4, it is expected that the mutual information in Sites 1 and 2 would be greater than for Site 3. The results of the final parameters Site 3 appear to be consistent with the amount of overlap seen in the scatter plots in feature space.

There are several factors that contribute to the uncertainty in the computation of mutual information. These uncertainty factors include the uncertainty in the data, the uncertainty in the distribution and the uncertainty in the computation of mutual information.

The uncertainty in the data arises from the fact that no data measurement is perfect. There will be both systematic errors and random errors in the measurements. The systematic errors are reduced to a negligible amount through the use of calibration. Therefore, the random component dominates the measurement errors. In the UF ALSM system, the uncertainty in the elevation data is at the 10 to 15 cm level.

The uncertainty in the distribution will arise from two factors, the uncertainty in the representativeness of the data and the uncertainty in the estimation of the probability distribution function (pdf). The question of representativeness is one of whether the data points chosen for your test data are representative of the classes being studied. In essence, this is a question of statistical sampling. In this analysis, there is no analytical equation that can be used either in the computation of the pdf or the uncertainties. Given

the possible diversity of classes, this uncertainty can be computed through a Monte Carlo approach by making multiple trials of choosing training data across an expanse of area and calculating the variation of features in feature space.

The uncertainty in the estimation of the pdf arises from the fact that the pdf may not be exactly determined from the data if the pdf is computed numerically instead of theoretically. In this analysis, a Parzen windows approach is used to estimate the pdf and there are two parameters that must be specified for the procedure to work. Variations in the choice of these parameters will result in differences in the pdf. There is no way to compute an optimal pdf using Parzen windows. Unfortunately, since poor choices in parameters can make the pdf arbitrarily bad, there is no way to rigorously estimate an uncertainty. Several “reasonable” pdfs can be calculated and the variations between these pdfs can be computed, but this will hardly produce a robust measure of the uncertainty in the estimation of the pdf. Another method of determining the validity of the pdfs is through the use of cross-validation. By using different combinations of training data, a set of pdfs can be created. By averaging these results, a more reasonable estimate of the pdf can be created. It is important to remember that while there are negative aspects to using a Parzen windows approach, the benefit is substantial. This approach frees the researcher from assuming that the data being analyzed fit a prescribed analytical model exactly. Without this freedom, much of this research would not have been possible.

There is also an uncertainty associated with the computation of the mutual information. However, this uncertainty can be made arbitrarily small. Since the computation of mutual information amounts to a calculation analogous to a numerical

integration, the error in this calculation can be reduced to an arbitrarily small number simply by reducing the width of the interval of integration (say, for example,  $\Delta x$ ) to an appropriately small number. In the case of an infinitesimally small width of the interval of integration ( $dx$ ), the computational error would be reduced to the round-off error in the computer.

The use of mutual information in the determination of the separation between classes is an elegant approach that has many benefits. This method is rigorous in that the exact amount of overlap between classes is estimated directly from the data in feature space. It is not necessary to make indirect computations through the use of statistics and/or assumptions about the distribution of the data. However, this rigor comes at a cost. The computational complexity of these procedures is significantly greater than for a method such as the median-based distance measure. As an example, when analyzing these data using the median-based measure, the computer programs would run for approximately 5 minutes. When performing the same analysis using mutual information, the computer would run for roughly 2 hours. This increase in computational complexity could play a significant role if the data set to be analyzed is large.

Strictly speaking, mutual information is independent of the number of data points in the training set since mutual information is only computed through the pdfs. However, realistically speaking, the computation of the pdfs will be dependent on the size of the training data set. If the data have an easily describable distribution, then the computational complexity of the pdf determination will grow linearly with the number of data points. However, with a method for determining arbitrary distributions, such as Parzen windows, the number of data points is not nearly as important as the range of the

data points and the fine structure apparent in the distribution. To estimate a pdf using Parzen windows, the number and distribution of the training data throughout the entire feature space needs to be calculated. Therefore, depending on the training data, as the number of points increases, the size of the distribution will increase and the computational complexity will increase.

### **Analysis of Information Gain**

By measuring directly the amount of information each feature or feature pair tells us about the building and tree classes, it is possible to quantitatively determine the separability between the classes. In this case, we know exactly how much information is needed to determine the class because for this analysis, the “question” being asked is whether the cell being queried is a building cell or a tree cell. For this binary question, the total amount of information possible is 1 bit.

The mutual information was estimated for each of the single features using the formulation presented in Description of Information Theory section. A Parzen window approach was used to estimate the probability distribution functions. The results for the single features are given in Table 6-I.

Table 6-I. Mutual information between the classes for each of the 11 features between building and tree classes. Information is being measured in bits.

Feature	Site 1	Site 2	Site 3
St. Dev. of first stop elevation	0.027	0.182	0.019
St. Dev. of last stop elevation	0.118	0.319	0.018
Difference between mean first and mean last elevation	0.146	0.194	0.008
Difference between mean first elevation and ground	0.194	0.218	0.006
Difference between mean last elevation and ground	0.153	0.156	0.014
Difference between mean first and low last elevation	0.16	0.251	0.009
Mean first stop intensity	0.553	0.294	0.193

Table 6-I—Continued.

Mean last stop intensity	0.543	0.26	0.144
St. Dev. of first stop intensity	0.4	0.473	0.113
St. Dev. of last stop intensity	0.345	0.497	0.16
Difference between mean first and mean last intensities	0.206	0.305	0.079

Table 6-I shows that the mean intensities and the standard deviation of the intensities do a reasonably good job of separating between the two classes across all three test sites. While the utility of intensities in classification is expected (Hug, 1997), the standard deviation of the intensities is not a feature that is in general use.

For Sites 1 and 2, the largest possible mutual information is desired to maximize separability. Since  $I(C; X) \in \{0,1\}$ , this means we expect “good features” to yield  $0 << I < 1$ . However, in Site 3, we find the mutual information for a given feature is generally smaller, than for Site 1 or 2. This is entirely desirable because large values for mutual information in Site 3 would imply that the vegetation changed dramatically between the two seasons, with respect to the ALSM features, and therefore that time of year for ALSM acquisitions would have to be factored into any subsequent landcover classification or analysis. However, mutual information values very close to zero would imply that seasonal changes in the canopies are virtually undetectable with ALSM. This would in turn imply that ALSM is not a useful data type for monitoring seasonal variations. The modest values we see for mutual information of Site 3 in Table I imply that seasonal variations can in fact be detected, but they do not appear to be so great as to severely hinder classification.

It is reasonable to expect that if a single feature provides good separation, that feature pairs will provide better separation. Mutual information was estimated for each of the 55 possible feature pairs using the formulation presented in Description of

Information Theory section. Table 6-II provides the mutual information for the four best feature pairs. For a complete list of mutual information for all 55 feature pairs, see Appendix E.

Table 6-II. The four feature pairs with the largest average ranking of mutual information between building and tree classes. The numbers in parentheses are the ranks of the measure for that test area and the information is being measured in bits.

Feature	Site 1	Site 2	Site 3
Mean First Stop Intensity and St. Dev. Last Stop Intensity	0.622 (2)	0.679 (2)	0.295 (2)
Mean Last Stop Intensity and St. Dev. Last Stop Intensity	0.615 (5)	0.684 (1)	0.3 (1)
Mean First Stop Intensity and St. Dev. First Stop Intensity	0.62 (3)	0.668 (3)	0.292 (3)
Mean Last Stop Intensity and St. Dev. First Stop Intensity	0.626 (1)	0.652 (4)	0.253 (7)

The rankings of the top four feature pairs are remarkably consistent considering that there are 55 possible feature pairs. Scatter plots of the best feature pair can be found in Figures 5-2 and 5-4

It is reasonable that Sites 1 and 2 have larger mutual information values than Site 3. To reiterate, in Sites 1 and 2, the mutual information is being computed between the building and tree classes while in Site 3, it is being computed between the May tree and the November tree classes. The small mutual information value expected from theory is borne out in the Figure 5-4 and Tables 6-I and 6-II.

### **Comparison of Mutual Information to Median-Based Measure**

This analysis uses the same features and the same input data as the median-based measures. It is informative to look at how the best features as measured using mutual information compare to the results using a median-based measure. Table 6-III shows the

rankings of the best separations based on mutual information and compares them to results using a median-based measure over the same features using the same input data.

Table 6-III. The rankings of the four feature pairs (out of 55) with the largest average ranking of mutual information between building and tree classes. These are the same feature pairs presented in Table II. The first two columns of numbers give the rankings according to mutual information for Sites 1 and 2 while the fourth and fifth columns give the rankings using a median-based measure.

<i>Feature</i>	<i>Mutual Information</i>			<i>Median-Based</i>		
	<i>Site 1</i>	<i>Site 2</i>	<i>Average</i>	<i>Site 1</i>	<i>Site 2</i>	<i>Average</i>
Mean 1st Stop Intensity and St. Dev. 2nd Stop Intensity	2	2	1	2	10	1
Mean 2nd Stop Intensity and St. Dev. 2nd Stop Intensity	5	1	2	9	20	8
Mean 1st Stop Intensity and St. Dev. 1st Stop Intensity	3	3	3	7	11	2
Mean 2nd Stop Intensity and St. Dev. 1st Stop Intensity	1	4	4	5	18	6

The rankings between the mutual information and a median-based measure are reasonably consistent. The feature pair containing the mean 1st stop intensity and st. dev. of the 2nd stop intensity has the best average rank as measured by mutual information (see Table II) and as determined by the median-based measure (see Luzum et al., 2004b, Table IV). In addition, the feature pair with the third best average ranking based on the mutual information separation (mean 1st stop intensity and st. dev. of the 1st stop intensity) has the second best average ranking when using the median-based measure. Note that the four feature pairs with the best average ranking as determined by mutual information all have average ranks in the top ten (out of 55 possible) as determined by the median-based measure.

In particular, the mutual information based ranks for Sites 1 and 2 and the median-based ranks of Site 1 show good agreement. For Site 2, there are known concerns with

homogeneity of the vegetation training data chosen (Luzum *et al.*, 2004b) which causes the rankings of the median-based measure to differ from Site 1. It is interesting that mutual information-based separation is not as severely altered by the heterogeneity of the vegetation but instead produces very consistent results in both Site 1 and 2.

### **Effects of Grid Cell Size**

The initial grid cell size is chosen as small as possible to still be able to compute meaningful statistics for the features. This choice was made in order to maximize the possibility of obtaining pure pixels in the analysis. This decision allows the analysis to take advantage of the precise planimetric and ranging capability of ALSM by not averaging the measurements over an area larger than necessary. However, it is not obvious that the choice of small grid cells will necessarily produce the best results in the separability of the classes. For instance, increasing the grid cell size incorporates contextual information because it is aggregating information from surrounding pixels.

In particular, there is concern that any feature or feature pair that uses the concept that buildings are hard targets (negligible penetration of the last return beyond the first return) and that trees are soft (diffuse) targets is likely to suffer from the fact that over a small area, trees can appear to be hard targets. To sense thick trees as a diffuse medium while using a small footprint (as we have), a laser shot must enter a gap in the canopy. Thus the notion of “trees are soft targets” only holds for large footprints or for comparing the spread in first stop returns over a large grid cell. There is an inherent scaling dependence. This problem is only exacerbated by the fact that the ALSM system used to collect the data only registers a true last return if the last return pulse is from a surface that is greater than 2.5m away from the first.

In addition to the argument above based on structure in Euclidian space, an argument can also be made using feature space. Although small grid cells should lead to purer samples, they also contribute to the non-Gaussianity of the clusters because we have chosen to only have one cluster for vegetation and one for buildings. This was desired because it minimizes the *a priori* knowledge required (e.g. it is not necessary to determine the number of special subclasses that exist within each class when clustering ALSM data. But as a result, the ‘trees’ class contains many different kinds of trees and will thus tend to appear multi-modal (non Gaussian). It is plausible that larger grid cells will reduce this multi-mode appearance in the clusters.

This concern about the effect of the scaling dependence is tested by reanalyzing the data using different grid cell sizes. The grid cell size is changed from the original 3m on a side to two other cases, 5m and 10m on a side. The mutual information for every feature and feature pair is recomputed for each of the different grid sizes. The results for the four best features are presented in Table 6-IV while the results for the four best feature pairs are presented in Table 6-V. For a complete list of mutual information for the 3m, 5m, and 10m grid cell sizes, see Appendix F.

Table 6-IV. The mutual information of the four best features when using the buildings and tree classes. The analysis at Sites 1 and 2 is made with grid cell sizes of 3m, 5m, and 10m. The mutual information is measured in bits.

Feature	Site 1			Site 2		
	3m	5m	10m	3m	5m	10m
Mean first stop intensity	0.55	0.58	0.60	0.29	0.30	0.35
Mean last stop intensity	0.54	0.56	0.59	0.26	0.21	0.22
St. Dev. of first stop intensity	0.40	0.43	0.49	0.47	0.63	0.58
St. Dev. of last stop intensity	0.35	0.34	0.41	0.50	0.63	0.67

Table 6-V. The mutual information of the four best feature pairs when using the buildings and tree classes. The analysis at Sites 1 and 2 is made with grid cell sizes of 3m, 5m, and 10m. The mutual information is measured in bits.

Feature	Site 1			Site 2		
	3m	5m	10m	3m	5m	10m
Mean First Stop Intensity and St. Dev. Last Stop Intensity	0.62	0.64	0.64	0.68	0.73	0.74
Mean Last Stop Intensity and St. Dev. Last Stop Intensity	0.62	0.63	0.64	0.68	0.74	0.73
Mean First Stop Intensity and St. Dev. First Stop Intensity	0.62	0.64	0.64	0.67	0.74	0.73
Mean Last Stop Intensity and St. Dev. First Stop Intensity	0.63	0.63	0.64	0.65	0.72	0.72

The four features presented in Table IV are substantially better than the other seven features studied in this analysis. The best features are roughly 0.2–0.3 better than the other features. For the feature pairs, the gap between the four listed in Table V and the other feature pairs is smaller with a difference of roughly 0.01–0.3 less than the best. However, even with the best features and feature pairs, there is still some overlap between the classes which causes the mutual information to be less than 1.0.

In general, the trend seen for most of the entries in these tables holds for most of the features and feature pairs, namely that as the grid cell size increases, the mutual information increases. This indicates that the larger grid cell size improves the separability between the two classes. This is due probably to the greater number of laser shots in each grid cell allowing the ALSM system to better estimate the features being studied. This result also seems to indicate that the better results could be obtained for a given grid cell size if the density of points was increased.

It is important to note that in Tables 6-IV and 6-V, the better feature and feature pairs maintain their high rank regardless of the grid cell size. This consistency is an

indication that the rankings of features and feature pairs are relatively invariant to the grid cell size.

While the larger grid cells may improve the separability of classes, these larger cells do not necessarily improve the classification in a real-world situation. The larger grid cells will inevitably lead to more mixed cells than would be obtained by using the smaller cells. Clearly, mixed cells will make classification more difficult and should be avoided when possible. This point also indicates that a higher density of laser points on the ground would probably produce more acceptable classification results.

### **Stability**

It is important to note in each of the test sites, it is possible to find features and feature pairs that provided good mutual information and therefore good separation between the classes. However, one of the most important considerations is whether the features and feature pairs are geographically and temporally stable. Stability offers the prospect that feature pairs that are ‘good’ for one data collection are likely to be ‘good’ for the next regardless of the location or time of year of the next data collection.

The most direct way to measure the stability of the features and feature pairs is to look at their ranking across the geographical, topographical and temporal variations provided by Sites 1, 2, and 3. Looking at the results in Tables 6-I and 6-II, it is clear that the rankings of the results are extremely consistent between the three sites. The stability of the rankings gives an indication that the feature pairs listed in Table 6-II would make good initial estimates for the feature pairs to be used to discriminate between buildings and trees regardless of location or the time of the year that the data are collected.

## CHAPTER 7 CONCLUSION

A novel set of features in ALSM data has been defined and studied to determine the geographic, topographic, and temporal stability of their ability to aid in the discrimination between two important classes of objects, buildings and trees. These two classes are generally difficult to separate automatically using height information alone because they exhibit similar elevations above ground. The separability is measured using a newly created median-based distance measure and a mutual information approach that does not assume that the data have a Gaussian distribution. This is critical because we chose to use only a single class for tree and a single class for building to minimize the a priori information required for training and testing. The goal is to develop a methodology that can apply to a broad spectrum of sites with little to know site-specific training. The tradeoff for this approach is that dissimilar types of vegetation and buildings are lumped together, leading to highly non-Gaussian clusters.

By employing these measures, it becomes possible to systematically quantify and rank the separability afforded by each feature. Thus the methodology will work no matter what data are presented to it. If separability is small in all features or is maximized for different features, the measures will clearly indicate so.

The features are found to be affected by certain geometrically-related parameters, in particular, the scan angle. While the ultimate impact of these effects on classification accuracy is not studied in this work, the overall separation measures remained high in spite of the effects of geometry. Some non-standard features such as standard deviations

of the intensities are found to provide some of the best separation between buildings and trees for both the median-based distance measure and mutual information.

The rankings of the separability appear to be stable over different geographic regions, topographic relief and different times. Although areas with heterogeneous samples (vegetation in this case) degrade the median-based distance measures, the rank of the distance measures is still relatively consistent with areas of homogeneous samples. The mutual information-based measures of separability appear to be more insensitive to heterogeneity of the samples. If heterogeneity of data is an undesirable consequence of the data collection, a mutual information-based measure may perform better than a median-based measure. If, on the other hand, heterogeneity is a property of the data that is to be exploited, it may be better to use a median-based measure. The fact that heterogeneous samples affect the measures is an indication that ALSM is sensitive to different kinds of vegetation and could potentially be used to remotely sense vegetation parameters.

While tree vegetation is found to be seasonally varying, the tree-building separation measures obtained with transferred building data are relatively consistent with those obtained using indigenous building data at other sites. Thus, the feature set described will provide reasonable separation over a wide range of seasonal conditions. Even though the temporal variation in the vegetation is undesirable for surface mapping, it can be highly desirable to quantify seasonal variation for biophysical studies. (Basically you want to see small, but nonzero distances for the seasonally varying case so that you are confident you can separate using these features regardless of season, yet you can also

detect some change. This is exactly what we see. ) This feature set provides a mechanism to quantify that variability using standard ALSM data.

When the features are analyzed using different grid cell sizes, the rankings of the best features and feature pairs are relatively invariant. This indicates that although the mutual information (and therefore separation) may be scale dependent, the features that are best for separation should be the same regardless of the scale being used to analyze the data. It is reasonable that scale dependencies would become more significant for grid cells smaller than 3 m x 3m in this case because that would drive the average number of points per cell lower.

The results of this research have implications on future hardware improvements to ALSM systems. For instance, it is reasonable to expect that the elevation-related features would have performed better had the laser pulses been shorter and the timing of the system been able to utilize these shorter pulses. This would have allowed the ALSM system to detect shorter distances between the surfaces struck by the first and last pulses. This change would have given the system a better chance of detecting the roughness of the tree canopy. However, this change would have to come with the concomitant change in pulse power. This, in turn, would have reduced the range of the intensity measurements. It is conceivable that any benefit gained in improving the elevation-based features could be lost in the intensity-based features.

Another change to the ALSM system that would benefit further research would be the increase in the number of data points on the ground. This can be accomplished either by overlaying additional strips or by increasing the repetition rate of the laser. This change would allow for a greater number density of points which would improve both the

statistical properties of the feature calculations as well as allow for the possibility of reducing the size of the grid cells. Smaller grid cells, in turn, would increase the likelihood of creating ‘pure pixels’ which would improve the separation between classes.

One last hardware change of interest is the measurement of polarization of the return signal. Laser light is highly polarized and measuring the polarization of the return would provide information about the reflecting surface. Surfaces like calm water and metal would most likely produce a consistent polarization of the return while irregular surfaces such as tree canopies would be more likely to produce random polarizations. This additional feature could provide significant insight into the objects being detected.

One extension of the features that would be potentially useful to analyze would be the three dimensional distribution of the data points for different classes. It is reasonable to expect that the distribution of data would be different for classes such as buildings and trees. As a result, the shapes of the distributions could potentially provide useful information for the discrimination between classes.

There are several interesting directions that were not addressed in this research which bear future investigation.

1. A cursory analysis was performed to determine the relation between the features and geometrically-related variables, but a more detailed analysis could provide a better interpretation of the relationship between these variables, particularly the maximum gradient of the ground. For instance, it is reasonable to assume that the direction at which the laser strikes the surface may play as significant a part as the angle. For instance, the possibility of specular reflections would not be accounted for in the current analysis.
2. The correlations between the features and geometrically-related variables were studied but the effect of correcting for these correlations on the ability to discriminate between classes was not. This correction could influence the separability between classes.
3. This direction of research would benefit from a more complete examination of ways in which the separation between the classes can be maximized. For instance, techniques like Fisher linear discriminant and principal component analysis could

potentially be used to transform the feature space to improve the separability between the classes.

4. If ALSM is to be used as a remote sensing device, additional work is necessary to determine the best classification schemes to be used. This work will need to be addressed while trying to discriminate between new classes of objects. Some of the more obvious new classes to consider are roads, grasslands, and water.
5. It appears as if the data density may influence the ability to differentiate between classes. Additional work quantifying this relationship would be useful in determining the how best to use ALSM in remote sensing.
6. It is postulated in this work that part of the reason that differences in first and last stop elevations doesn't work better in identifying vegetation is due to the hardware limited 2.5m distance between stops that is necessary to trigger a true last return. Performing a similar analysis with different hardware that allowed for a smaller distance between true first and last stops would help to determine the true potential elevation differences in discriminating between classes. Using waveform digitization of the return could also help in the discrimination between classes.
7. Initial indications are that ALSM can determine seasonal differences in vegetation. Considering the importance of monitoring vegetation health, the full extent of the ability of ALSM in this area should be determined.
8. Grid cell size apparently plays an important role in the separability between classes indicating that the features are scale dependent. A thorough investigation of this scale dependence could help to determine the optimal size over which the features work best.

## APPENDIX A CORRELATION STUDY

Tables A-I and A-II are the correlations between the geometric variables and the 11 features for Sites 1 and 2 respectively. The correlations for the buildings and trees for Sites 1 and 2 are listed separately. Many of the correlations are statistically significant but few are of practical significance. By plotting the data, it appears that practical significance is not achieved until the absolute correlation is greater than 0.5. Those results are presented in Chapter 5.

Table A-I. Correlations for Site 1 (Flagler County, FL).

### Building Correlations with First Stop Number Density

Mean First Stop Elevation 0.009; Significance = 0.23  
Mean Last Stop Elevation 0.013; Significance = 0.32  
St. Dev. First Stop Elevation 0.106; Significance = 1.00  
St. Dev. Last Stop Elevation 0.080; Significance = 0.99  
Difference in Stop Elevation -0.061; Significance = 0.95  
Difference between Mean First Stop and Ground Elevation -0.032; Significance = 0.69  
Difference between Mean Last Stop and Ground Elevation -0.027; Significance = 0.60  
Difference between Mean First and Low Last Stop Elevation 0.081; Significance = 0.99  
Mean First Stop Intensity 0.026; Significance = 0.59  
Mean Last Stop Intensity 0.024; Significance = 0.55  
St. Dev. First Stop Intensity 0.052; Significance = 0.90  
St. Dev. Last Stop Intensity 0.072; Significance = 0.98  
Difference in Stop Intensities 0.043; Significance = 0.83

### Tree Correlations with First Stop Number Density

Mean First Stop Elevation 0.008; Significance = 0.41  
Mean Last Stop Elevation 0.090; Significance = 1.00  
St. Dev. First Stop Elevation 0.056; Significance = 1.00  
St. Dev. Last Stop Elevation -0.078; Significance = 1.00  
Difference in Stop Elevation -0.139; Significance = 1.00  
Difference between Mean First Stop and Ground Elevation -0.059; Significance = 1.00  
Difference between Mean Last Stop and Ground Elevation 0.047; Significance = 1.00  
Difference between Mean First and Low Last Stop Elevation 0.012; Significance = 0.60  
Mean First Stop Intensity 0.089; Significance = 1.00

## Table A-I—Continued.

Mean Last Stop Intensity 0.144; Significance = 1.00  
 St. Dev. First Stop Intensity 0.051; Significance = 1.00  
 St. Dev. Last Stop Intensity -0.056; Significance = 1.00  
 Difference in Stop Intensities -0.101; Significance = 1.00

## Building Correlations with Last Stop Number Density

Mean First Stop Elevation -0.005; Significance = 0.12  
 Mean Last Stop Elevation -0.004; Significance = 0.09  
 St. Dev. First Stop Elevation 0.115; Significance = 1.00  
 St. Dev. Last Stop Elevation 0.115; Significance = 1.00  
 Difference in Stop Elevation -0.014; Significance = 0.35  
 Difference between Mean First Stop and Ground Elevation -0.048; Significance = 0.88  
 Difference between Mean Last Stop and Ground Elevation -0.047; Significance = 0.87  
 Difference between Mean First and Low Last Stop Elevation 0.126; Significance = 1.00  
 Mean First Stop Intensity 0.006; Significance = 0.15  
 Mean Last Stop Intensity 0.006; Significance = 0.15  
 St. Dev. First Stop Intensity 0.034; Significance = 0.72  
 St. Dev. Last Stop Intensity 0.067; Significance = 0.97  
 Difference in Stop Intensities 0.001; Significance = 0.03

## Tree Correlations with Last Stop Number Density

Mean First Stop Elevation -0.007; Significance = 0.38  
 Mean Last Stop Elevation -0.051; Significance = 1.00  
 St. Dev. First Stop Elevation 0.030; Significance = 0.97  
 St. Dev. Last Stop Elevation 0.140; Significance = 1.00  
 Difference in Stop Elevation 0.075; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.058; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.110; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation 0.229; Significance = 1.00  
 Mean First Stop Intensity 0.131; Significance = 1.00  
 Mean Last Stop Intensity -0.018; Significance = 0.80  
 St. Dev. First Stop Intensity -0.035; Significance = 0.99  
 St. Dev. Last Stop Intensity 0.064; Significance = 1.00  
 Difference in Stop Intensities 0.168; Significance = 1.00

## Building Correlations with Scan Angle

Mean First Stop Elevation -0.859; Significance = 1.00  
 Mean Last Stop Elevation -0.855; Significance = 1.00  
 St. Dev. First Stop Elevation -0.027; Significance = 0.61  
 St. Dev. Last Stop Elevation -0.062; Significance = 0.95  
 Difference in Stop Elevation -0.068; Significance = 0.97  
 Difference between Mean First Stop and Ground Elevation -0.537; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.532; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.071; Significance = 0.98  
 Mean First Stop Intensity -0.460; Significance = 1.00

## Table A-I—Continued.

Mean Last Stop Intensity -0.463; Significance = 1.00  
 St. Dev. First Stop Intensity -0.176; Significance = 1.00  
 St. Dev. Last Stop Intensity -0.187; Significance = 1.00  
 Difference in Stop Intensities 0.013; Significance = 0.31

## Tree Correlations with Scan Angle

Mean First Stop Elevation -0.797; Significance = 1.00  
 Mean Last Stop Elevation -0.750; Significance = 1.00  
 St. Dev. First Stop Elevation 0.187; Significance = 1.00  
 St. Dev. Last Stop Elevation 0.086; Significance = 1.00  
 Difference in Stop Elevation 0.115; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.227; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.300; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation 0.031; Significance = 0.97  
 Mean First Stop Intensity -0.413; Significance = 1.00  
 Mean Last Stop Intensity -0.397; Significance = 1.00  
 St. Dev. First Stop Intensity 0.112; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.063; Significance = 1.00  
 Difference in Stop Intensities 0.096; Significance = 1.00

## Building Correlations with Gradient

Mean First Stop Elevation 0.078; Significance = 0.99  
 Mean Last Stop Elevation 0.086; Significance = 0.99  
 St. Dev. First Stop Elevation -0.269; Significance = 1.00  
 St. Dev. Last Stop Elevation -0.285; Significance = 1.00  
 Difference in Stop Elevation -0.110; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.217; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.208; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.265; Significance = 1.00  
 Mean First Stop Intensity -0.101; Significance = 1.00  
 Mean Last Stop Intensity -0.101; Significance = 1.00  
 St. Dev. First Stop Intensity -0.311; Significance = 1.00  
 St. Dev. Last Stop Intensity -0.306; Significance = 1.00  
 Difference in Stop Intensities -0.020; Significance = 0.48

## Tree Correlations with Gradient

Mean First Stop Elevation -0.155; Significance = 1.00  
 Mean Last Stop Elevation 0.079; Significance = 1.00  
 St. Dev. First Stop Elevation -0.023; Significance = 0.91  
 St. Dev. Last Stop Elevation -0.432; Significance = 1.00  
 Difference in Stop Elevation -0.353; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.498; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.209; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.441; Significance = 1.00  
 Mean First Stop Intensity 0.063; Significance = 1.00

## Table A-I—Continued.

Mean Last Stop Intensity 0.247; Significance = 1.00  
 St. Dev. First Stop Intensity -0.049; Significance = 1.00  
 St. Dev. Last Stop Intensity -0.333; Significance = 1.00  
 Difference in Stop Intensities -0.271; Significance = 1.00

## Number of First Stop Points Correlation

Number of Last Stop Points 0.768; Significance = 1.00  
 Scan Angle -0.005; Significance = 0.30  
 Gradient 0.046; Significance = 1.00

## Number of Last Stop Points Correlation

Scan Angle 0.009; Significance = 0.49  
 Gradient 0.012; Significance = 0.63

## Scan Angle Correlation

Gradient -0.044; Significance = 1.00

## Table A-II. Correlations for Site 2 (Mare Island, CA).

## House Correlations with First Stop Number Density

Mean First Stop Elevation 0.005; Significance = 0.14  
 Mean Last Stop Elevation 0.005; Significance = 0.15  
 St. Dev. First Stop Elevation 0.331; Significance = 1.00  
 St. Dev. Last Stop Elevation 0.332; Significance = 1.00  
 Difference in Stop Elevation -0.011; Significance = 0.32  
 Difference between Mean First Stop and Ground Elevation -0.095; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.094; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation 0.346; Significance = 1.00  
 Mean First Stop Intensity 0.004; Significance = 0.12  
 Mean Last Stop Intensity -0.004; Significance = 0.11  
 St. Dev. First Stop Intensity 0.137; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.160; Significance = 1.00  
 Difference in Stop Intensities 0.166; Significance = 1.00

## Tree Correlations with First Stop Number Density

Mean First Stop Elevation 0.120; Significance = 1.00  
 Mean Last Stop Elevation 0.154; Significance = 1.00  
 St. Dev. First Stop Elevation 0.055; Significance = 1.00  
 St. Dev. Last Stop Elevation 0.028; Significance = 0.95  
 Difference in Stop Elevation -0.104; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.001; Significance = 0.05  
 Difference between Mean Last Stop and Ground Elevation 0.064; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation 0.069; Significance = 1.00  
 Mean First Stop Intensity 0.202; Significance = 1.00  
 Mean Last Stop Intensity 0.215; Significance = 1.00

## Table A-II—Continued.

St. Dev. First Stop Intensity 0.095; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.101; Significance = 1.00  
 Difference in Stop Intensities -0.055; Significance = 1.00

## House Correlations with Last Stop Number Density

Mean First Stop Elevation 0.006; Significance = 0.20  
 Mean Last Stop Elevation 0.006; Significance = 0.20  
 St. Dev. First Stop Elevation 0.328; Significance = 1.00  
 St. Dev. Last Stop Elevation 0.327; Significance = 1.00  
 Difference in Stop Elevation 0.004; Significance = 0.11  
 Difference between Mean First Stop and Ground Elevation -0.100; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.100; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation 0.357; Significance = 1.00  
 Mean First Stop Intensity -0.003; Significance = 0.08  
 Mean Last Stop Intensity -0.010; Significance = 0.30  
 St. Dev. First Stop Intensity 0.134; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.155; Significance = 1.00  
 Difference in Stop Intensities 0.165; Significance = 1.00

## Tree Correlations with Last Stop Number Density

Mean First Stop Elevation 0.111; Significance = 1.00  
 Mean Last Stop Elevation 0.131; Significance = 1.00  
 St. Dev. First Stop Elevation 0.032; Significance = 0.97  
 St. Dev. Last Stop Elevation 0.076; Significance = 1.00  
 Difference in Stop Elevation -0.044; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.005; Significance = 0.26  
 Difference between Mean Last Stop and Ground Elevation 0.021; Significance = 0.85  
 Difference between Mean First and Low Last Stop Elevation 0.127; Significance = 1.00  
 Mean First Stop Intensity 0.248; Significance = 1.00  
 Mean Last Stop Intensity 0.151; Significance = 1.00  
 St. Dev. First Stop Intensity 0.042; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.188; Significance = 1.00  
 Difference in Stop Intensities 0.140; Significance = 1.00

## House Correlations with Scan Angle

Mean First Stop Elevation -0.403; Significance = 1.00  
 Mean Last Stop Elevation -0.405; Significance = 1.00  
 St. Dev. First Stop Elevation 0.062; Significance = 0.98  
 St. Dev. Last Stop Elevation 0.067; Significance = 0.99  
 Difference in Stop Elevation 0.044; Significance = 0.91  
 Difference between Mean First Stop and Ground Elevation -0.015; Significance = 0.44  
 Difference between Mean Last Stop and Ground Elevation -0.016; Significance = 0.45  
 Difference between Mean First and Low Last Stop Elevation 0.047; Significance = 0.93  
 Mean First Stop Intensity 0.386; Significance = 1.00  
 Mean Last Stop Intensity 0.384; Significance = 1.00

## Table A-II—Continued.

St. Dev. First Stop Intensity 0.166; Significance = 1.00  
 St. Dev. Last Stop Intensity 0.180; Significance = 1.00  
 Difference in Stop Intensities 0.047; Significance = 0.93

## Tree Correlations with Scan Angle

Mean First Stop Elevation -0.560; Significance = 1.00  
 Mean Last Stop Elevation -0.482; Significance = 1.00  
 St. Dev. First Stop Elevation -0.259; Significance = 1.00  
 St. Dev. Last Stop Elevation -0.669; Significance = 1.00  
 Difference in Stop Elevation -0.678; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.768; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.600; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.701; Significance = 1.00  
 Mean First Stop Intensity 0.687; Significance = 1.00  
 Mean Last Stop Intensity 0.670; Significance = 1.00  
 St. Dev. First Stop Intensity -0.011; Significance = 0.53  
 St. Dev. Last Stop Intensity 0.069; Significance = 1.00  
 Difference in Stop Intensities -0.071; Significance = 1.00

## House Correlations with Gradient

Mean First Stop Elevation 0.029; Significance = 0.73  
 Mean Last Stop Elevation 0.027; Significance = 0.70  
 St. Dev. First Stop Elevation -0.483; Significance = 1.00  
 St. Dev. Last Stop Elevation -0.486; Significance = 1.00  
 Difference in Stop Elevation 0.081; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation 0.219; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation 0.217; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.473; Significance = 1.00  
 Mean First Stop Intensity -0.022; Significance = 0.60  
 Mean Last Stop Intensity -0.014; Significance = 0.41  
 St. Dev. First Stop Intensity -0.252; Significance = 1.00  
 St. Dev. Last Stop Intensity -0.274; Significance = 1.00  
 Difference in Stop Intensities -0.171; Significance = 1.00

## Tree Correlations with Gradient

Mean First Stop Elevation -0.275; Significance = 1.00  
 Mean Last Stop Elevation -0.242; Significance = 1.00  
 St. Dev. First Stop Elevation -0.189; Significance = 1.00  
 St. Dev. Last Stop Elevation -0.289; Significance = 1.00  
 Difference in Stop Elevation -0.305; Significance = 1.00  
 Difference between Mean First Stop and Ground Elevation -0.178; Significance = 1.00  
 Difference between Mean Last Stop and Ground Elevation -0.047; Significance = 1.00  
 Difference between Mean First and Low Last Stop Elevation -0.300; Significance = 1.00  
 Mean First Stop Intensity 0.295; Significance = 1.00  
 Mean Last Stop Intensity 0.312; Significance = 1.00

## Table A-II—Continued.

St. Dev. First Stop Intensity -0.114; Significance = 1.00  
St. Dev. Last Stop Intensity -0.136; Significance = 1.00  
Difference in Stop Intensities -0.077; Significance = 1.00

## Number of First Stop Points Correlation

Number of Last Stop Points 0.902; Significance = 1.00  
Scan Angle 0.154; Significance = 1.00  
Gradient -0.091; Significance = 1.00

## Number of Last Stop Points Correlation

Scan Angle 0.147; Significance = 1.00  
Gradient -0.099; Significance = 1.00

## Scan Angle Correlation

Gradient 0.277; Significance = 1.00

APPENDIX B  
COMPLETE LIST OF MEDIAN DISTANCES

Table B-I below is a complete list of the median distance measures for each of the 55 feature pairs analyzed.

Table B-I. A complete list of the median distance measures for all 55 feature pairs.

Feature	Site 1		Site 2		Site 3	
	distance	rank	distance	rank	distance	rank
StD First Stop Elevation vs. StD Last Stop Elevation	0.924	50	1.538	39	0.174	47
StD First Stop Elevation vs. Difference in Elevation	1.176	47	1.841	32	0.152	48
StD First Stop Elevation vs. Difference between First Stop and ground	0.551	53	2.703	16	0.19	46
StD First Stop Elevation vs. Difference between Last Stop and ground	0.295	54	2.476	19	0.243	41
StD First Stop Elevation vs. Difference between First Stop and lowz2	1.911	38	1.59	36	0.311	38
StD First Stop Elevation vs. Mean First Stop Intensity	3.113	18	1.637	34	1.221	5
StD First Stop Elevation vs. Mean Last Stop Intensity	2.818	25	1.152	50	0.932	18
StD First Stop Elevation vs. StD First Stop Intensity	2.029	36	2.709	15	0.894	22
StD First Stop Elevation vs. Std Last Stop Intensity	2.651	29	2.797	12	0.951	16
StD First Stop Elevation vs. Difference in Intensity	1.145	48	1.43	43	0.335	27
StD Last Stop Elevation vs. Difference in Elevation	1.425	43	1.628	35	0.102	53
StD Last Stop Elevation vs. Difference between First Stop and ground	1.383	44	2.281	21	0.131	51

Table B-I—Continued.

Feature	Site 1		Site 2		Site 3	
	distance	rank	distance	rank	distance	rank
StD Last Stop Elevation vs. Difference between Last Stop and ground	1.033	49	2.259	22	0.197	45
StD Last Stop Elevation vs. Difference between First Stop and lowz2	1.433	42	1.203	49	0.15	49
StD Last Stop Elevation vs. Mean First Stop Intensity	4.435	6	1.651	33	1.184	7
StD Last Stop Elevation vs. Mean Last Stop Intensity	3.158	17	1.052	51	0.888	25
StD Last Stop Elevation vs. StD First Stop Intensity	3.321	15	3.182	8	0.862	27
StD Last Stop Elevation vs. StD Last Stop Intensity	2.496	31	2.699	17	0.915	21
StD Last Stop Elevation vs. Difference in Intensity	1.751	39	1.459	41	0.287	39
Difference in Elevation vs. Difference between First Stop and ground	1.199	46	2.118	26	0.091	54
Difference in Elevation vs. Difference between Last Stop and ground	0.826	52	2.198	23	0.125	52
Difference in Elevation vs. Difference between First Stop and lowz2	1.912	37	1.558	37	0.144	50
Difference in Elevation vs. Mean First Stop Intensity	4.304	8	1.431	42	1.122	10
Difference in Elevation vs. Mean Last Stop Intensity	2.999	20	0.839	53	0.828	32
Difference in Elevation vs. StD First Stop Intensity	3.569	13	3.432	5	0.799	33
Difference in Elevation vs. Std Last Stop Intensity	2.887	24	3.123	9	0.856	28
Difference in Elevation vs. Difference in Intensity	1.567	40	1.338	46	0.225	44
Difference between First Stop and ground vs. Difference between Last Stop and ground	0.104	55	1.539	38	0.053	55
Difference between First Stop and ground vs. Difference between First Stop and lowz2	2.452	33	2.009	28	0.235	42

Table B-I—Continued.

Feature	Site 1		Site 2		Site 3	
	distance	rank	distance	rank	distance	rank
Difference between First Stop and ground vs. Mean First Stop Intensity	2.687	27	1.982	29	1.123	9
Difference between First Stop and ground vs. Mean Last Stop Intensity	2.494	32	1.391	44	0.831	31
Difference between First Stop and ground vs. StD First Stop Intensity	2.986	21	4.224	1	0.793	34
Difference between First Stop and ground vs. Std Last Stop Intensity	3.042	19	3.92	3	0.851	29
Difference between First Stop and ground vs. Difference in Intensity	1.275	45	2.122	25	0.227	43
Difference between Last Stop and ground vs. Difference between First Stop and lowz2	2.243	34	2.022	27	0.343	36
Difference between Last Stop and ground vs. Mean First Stop Intensity	2.161	35	1.909	31	1.167	8
Difference between Last Stop and ground vs. Mean Last Stop Intensity	1.488	41	1.368	45	0.876	26
Difference between Last Stop and ground vs. StD First Stop Intensity	2.616	30	3.941	2	0.834	30
Difference between Last Stop and ground vs. Std Last Stop Intensity	2.658	28	3.702	4	0.891	24
Difference between Last Stop and ground vs. Difference in Intensity	0.858	51	1.949	30	0.272	40
Difference between First Stop and lowz2 vs. Mean First Stop Intensity	5.943	1	1.519	40	1.328	2
Difference between First Stop and lowz2 vs. Mean Last Stop Intensity	4.678	4	0.918	52	1.02	13
Difference between First Stop and lowz2 vs. StD First Stop Intensity	4.819	3	3.262	7	1.027	12
Difference between First Stop and lowz2 vs. Std Last Stop Intensity	4.016	11	2.76	13	1.083	11
Difference between First Stop and lowz2 vs. Difference in Intensity	2.697	26	1.322	47	0.424	35
Mean First Stop Intensity vs. Mean Last Stop Intensity	3.543	14	0.322	55	1.199	6
Mean First Stop Intensity vs. StD First Stop Intensity	4.391	7	2.926	11	1.229	4
Mean First Stop Intensity vs. Std Last Stop Intensity	5.485	2	3.072	10	1.333	1
Mean First Stop Intensity vs. Difference in Intensity	4.194	10	1.296	48	1.246	3
Mean Last Stop Intensity vs. StD First Stop Intensity	4.54	5	2.558	18	0.997	14

Table B-I—Continued.

Feature	Site 1		Site 2		Site 3	
	distance	rank	distance	rank	distance	rank
Mean Last Stop Intensity vs. Std Last Stop Intensity	4.229	9	2.283	20	0.967	15
Mean Last Stop Intensity vs. Difference in Intensity	2.977	22	0.712	54	0.894	23
StD First Stop Intensity vs. Std Last Stop Intensity	3.919	12	3.316	6	0.941	17
StD First Stop Intensity vs. Difference in Intensity	3.256	16	2.72	14	0.922	20
StD Last Stop Intensity vs. Difference in Intensity	2.966	23	2.156	24	0.931	19

In general, the results indicate that the intensity-based features have larger distances than the elevation-based features. This is an indication that the intensity-based measures would be better candidates for use in the classification of buildings and trees.

It is interesting to note that the distances between Sites 1 and 3 agree better than the distances between Sites 1 and 2. This appears to be due to the heterogeneous nature of the training data in Site 2 as was discussed in Chapter 5.

## APPENDIX C RULE-BASED CLASSIFICATION SYSTEM

To give a better idea of how distance between classes would translate into separability, a simple rule-based classification system was designed. This system was not intended for robust use on arbitrary data sets but was instead designed to test the separability of the buildings and tree classes in pre-selected feature space. The results of the classifications performed using this system should not be seen as being optimal.

The first step in the rule-based system was to eliminate all pixels which were located sufficiently close to the ground. In all cases, a cutoff of two meters was used to eliminate pixels near the ground. Once all aboveground pixels had been located, rules derived from the clustering of data in feature space were applied to discriminate between buildings and trees. The rule-based logic for classification done in Site 1 are presented in Tables C-I – C-IV while the logic for Site 2 are presented in Table C-V – C-VIII. All of these rules already assume that ground points have been eliminated.

Table C-I. Rule-based logic for Site 1 using mean first stop intensity and standard deviation of the last stop intensity for Site 1.

```
if (mean first stop intensity > 100 and mean first stop intensity < 200 and
    standard deviation of the last stop intensity > 8 and
    standard deviation of the last stop intensity < 103) then the pixel is a tree
else
    the pixel is a building
```

Table C-II. Rule-based logic for Site 1 using the difference between mean first stop and low last stop elevation and mean first stop intensity for Site 1.

```
if (difference between mean first stop and low last stop elevation > 0 and
    difference between mean first stop and low last stop elevation < 8 and
    mean first stop intensity > 100 and mean first stop intensity < 200)
```

then the pixel is a tree  
 else  
 the pixel is a building

Table C-III. Rule-based logic for Site 1 using the difference between mean stop elevations and standard deviation of first stop intensity for Site 1.

if (difference between mean stop elevations > 0 and  
 standard deviation of first stop intensity > 10 and  
 standard deviation of first stop intensity < 70) then the pixel is a tree  
 else  
 the pixel is a building

Table C-IV. Rule-based logic for Site 1 using the standard deviation of first stop elevation and difference between mean stop elevations for Site 1.

if (standard deviation of first stop elevation > 0.1 and  
 standard deviation of first stop elevation < 1 and  
 difference between mean stop elevations > 0) then the pixel is a tree  
 else  
 the pixel is a building

Table C-V. Rule-based logic for Site 1 using mean first stop intensity and standard deviation of the last stop intensity for Site 2.

if (mean first stop intensity > 100 and mean first stop intensity < 150 and  
 standard deviation of the last stop intensity > 0 and  
 standard deviation of the last stop intensity < 30) then the pixel is a building  
 else  
 the pixel is a tree

Table C-VI. Rule-based logic for Site 1 using the difference between mean first stop and low last stop elevation and mean first stop intensity for Site 2.

if (difference between mean first stop and low last stop elevation > 0 and  
 difference between mean first stop and low last stop elevation < 2 and  
 mean first stop intensity > 100 and mean first stop intensity < 150)  
 then the pixel is a building  
 else  
 the pixel is a tree

Table C-VII. Rule-based logic for Site 1 using the difference between mean stop elevations and standard deviation of first stop intensity for Site 2.

if (difference between mean stop elevations > 0.5 and  
 standard deviation of first stop intensity > 0 and

standard deviation of first stop intensity  $< 25$ ) then the pixel is a building  
else  
the pixel is a tree

Table C-VIII. Rule-based logic for Site 1 using the standard deviation of first stop elevation and difference between mean stop elevations for Site 2.  
if (standard deviation of first stop elevation  $> 0$  and  
standard deviation of first stop elevation  $< 3.2$  and  
difference between mean stop elevations  $> 0.5$ ) then the pixel is a building  
else  
the pixel is a tree

APPENDIX D  
ALGORITHM USED FOR COMPUTING MUTUAL INFORMATION AND PARZEN  
WINDOWS

Table D-I below is a Matlab program listing of the algorithm used to compute the mutual information for two dimensional features.

Table D-I. A program listing of the algorithm for computing the two dimensional mutual information.

```
function [ICX] = MIG2_alt2(x,y,dx,dy,totx,toty,housex,housey,treex,trey)

clear p_totx p_house p_tree;

lentot = length(housex) + length(treex);
p_c(1) = length(housex)/lentot;
p_c(2) = length(treex)/lentot;

hx = dx * 10;
hy = dy * 10;
V = hx * hy;

p_tot = parzen2_alt(x,y,totx,toty,hx,hy,V);
p_house = parzen2_alt(x,y,housex,housey,hx,hy,V);
p_tree = parzen2_alt(x,y,treex,trey,hx,hy,V);

% Try to compute MIG

% First need to compute Entropy of the class

HC = -(p_c(1)*log(p_c(1))/log(2) + p_c(2)*log(p_c(2))/log(2));

% Second need to compute the conditional entropy

HCX1 = 0;
HCX2 = 0;
for i = 1:length(x)
    for j = 1:length(y)
        if (p_tot(i,j) ~= 0)
            pcx = p_house(i,j)*p_c(1)/p_tot(i,j);
            if (pcx ~=0)
```

Table D-I—Continued.

```

        HCX1 = HCX1 + p_house(i,j)*log(pcx)/log(2)*dx*dy;
    end
end
end
end
HCX1 = p_c(1) * (-HCX1);
for i = 1:length(x)
    for j = 1:length(y)
        if (p_tot(i,j) ~= 0)
            pcx = p_tree(i,j)*p_c(2)/p_tot(i,j);
            if (pcx ~= 0)
                HCX2 = HCX2 + p_tree(i,j)*log(pcx)/log(2)*dx*dy;
            end
        end
    end
end
end
HCX2 = p_c(2) * (-HCX2);

HCX = HCX1 + HCX2;

% Finally compute the mutual information

ICX = HC - HCX;

```

Table D-II below is a Matlab program listing of the algorithm used to compute the Parzen windows estimation of the two-dimensional probability distribution function.

This program is called from the routine presented in Table D-I.

Table D-II. A program listing of the algorithm for computing the Parzen windows estimation of a two-dimensional probability distribution function.

```

function p = parzen2(x,y,classx,classy,hx,hy,V)

clear p;
p = zeros(length(x),length(y));

for i = 1:length(x)
    for j = 1:length(y)
        for k = 1:length(classx)
            distx = (x(i)-classx(k))/hx;
            disty = (y(j)-classy(k))/hy;

```

Table D-II—Continued.

```
        if (abs(distx) < 0.5 && abs(disty) < 0.5)
            p(i,j) = p(i,j) + 1/V;
        end
    end
    p(i,j) = p(i,j) / length(classx);
end
end
```

APPENDIX E  
COMPLETE LIST OF MUTUAL INFORMATION MEASURES

Table E-I below is a complete list of the mutual information for each of the 55 feature pairs analyzed.

Table E-I. A complete list of the mutual information for all 55 feature pairs. The mutual information is given in bits.

Feature	Site 1		Site 2		Site 3	
	MIG	rank	MIG	rank	MIG	rank
StD First Stop Elevation vs. StD Last Stop Elevation	0.18	53	0.394	36	0.061	41
StD First Stop Elevation vs. Difference in Elevation	0.171	55	0.329	48	0.045	43
StD First Stop Elevation vs. Difference between First Stop and ground	0.244	47	0.363	43	0.035	48
StD First Stop Elevation vs. Difference between Last Stop and ground	0.202	51	0.309	51	0.042	46
StD First Stop Elevation vs. Difference between First Stop and lowz2	0.211	50	0.362	44	0.04	47
StD First Stop Elevation vs. Mean First Stop Intensity	0.603	11	0.455	27	0.238	9
StD First Stop Elevation vs. Mean Last Stop Intensity	0.591	19	0.423	31	0.194	22
StD First Stop Elevation vs. StD First Stop Intensity	0.437	36	0.53	19	0.162	30
StD First Stop Elevation vs. Std Last Stop Intensity	0.39	38	0.552	9	0.203	19
StD First Stop Elevation vs. Difference in Intensity	0.236	49	0.411	34	0.123	37
StD Last Stop Elevation vs. Difference in Elevation	0.172	54	0.344	46	0.046	42
StD Last Stop Elevation vs. Difference between First Stop and ground	0.289	41	0.393	37	0.035	48

Table E-I—Continued.

Feature	Site 1		Site 2		Site 3	
	MIG	rank	MIG	rank	MIG	rank
StD Last Stop Elevation vs. Difference between Last Stop and ground	0.278	44	0.386	38	0.044	44
StD Last Stop Elevation vs. Difference between First Stop and lowz2	0.186	52	0.354	45	0.043	45
StD Last Stop Elevation vs. Mean First Stop Intensity	0.596	16	0.541	14	0.237	11
StD Last Stop Elevation vs. Mean Last Stop Intensity	0.599	12	0.548	12	0.197	20
StD Last Stop Elevation vs. StD First Stop Intensity	0.474	30	0.552	9	0.156	31
StD Last Stop Elevation vs. StD Last Stop Intensity	0.454	35	0.558	7	0.209	18
StD Last Stop Elevation vs. Difference in Intensity	0.239	48	0.423	31	0.128	36
Difference in Elevation vs. Difference between First Stop and ground	0.466	32	0.28	52	0.023	53
Difference in Elevation vs. Difference between Last Stop and ground	0.465	33	0.276	53	0.028	51
Difference in Elevation vs. Difference between First Stop and lowz2	0.436	37	0.27	54	0.033	50
Difference in Elevation vs. Mean First Stop Intensity	0.613	8	0.436	29	0.238	9
Difference in Elevation vs. Mean Last Stop Intensity	0.615	5	0.423	31	0.196	21
Difference in Elevation vs. StD First Stop Intensity	0.562	20	0.531	17	0.142	33
Difference in Elevation vs. Std Last Stop Intensity	0.554	21	0.519	21	0.188	24
Difference in Elevation vs. Difference in Intensity	0.455	34	0.335	47	0.106	38
Difference between First Stop and ground vs. Difference between Last Stop and ground	0.255	46	0.262	55	0.019	54
Difference between First Stop and ground vs. Difference between First Stop and lowz2	0.285	43	0.323	49	0.018	55

Table E-I—Continued.

Feature	Site 1		Site 2		Site 3	
	MIG	rank	MIG	rank	MIG	rank
Difference between First Stop and ground vs. Mean First Stop Intensity	0.612	9	0.444	28	0.217	14
Difference between First Stop and ground vs. Mean Last Stop Intensity	0.615	5	0.426	30	0.172	28
Difference between First Stop and ground vs. StD First Stop Intensity	0.518	22	0.548	12	0.136	34
Difference between First Stop and ground vs. Std Last Stop Intensity	0.511	24	0.55	11	0.18	26
Difference between First Stop and ground vs. Difference in Intensity	0.323	39	0.379	39	0.1	40
Difference between Last Stop and ground vs. Difference between First Stop and lowz2	0.287	42	0.319	50	0.028	51
Difference between Last Stop and ground vs. Mean First Stop Intensity	0.616	4	0.402	35	0.22	13
Difference between Last Stop and ground vs. Mean Last Stop Intensity	0.608	10	0.379	39	0.176	27
Difference between Last Stop and ground vs. StD First Stop Intensity	0.508	25	0.531	17	0.143	32
Difference between Last Stop and ground vs. Std Last Stop Intensity	0.512	23	0.541	14	0.185	25
Difference between Last Stop and ground vs. Difference in Intensity	0.317	40	0.376	41	0.211	17
Difference between First Stop and lowz2 vs. Mean First Stop Intensity	0.597	13	0.488	25	0.212	16
Difference between First Stop and lowz2 vs. Mean Last Stop Intensity	0.597	13	0.49	24	0.169	29
Difference between First Stop and lowz2 vs. StD First Stop Intensity	0.48	28	0.535	16	0.135	35
Difference between First Stop and lowz2 vs. Std Last Stop Intensity	0.468	31	0.525	20	0.193	23
Difference between First Stop and lowz2 vs. Difference in Intensity	0.265	45	0.364	42	0.104	39
Mean First Stop Intensity vs. Mean Last Stop Intensity	0.597	13	0.476	26	0.272	4
Mean First Stop Intensity vs. StD First Stop Intensity	0.62	3	0.668	3	0.292	3
Mean First Stop Intensity vs. Std Last Stop Intensity	0.622	2	0.679	2	0.295	2
Mean First Stop Intensity vs. Difference in Intensity	0.593	17	0.509	23	0.272	4

Table E-I—Continued.

Feature	Site 1		Site 2		Site 3	
	MIG	rank	MIG	rank	MIG	rank
Mean Last Stop Intensity vs. StD First Stop Intensity	0.626	1	0.652	4	0.253	7
Mean Last Stop Intensity vs. Std Last Stop Intensity	0.615	5	0.684	1	0.3	1
Mean Last Stop Intensity vs. Difference in Intensity	0.592	18	0.51	22	0.271	6
StD First Stop Intensity vs. Std Last Stop Intensity	0.489	27	0.555	8	0.227	12
StD First Stop Intensity vs. Difference in Intensity	0.496	26	0.57	5	0.217	14
StD Last Stop Intensity vs. Difference in Intensity	0.479	29	0.566	6	0.245	8

The results indicate that the intensity-based features have larger mutual information than the elevation-based features. This is an indication that the intensity-based measures would be better candidates for use in the classification of buildings and trees.

The consistency between the 3 sites appears to be better than for the median-based distances. This indicates that mutual information is less sensitive to the heterogeneity of the data found in Site 2.

APPENDIX F  
GRID CELL SIZE ANALYSIS

Table F-I below is a complete list of the mutual information for each of the 11 features and 55 feature pairs analyzed using 3 different grid sizes.

Table F-I. A complete list of the mutual information for all 11 features and 55 feature pairs. The mutual information is given in bits.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
MIG for StD First Stop Elevation	0.03	0.10	0.43	0.18	0.18	0.15	0.02	0.01	0.02
MIG for Std Last Stop Elevation	0.12	0.15	0.13	0.32	0.37	0.31	0.02	0.01	0.01
MIG for Difference in Elevation	0.15	0.21	0.28	0.19	0.23	0.25	0.01	0.01	0.02
MIG for Difference between First Stop and Ground	0.19	0.18	0.28	0.22	0.20	0.36	0.01	0.00	0.00
MIG for Difference between Last Stop and Ground	0.15	0.16	0.23	0.16	0.13	0.31	0.01	0.01	0.00
MIG for Difference between First stop and lowz2	0.16	0.19	0.28	0.25	0.33	0.28	0.01	0.01	0.01
MIG for Mean First Stop Intensity	0.55	0.58	0.60	0.29	0.30	0.35	0.19	0.29	0.57
MIG for Mean Last Stop Intensity	0.54	0.56	0.59	0.26	0.21	0.22	0.14	0.18	0.40
MIG for StD First Stop Intensity	0.40	0.43	0.49	0.47	0.63	0.58	0.11	0.18	0.44
MIG for StD Last Stop Intensity	0.35	0.34	0.41	0.50	0.63	0.67	0.16	0.28	0.58
MIG for Difference of Intensities	0.21	0.23	0.32	0.31	0.31	0.31	0.08	0.09	0.14
StD First Stop Elevation vs. StD Last Stop Elevation	0.18	0.34	0.56	0.39	0.43	0.40	0.06	0.04	0.09

Table F-I—Continued.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
StD First Stop Elevation vs. Difference in Elevation	0.17	0.30	0.54	0.33	0.36	0.35	0.05	0.04	0.07
StD First Stop Elevation vs. Difference between First Stop and Ground	0.24	0.29	0.58	0.36	0.40	0.58	0.04	0.03	0.07
StD First Stop Elevation vs. Difference between Last Stop and Ground	0.20	0.27	0.53	0.31	0.32	0.55	0.04	0.03	0.07
StD First Stop Elevation vs. Difference between First Stop and lowz2	0.21	0.36	0.59	0.36	0.42	0.40	0.04	0.03	0.06
StD First Stop Elevation vs. Mean First Stop Intensity	0.60	0.64	0.64	0.46	0.49	0.53	0.24	0.35	0.64
StD First Stop Elevation vs. Mean Last Stop Intensity	0.59	0.64	0.64	0.42	0.40	0.42	0.19	0.24	0.49
StD First Stop Elevation vs. StD First Stop Intensity	0.44	0.52	0.62	0.53	0.65	0.64	0.16	0.23	0.50
StD First Stop Elevation vs. Std Last Stop Intensity	0.39	0.47	0.63	0.55	0.67	0.71	0.20	0.31	0.61
StD First Stop Elevation vs. Difference in Intensity	0.24	0.31	0.54	0.41	0.44	0.44	0.12	0.13	0.21
StD Last Stop Elevation vs. Difference in Elevation	0.17	0.31	0.50	0.34	0.39	0.37	0.05	0.04	0.07
StD Last Stop Elevation vs. Difference between First Stop and Ground	0.29	0.32	0.43	0.39	0.48	0.59	0.04	0.03	0.06
StD Last Stop Elevation vs. Difference between Last Stop and Ground	0.28	0.33	0.45	0.39	0.48	0.58	0.04	0.04	0.06

Table F-I—Continued.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
StD Last Stop Elevation vs. Difference between First Stop and lowz2	0.19	0.26	0.47	0.35	0.40	0.40	0.04	0.04	0.05
StD Last Stop Elevation vs. Mean First Stop Intensity	0.60	0.62	0.61	0.54	0.57	0.56	0.24	0.35	0.64
StD Last Stop Elevation vs. Mean Last Stop Intensity	0.60	0.61	0.60	0.55	0.55	0.54	0.20	0.25	0.51
StD Last Stop Elevation vs. StD First Stop Intensity	0.47	0.50	0.54	0.55	0.65	0.63	0.16	0.23	0.48
StD Last Stop Elevation vs. Std Last Stop Intensity	0.45	0.47	0.53	0.56	0.67	0.72	0.21	0.32	0.61
StD Last Stop Elevation vs. Difference in Intensity	0.24	0.29	0.44	0.42	0.46	0.43	0.13	0.14	0.20
Difference in Elevation vs. Difference between First Stop and Ground	0.47	0.45	0.47	0.28	0.30	0.43	0.02	0.03	0.04
Difference in Elevation vs. Difference between Last Stop and Ground	0.47	0.45	0.47	0.28	0.29	0.41	0.03	0.03	0.04
Difference in Elevation vs. Difference between First Stop and lowz2	0.44	0.44	0.47	0.27	0.35	0.34	0.03	0.04	0.04
Difference in Elevation vs. Mean First Stop Intensity	0.61	0.62	0.63	0.44	0.44	0.47	0.24	0.36	0.63
Difference in Elevation vs. Mean Last Stop Intensity	0.62	0.61	0.63	0.42	0.41	0.43	0.20	0.26	0.51
Difference in Elevation vs. StD First Stop Intensity	0.56	0.57	0.60	0.53	0.64	0.62	0.14	0.21	0.47

Table F-I—Continued.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
Difference in Elevation vs. Std Last Stop Intensity	0.55	0.56	0.59	0.52	0.64	0.68	0.19	0.32	0.59
Difference in Elevation vs. Difference in Intensity	0.46	0.45	0.47	0.34	0.37	0.38	0.11	0.12	0.16
Difference between First Stop and Ground vs. Difference between Last Stop and Ground	0.26	0.27	0.38	0.26	0.24	0.39	0.02	0.01	0.01
Difference between First Stop and Ground vs. Difference between First Stop and lowz2	0.29	0.29	0.36	0.32	0.43	0.56	0.02	0.02	0.02
Difference between First Stop and Ground vs. Mean First Stop Intensity	0.61	0.62	0.64	0.44	0.41	0.56	0.22	0.33	0.63
Difference between First Stop and ground vs. Mean Last Stop Intensity	0.62	0.60	0.64	0.43	0.37	0.53	0.17	0.23	0.49
Difference between First Stop and Ground vs. StD First Stop Intensity	0.52	0.54	0.57	0.55	0.66	0.73	0.14	0.21	0.48
Difference between First Stop and Ground vs. Std Last Stop Intensity	0.51	0.48	0.56	0.55	0.66	0.75	0.18	0.30	0.61
Difference between First Stop and Ground vs. Difference in Intensity	0.32	0.34	0.45	0.38	0.40	0.45	0.10	0.12	0.18
Difference between Last Stop and Ground vs. Difference between First Stop and lowz2	0.29	0.32	0.41	0.32	0.42	0.52	0.03	0.03	0.02
Difference between Last Stop and Ground vs. Mean First Stop Intensity	0.62	0.63	0.64	0.40	0.40	0.53	0.22	0.33	0.62

Table F-I—Continued.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
Difference between Last Stop and Ground vs. Mean Last Stop Intensity	0.61	0.62	0.64	0.38	0.30	0.49	0.18	0.22	0.48
Difference between Last Stop and Ground vs. StD First Stop Intensity	0.51	0.53	0.60	0.53	0.66	0.73	0.14	0.21	0.47
Difference between Last Stop and Ground vs. Std Last Stop Intensity	0.51	0.50	0.54	0.54	0.66	0.75	0.19	0.30	0.62
Difference between Last Stop and Ground vs. Difference in Intensity	0.32	0.35	0.43	0.38	0.40	0.43	0.21	0.26	0.44
Difference between First Stop and lowz2 vs. Mean First Stop Intensity	0.60	0.62	0.64	0.49	0.53	0.51	0.21	0.32	0.61
Difference between First Stop and lowz2 vs. Mean Last Stop Intensity	0.60	0.60	0.64	0.49	0.51	0.50	0.17	0.22	0.47
Difference between First Stop and lowz2 vs. StD First Stop Intensity	0.48	0.53	0.59	0.54	0.65	0.66	0.14	0.21	0.46
Difference between First Stop and lowz2 vs. Std Last Stop Intensity	0.47	0.48	0.57	0.53	0.66	0.70	0.19	0.32	0.59
Difference between First Stop and lowz2 vs. Difference in Intensity	0.27	0.32	0.46	0.36	0.43	0.39	0.10	0.13	0.17
Mean First Stop Intensity vs. Mean Last Stop Intensity	0.60	0.61	0.63	0.48	0.46	0.43	0.27	0.35	0.60
Mean First Stop Intensity vs. StD First Stop Intensity	0.62	0.64	0.64	0.67	0.74	0.73	0.29	0.38	0.66

Table F-I—Continued.

Feature	Site 1			Site 2			Site 3		
	3m	5m	10m	3m	5m	10m	3m	5m	10m
Mean First Stop Intensity vs. Std Last Stop Intensity	0.62	0.64	0.64	0.68	0.73	0.74	0.30	0.41	0.70
Mean First Stop Intensity vs. Difference in Intensity	0.59	0.61	0.63	0.51	0.50	0.46	0.27	0.36	0.61
Mean Last Stop Intensity vs. StD First Stop Intensity	0.63	0.63	0.64	0.65	0.72	0.72	0.25	0.32	0.58
Mean Last Stop Intensity vs. Std Last Stop Intensity	0.62	0.63	0.64	0.68	0.74	0.73	0.30	0.39	0.66
Mean Last Stop Intensity vs. Difference in Intensity	0.59	0.61	0.63	0.51	0.50	0.44	0.27	0.35	0.59
StD First Stop Intensity vs. Std Last Stop Intensity	0.49	0.52	0.59	0.56	0.66	0.70	0.23	0.34	0.67
StD First Stop Intensity vs. Difference in Intensity	0.50	0.51	0.57	0.57	0.66	0.67	0.22	0.27	0.53
StD Last Stop Intensity vs. Difference in Intensity	0.48	0.48	0.53	0.57	0.65	0.69	0.25	0.35	0.62

The results indicate that the mutual information increases as grid cell size increases.

This is an indication that larger grid cell sizes provide the ability to better represent, in a statistical sense, the features. In general, the rankings are relatively consistent across the different grid cell sizes.

## LIST OF REFERENCES

- Axelsson, P., 1999, "Processing of laser scanner data — algorithms and applications," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54(2-3)**, pp. 138–147.
- Baltsavias, E. P., 1999a, "Airborne laser scanning: existing systems and firms and other resources," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54(2-3)**, pp. 164–198.
- Baltsavias, E. P., 1999b, "Airborne laser scanning: basic relations and formulas," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54(2-3)**, pp. 199–214.
- Baltsavias, E., Mason, S., and Stallmann, D., 1995, "Use of DTMs/DSMs and orthoimages to support building extraction," in *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Gruen, A., Kübler, O., and Agouris, P. (eds.), Birkhäuser, Basel, pp. 199 – 210.
- Bezdek, J. C. and Pal, N. R., 1998, "Some new indexes of cluster validity," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, **Vol. 28(3)**, pp. 301–315.
- Brenner, C., 2000, "Towards fully automated generation of city models," *Proceedings of the ISPRS*, **Vol. 33**, Amsterdam.
- Brown, R. G. and Hwang, P. Y. C., 1996, *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises and Solution* (3<sup>rd</sup> Edition), Wiley Text Books, New York, 484 pp.
- Brunn, A. and Weidner, U., 1998, "Hierarchical Bayesian nets for building extraction using dense digital surface models," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 53**, pp. 296–307.
- Cover, T. M. and Thomas, J. A., 1991, *Elements of Information Theory*, Wiley, New York, 542 pp.
- Davies, D. L. and Bouldin, D. W., 1979, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **Vol. 1(2)**, pp. 224–227.
- Devore, J. L., 1982, *Probability & Statistics for Engineering and the Sciences*, Brooks/Cole, Monterey, CA.

- Duda, R. O., Hart, P. E. and Stork, D. G., 2001, *Pattern Classification* (2nd Edition), Wiley Interscience, New York, 654 pp.
- Dunn, J. C., 1973, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, **Vol. 3(3)**, pp. 32–57.
- Dunn, J. C., 1974, "Well separated clusters and optimal fuzzy partitions," *J. Cybern.*, **Vol. 4**, pp. 95–104.
- Eckstein, W. and Munkelt, O., 1995, "Extracting objects from digital terrain models," in *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, Schenk, T., editor, Proc. Spie 2572, pages 222–231.
- Elberink, S. O. and Maas, H.-G., 2000, "The use of anisotropic height texture measures for the segmentation of airborne laser scanner data," *Proceedings of the ISPRS*, **Vol. 33**, Amsterdam.
- Elaksher, A. F. and Bethel, J. S., 2001, "Reconstructing 3D buildings from LIDAR data," presented at the ISPRS Commission 3 Symposium, 9–13 September 2002, Graz, Austria.
- El-Rabbany, A., 2002, *Introduction to GPS: The Global Positioning System*, Artech House, Boston, 196 pp.
- Ewing, C. E. and Mitchell, M. M., 1970, *Introduction to Geodesy*, American Elsevier Publishing Co., Inc., New York, 304 pp.
- Fradkin, M., Roux, M., and Maître, H., 1999, "Building detection from multiple views," presented at the Automatic Extraction of GIS Objects from Digital Imagery, 8–10 September, 1999, Technische Universität München, Germany.
- Fujii, K. and Arikawa, T., 2002, "Urban object reconstruction using airborne laser elevation image and serial image," *IEEE Transactions on Geoscience and Remote Sensing*, **Vol. 40(10)**, pp. 2234–2240.
- GEM Center, "Normalizing ALSM Intensities," Report No. Rep\_2004\_06\_02.
- Grall-Maës, E. and Beausery, P., "Mutual information-based feature extraction on the time-frequency plane," *IEEE Trans. Sig. Proc.*, **Vol. 50(4)**, 2002, pp. 779–790.
- Haala, N., 1995, "3D building reconstruction using linear edge segments," in *Photogrammetric Week*, Fritsch, D. and Hobbler, D. (eds.), Wichmann, Karlsruhe, pp. 19–28.
- Haala, N. and Brenner, C., 1998, "Interpretation of urban surface models using 2D building information," *Computer Vision and Image Understanding*, **Vol. 27(2)**, pp. 204–214.

- Haala, N. and Brenner, C., 1999, "Extraction of buildings and trees in urban environments," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54**, pp. 130–137.
- Haithcoat, T. L., Song, W., and Hipple, J. D., 2001, "Building footprint extraction and 3-D reconstruction from LIDAR data," presented at the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, 8–9 November 2001, Rome, Italy.
- Halkidi, M., Batistakis, Y., and Vazirgiannis, M., 2001, "On clustering validation techniques," *J. Intelligent Information Systems*, **Vol. 17(2–3)**, pp. 107–145.
- Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J., 2001, *Global Positioning System Theory and Practice* (5<sup>th</sup> edition), Springer Verlag, Berlin, 382 pp.
- Hug, C., 1997, "Extracting artificial surface objects from airborne laser scanner data," presented at the Automatic Extraction of Man-Made Objects from Aerial and Space Images, 2nd Workshop, Ascona, Switzerland, 5–9 May, 1997.
- Hug, C. and Wehr, A., 1997, "Detecting and identifying topographic objects in imaging laser altimetry data," *IAPRS*, **Vol. 32, part 3-2w3**, pp. 19–26.
- Huising, E. J. and Gomes Pereira, L. M., 1998, "Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 53**, pp. 245–261.
- Jain, A. K. and Farrokhnia, F., 1991, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, **Vol. 24(12)**, pp. 1167–1186.
- Jain, A. K., Murty, M. N., and Flynn, P. J., 1999, "Data clustering: a review," *ACM Computing Surveys*, **Vol. 31(3)**, pp. 264–323.
- Jekeli, C., 2000, *Inertial Navigation Systems with Geodetic Applications*, Walter de Gruyter, Inc., New York, 500 pp.
- Kaplan, E. D., 1996, *Understanding GPS: Principles and Applications*, Artech House, Boston, 554 pp.
- Krishnamoorthy, P., Boyer, K. L., and Flynn, P. J., 2002, "Robust detection of buildings in digital surface models," in the Proceedings of the 16th International Conference on Pattern Recognition, **Vol. 1**, 11–15 Aug. 2002, pp. 159–163.
- Laws, K., 1980, "Textured image segmentation," Ph.D. Dissertation, University of Southern California.
- Leon-Garcia, A., 1994, *Probability and Random Processes for Electrical Engineering* (2nd Edition), Prentice-Hall, Upper Saddle River, NJ, 589 pp.

- Levy, L., 1997, "The Kalman filter: navigation's integration workhorse," *GPS World*, **Vol. 8(7)**, pp. 65–71.
- Luzum, B. J., Slatton, K. C., and Shrestha, R. L., 2004a, "Identification and analysis of airborne laser swath mapping data in a novel feature space," accepted by *Geoscience and Remote Sensing Letters*.
- Luzum, B. J., Slatton, K. C., and Shrestha, R. L., 2004b, "Analysis of spatial and temporal stability of airborne laser swath mapping data in a novel feature space," submitted to *Transactions on Geoscience and Remote Sensing*.
- Luzum, B. J., Slatton, K. C., and Shrestha, R. L., 2004c, "Analysis of spatial and temporal stability of airborne laser swath mapping data using information theory," submitted to *Transactions on Geoscience and Remote Sensing*.
- Maas, H.-G., 1999a, "The potential of height texture measures for the segmentation of airborne laserscanner data," presented at the Fourth International Airborne Remote Sensing Conference, 21–24 June 1999, Ottawa, Canada.
- Maas, H.-G., 1999b, "Closed solutions for the determination of parametric building models from invariant moments of airborne laserscanner data," presented at the ISPRS Conference 'Automatic Extraction of GIS Objects from Digital Imagery, 8–10 September, München, Germany, IAPRS **Vol. 32, Part 3-2W5**, pp. 193–199.
- Maas, H.-G. and Vosselman, G., 1999, "Two algorithms for extracting building models from raw laser altimetry data," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54**, pp. 153–163.
- Mader, G. L., 1992, "Rapid static and kinematic Global Positioning System solutions using the ambiguity function technique," *Journal of Geophysical Research*, **Vol. 97**, pp. 3271–3283
- Matikainen, L., Hyppä, J., and Hyppä, J., 2003, "Automatic detection of buildings from laser scanner data for map updating," in Proceedings of the ISPRS Working Group III/3 Workshop '3-D Reconstruction from Airborne Laserscanner and InSAR Data,' 8-10 October 2003, Dresden, Germany.
- Maune, D. F. (ed.), 2001, *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, The American Society of Photogrammetry and Remote Sensing, Bethesda, MD, 539 pp.
- McIntosh, K. and Krupnik, A., 2000, "Integration of laser-derived DSMs and matched image edges for generating an accurate surface model," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 56**, pp. 167–176.
- Pal, N. R. and Biswas, J., 1997, "Clustering validation using graph theoretic concepts," *Pattern Recognition*, **Vol. 30(6)**, pp. 847–857.

- Park, J.-Y., 2002, "Data fusion techniques for object space classification using airborne laser data and airborne digital photographs," Ph.D. Dissertation, University of Florida.
- Priestnall, G., Jaafar, J., Duncan, A., 2000, "Extracting urban features from LiDAR digital surface models," *Computers, Environment, and Urban Systems*, **Vol. 24**, pp. 65–78.
- Rajan, S. and Slatton, K. C., 2004, "Clustering methods for nonstationary extremely low frequency digitally modulated signals," to be published as an *ASPL Technical Report*.
- Rogers, R. M., 2000, *Applied Mathematics in Integrated Navigation Systems*, AIAA Education Series, Reston, VA, 300 pp.
- Rüther, H., Martine, H. M., Mtalo, E. G., "Application of snakes and dynamic programming optimization technique in modeling of buildings in informal settlement areas," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 56**, pp. 269–282.
- Shannon, C. E., 1948, "A mathematical theory of communication," *The Bell System Technical Journal*, **Vol. 27**, pp. 379–423, 623–656.
- Shapiro, L. G. and Stockman, G. C., 2001, *Computer Vision*, Prentice Hall, Upper Saddle River, NJ, 580 pp.
- Thuy, T. and Tokunaga, M., 2001, "Wavelet and scale-space theory in segmentation of airborne laser scanner data," paper presented at 2nd Asian Conference on Remote Sensing, 5 – 9 November 2001, Singapore.
- Vaughn, C. R., Bufton, J. L., Krabill, W. B., and Rabine, D. L., 1996, "Georeferencing of airborne laser altimeter measurements," *International Journal of Remote Sensing*, **Vol. 17(11)**, pp. 2185–2200.
- Weed, C.A., Crawford, M.M., Neuenschwander, A.L., and Gutierrez, R., 2002, "Classification of LIDAR data using a lower envelope follower and gradient-based operator," in *IGARSS '02. 2002 IEEE International* **Vol. 3**, pp. 1384–1386, 24-28 June 2002.
- Wehr, A. and Lohr, 1999, "Airborne laser scanning — an introduction and overview," *ISPRS Journal of Photogrammetry & Remote Sensing*, **Vol. 54**, pp. 68–82.

## BIOGRAPHICAL SKETCH

Brian J. Luzum received his B.S. in physics and mathematics from Iowa State University in 1987 and his M.S. in civil and construction engineering from Iowa State University in 1993. From 1987-2001 he was an astronomer in the Earth Orientation Parameter Division/Earth Orientation Department at the U.S. Naval Observatory.