FUSION OF FULL WAVEFORM LIDAR AND PASSIVE REMOTE SENSING FOR IMPROVED LAND-COVER CLASSIFICATION

A Dissertation

Presented to

the Faculty of the Department of Civil and Environmental Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Geosensing Systems Engineering and Sciences

by

Hongzhou Wang

May 2016

FUSION OF FULL WAVEFORM LIDAR AND PASSIVE REMOTE SENSING

FOR IMPROVED LAND-COVER CLASSIFICATION

_____

Hongzhou Wang

Approved:

_____
Chair of the Committee
Craig L. Glennie, Ph.D.
Assistant Professor
Civil & Environmental Engineering

Committee Member:

_____
Ramesh L. Shrestha, Ph.D.
Hugh Roy & Lillie Cranz Cullen
Distinguished Professor
Civil & Environmental Engineering

_____
Hyongki Lee, Ph.D.
Assistant Professor
Civil & Environmental Engineering

_____
Saurabh Prasad, Ph.D.
Assistant Professor
Electrical & Computer Engineering

_____
Jinha Jung, Ph.D.
Assistant Professor
Engineering & Computer Sciences
Texas A&M University,
Corpus Christi

_____
Suresh K. Khator, Ph.D.
Associate Dean
Cullen College of Engineering

_____
Craig L. Glennie, Ph.D.
Assistant Professor
Civil & Environmental Engineering
Geosensing Systems Engineering &
Sciences Program Director

FUSION OF FULL WAVEFORM LIDAR AND PASSIVE REMOTE SENSING FOR IMPROVED LAND-COVER CLASSIFICATION

An Abstract

of a

Dissertation

Presented to

The Faculty of the Civil and Environmental Engineering

University Of Houston

In Partial Fulfilment

of the Requirements for The Degree of

Doctor of Philosophy

in Geosensing Systems Engineering and Sciences

by

Hongzhou Wang

May 2016

Abstract

Land-cover classification is a crucial step in interpreting remote sensing data, and the accuracy determines the reliability of the product for further downstream applications. Hyperspectral sensors have been widely utilized for classification because of the discrimination afforded by its rich spectral information and high resolution in both the spatial and spectral domains. On the other hand, LiDAR (Light Detection And Ranging) data has gained increasing interest for use in classification because it provides precise three-dimensional (3-D) data for large areas with precise 3-D location information, and therefore greatly expands the domain of available spatial information. Reflected laser energy from targets is also collected by LiDAR systems, and contains information regarding target backscattering properties. With the introduction of full waveform LiDAR (FWL), the possibility of using LiDAR for target discrimination has been enhanced due to the additional structural information acquired. The geometrical information and backscattering properties measured by FWL is complementary to the reflectance characteristics recorded within in Hyperspectral imagery (HI). Thus, the fusion of FWL and HI is highly desirable.

There has been a fair amount of research investigating the fusion of LiDAR and HI for target characterization and land-cover classification. However discrete-return LiDAR point clouds were more thoroughly investigated in this area than FWL because of their wider availability and easier interpretation. In those studies that utilized FWL, the application of waveform data was mainly limited as a reference data source to provide height information pertaining to observed targets. Furthermore, application of fused FWL

and HI data for target identification has been mostly limited to selected objects, such as trees or buildings, while the subject of land cover classification has been investigated in only a few works.

This dissertation aims to build a framework for fusing FWL and HI and to demonstrate the application of the combined data set for land-cover classification without being limited to a small sample of objects. Feature extraction methods and classifier designs are proposed considering characteristics of both data sets, and performance of the proposed methods are evaluated using two data sets collected in complex scenes by the National Center for Airborne Laser Mapping (NCALM). Experimental results show that the proposed methods are successful in extracting features from reconstructed FWL data, and the proposed classification scheme effectively utilizes the combined FWL and HI features for separating ground cover features in both data sets with over 95% accuracy.

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronyms are listed in the order of appearance in the dissertation.

LiDAR	Light Detection and Ranging
HI	Hyperspectral Imagery
RGB	Red, Green, Blue
FWL	Full Waveform LiDAR
DEM	Digital Elevation Model
DTM	Digital Terrain Model
InSAR	Interferometric Synthetic Aperture Radar
SAM	Spectral Angel Mapper
SVM	Support Vector Machine
ML	Maximum Likelihood
MLC	Maximum Likelihood Classifier
ALS	Airborne Laser Scanning
POS	Position and Orientation System
INS	Inertial Navigation Sensor
GNSS	Global Navigation Satellite System

IFOV	Instantaneous Field of View
SWF	Synthesized Waveform
PCA	Principal Component Analysis
VEDC	Vertical Energy Distribution Coefficient
SFFS	Sequential Forward Feature Selection
CCR	Correct Classification Rate

Chapter 1 Introduction

1.1 Background

1.1.1 LiDAR and Hyperspectral Imagery Fusion

LiDAR (Light Detection and Ranging) is an active remote sensing technique that measures the time of flight of short laser radiation pulses traveling to the target and back, so that the range between the sensor and target can be calculated, converted to discrete 3-D points with location and amplitude (Glennie et al., 2013). An important enhancement of LiDAR technology is the introduction of full waveform recording. Full waveform LiDAR (FWL) records the backscattered return echo as a function of time, enabling the scattering properties and geometric target characteristics to be derived from the waveform shape using appropriate signal processing methods (Mallet & Bretar, 2009). Compared to discrete return LiDAR, FWL provides denser sampling of the vertical (with airborne LiDAR) structure of the illuminated area, as well as the capability to estimate object geometric properties through waveform shape analysis. Airborne LiDAR with waveform digitizing can be particularly useful for discovering obscured targets, especially in cluttered or complex environments due to its unique ability to resolve echoes from multiple closely spaced reflectors (Tolt & Larsson, 2007).

Because LiDARs capability of obtaining 3D geometrical information of targets is not offered by other remote sensing techniques, the fusion of discrete-return LiDAR or FWL data with other sensors (especially optical and hyperspectral cameras) has become a

research topic generating significant interest in the literature. Much of the early work focused on using discrete-return LiDAR or DEM/DTM products generated from discrete return LiDAR data only. Miller conducted a study using a LiDAR derived DEM and hyperspectral imagery (HI) for canopy parameter analysis where the trees in the DEM were separated from ground by computing local slope (Miller, 2001). A similar empirical ground-tree separation method was applied to LiDAR data for merging with InSAR data and providing the measurements of "true surface topography" to estimate surface topography and vegetation heights (Slatton et al. 2001). Discrete-return LiDAR has also been broadly used in forest characterization in combination with hyperspectral imagery or visible photograph with local filtering algorithms (McCombs et al., 2003; Popescu et al., 2004) or DTM/DSM generation (Suárez et al. 2005; Geerling et al. 2007; Dalponte et al. 2008; Asner et al. 2008; Kempeneers et al. 2009; Breidenbach et al. 2010). An improvement in the fusion of LiDAR and hyperspectral imagery was also made by introducing both height and intensity metrics from the LiDAR point cloud data based on existing vegetation models (Erdody & Moskal, 2010; Forzieri et al., 2010; Swatantran et al., 2011; Dalponte et al., 2012). Other applications that have fused discrete-return LiDAR and visible photography or HI include building detection (Rottensteiner et al., 2005; Sohn & Dowman, 2007) and natural hazard monitoring (Mason et al., 2007).

The fusion of FWL with passive remote sensing data has been addressed in the literature; however, the work has predominantly been limited to applications where the FWL data is first converted to discrete return point clouds prior to analysis (Asner et al., 2007; Anderson et al., 2008; Kulawardhana et al., 2014; Paris & Bruzzone, 2015). In these instances, the FWL return energy distribution is discarded before the analysis of the

fused FWL and HI dataset is undertaken. Alternatively, several researchers have first decomposed the FWL data based on waveform modeling before undertaking feature level fusion with HI (Sarrazin et al., 2011; Jung, 2011). Each waveform is decomposed into separate return echoes reflected from different targets, and the characteristics of the echoes (e.g., amplitude and width) are used as features for classification. This is an improvement over simply using target locations, but still involves simplification of the waveform based on the assumptions of the decomposition model, e.g., a Gaussian model. While this may be suitable for land targets in simple environment, it is often not the case in complex environments where pulse shape may be distorted by cluttered targets. In these cases, a model based decomposition of the waveform may remove significant information from the full waveform backscatter signature.

1.1.2 Fused Data Classification

Based on the techniques developed for fusion of LiDAR and HI or other passive remotely sensed data, it is possible to utilize the fused data set for land cover classification. This topic was first analyzed theoretically in 1998 (Perry et al., 1998), where basic problems such as geometric and radiometric correction and HI band selection were briefly presented. In most early applications that fused LiDAR with HI for improved classification, only elevation data from LiDAR was used. For example, (Mundt et al., 2006) examined the segmentation of sagebrush using fused LiDAR and HI data, where the LiDAR data was used to generate a ground surface model, as well as for computation of some physical properties for targets; the interpolated LiDAR data was registered to HI by aligning the LiDAR intensity map with the HI pixels. Another study of fused FWL/HI classification was performed in 2007 for a coastal area with eight

segmented classes, based on data from The Compact Hydrographic Airborne Rapid Total Survey (CHARTS) system, which collects LiDAR data, HI, and RGB images simultaneously (Wozencraft et al., 2007). Discrete return LiDAR elevation data was used to generate a DEM that was mainly used as ancillary information, and the evaluation of classification accuracy was not presented. Another study of supervised classification with a relatively complete accuracy evaluation is presented in (Sugumaran & Voss, 2007) for separating tree classes in an urban environment (only trees were considered for classification; all other targets were discarded before classification). Therein, LiDAR elevation data was used to pre-segment the data, and enhance the classification by using some elevation based rules. In addition to the use of absolute elevation, the penetration depth (elevation different between top and bottom layer returns) of LiDAR is very useful for applications that characterize plant distributions in combination with LiDAR derived DEMs (Sadro et al., 2007).

With the continued evolution of studies attempting LiDAR data fusion and classification, the variety of LiDAR features extracted and utilized has been significantly expanded. Intensity information from LiDAR data, which is correlated with the backscattering properties of the illuminated surface, have been gradually introduced into the fusion process to assist in classification. The intensity data can be transformed into raster images in the same manner as DEM generation (Voss & Sugumaran, 2008; Dalponte et al., 2008). Then the registered LiDAR elevation and intensity images, together with the selected HI bands are fed into the chosen classifiers together to be employed in a supervised classification scheme for a multi-class separation problem (Dalponte et al., 2008). In addition to elevation and intensity, other kinds of information

that can be extracted from LiDAR point clouds have also been used as features for classification including: point density and height distribution (Koetz et al., 2008; Puttonen et al., 2010), morphological attribute features (Pedergnana et al., 2011; Debes et al., 2014), and metrics representing vegetation structure (Alonzo et al., 2014). These features are mostly generated from rasterized LiDAR point cloud data, and are used as a layer of features both independently and in association with other layers.

Limitations on the use of LiDAR point clouds exist because of their sparse and irregular spatial sampling and the limited metrics that can be extracted from intensity and elevation information. Waveform LiDAR, however, is expected to enhance classification tasks because it records the properties of all targets illuminated within the laser pulse diffraction cone. For example, assuming a portion of a laser pulse is able to hit the ground after interacting with top layer of an object, then the elevation of the top surface can be determined as the full width half maximum (FWHM) difference between the first peak and last peak of the waveform; this is a commonly used waveform metric (Sarrazin et al., 2010). Furthermore, height metrics derived from discrete LiDAR point cloud features can also be derived from LiDAR waveforms; and applied more precisely given the finer resolution offered by waveform processing (Swatantran et al., 2011). Waveforms can also be used as a multi-dimensional array collected at a specific location; dimension reduction and feature analysis approaches can then be applied to the raw waveform data, or even derivatives of different orders (Sarrazin et al., 2012). A waveform usually has more than one peak if acquired over complex environments, and thus the width and amplitude of each individual echo decomposed from the original waveform are often useful information for target discrimination (Heinzel & Koch, 2012).

However, to take full advantage of the FWL data, it would be optimal to collect all the raw waveforms corresponding to one HI pixel and utilize the collocated HI spectral and LiDAR waveform characteristics for efficient classification. In order to register raw waveforms to one HI pixel, voxels need to be built on the basis of the HI sample frame to essentially rasterize the FWL data (Buddenbaum et al., 2013; Wang et al., 2013). Then the return intensity profile of each voxel can then be used as a normalized feature derived from multiple laser pulses. Similar methodology has also been proposed for generating "pseudo-waveform" from voxelized LiDAR point clouds (Jung et al., 2014; Y. Zhang et al., 2015; Muss et al., 2011), which are then used in combination with HI for object-based feature extraction and classification. In most of these works, however, the waveform of each column of voxels is reconstructed at a much coarser spatial resolution compared to the original waveform.

From the perspective of classification algorithms, many of the popular HI classification algorithms have been applied to fused FWL/HI data sets. These methods include principal component analysis (PCA) associated with discrimination analysis for class separation (Sarrazin et al., 2010; Sarrazin et al., 2012), random forests (Guo et al., 2011), support vector machines (SVM) (Heinzel & Koch, 2012; Buddenbaum et al., 2013), spectral angle mapper (SAM) (Buddenbaum et al., 2013), and maximum likelihood (ML) (Buddenbaum et al., 2013). A comparison of SVM, SAM and ML can be found in (Buddenbaum et al., 2013), where waveform data was used in the form of multi-band images in a similar manner as HI and the results suggested that SVM outperforms SAM and ML for the combined data set. This dissertation will use and evaluate SVM and

ML classifiers within a proposed classification framework, because of their proven efficiency in similar applications.

Multiple classes classification problems are often solved by training the classifier using data samples from all classes, and all validation samples are labeled using a common classifier (Lu & Weng, 2007). However, with a large number of classes it may be difficult to train a single classifier, especially when input data is from multiple sources. The most important reason is because a single classifier use a common feature subset and optimization parameters for separating all classes, whereas each class may have its own unique optimal feature signature, thus the performance of the whole system may be sacrificed. On the other hand, pairwise classification solves the multiple classification problem as a series of two class separation problems (Hastie & Tibshirani, 1998), an individual classifier with specific feature subset is available for each two class pair. Because of this characteristic, pairwise classification can be used for both single data source classification (Kumar et al., 2001) and fused multiple data sets (Crawford, 1999). Pairwise classification framework is adapter for use on fused FWL and HI features in this dissertation.

1.2 Research Opportunities

1.2.1 LiDAR Waveform and Raster Data Registration

A raw data fusion strategy for merging of FWL and HI data at the same spatial level, e.g., by interpolating the FWL measurements to the same sampling interval as the HI imagery (i.e., same raster resolution) is required. However, the fusion of HI and FWL at the raw data level is difficult due to the different data acquisition modalities of each

sensor and the disparate data architecture; hyperspectral imagery records spectral response in a pixel-based raster format, whereas FWL data is acquired with an uneven structure and irregular distribution over the imaged scene. A few references have discussed methods of deriving vertical waveforms over a raster grid from the original waveforms (e.g., Wu et al. 2012; Hermosilla et al. 2014), but the fusion of vertical waveforms with other remote sensing observations is not discussed, or limited to chosen tree types utilizing height information from the FWL in combination with HI (Buddenbaum et al., 2013).

To address the issues involved in proper registration, a method for locating the FWL data and HI in a common geographical frame is required utilizing the high precision navigation systems integrated with the laser scanner. Since a single target can receive multiple laser return pulses from various scan angles, especially when located in an area of multiple overlapping scan strips, it is desirable that a single waveform is generated for each column of voxels in order to preserve the features from all waveforms intersecting the column. This requires the development and evaluation of appropriate waveform synthesizing methods.

1.2.2 Feature extraction

Appropriate information retrieval methods are necessary for extracting features from FWL data after they have been reconstructed in a common raster framework with HI. Information contained in FWL can be considered as four-dimensional arrays, i.e., three-dimensional coordinates plus an intensity value for each sample. However, FWL sample intensity values are not only dependent on target characteristics, but also on the laser scanner pulse parameters and the incidence angle between the laser pulse and

targets. The process of converting intensity into a radar cross-section of the target is normally referred to as calibration (Wagner et al., 2006). Direct calibration of FWL, however, is hard to achieve because the power received by each target and the incidence angle is difficult to model (Abed et al., 2014), especially when laser pulses travel through complex, multiple target environments. Therefore, structural information needs to be included in feature extraction for FWL, i.e., the trend of intensity changes in the vertical direction, instead of directly using the individual intensity values in classification. To extract such information from FWL data, features need to be designed such that geometric information and reflected laser energy distribution are both included, and the use of feature is not affected by the absolute intensity value of the waveforms. With the synthesized waveforms reconstructed for correspondence with HI pixels, such features extracted from synthesized waveform properly represent the properties of the targets within the voxel column.

1.2.3 Classification and Feature Selection

A stacked feature set comprised of both FWL and HI features offers information on spectral reflectance, backscattering properties and target geometry. Such a feature set should be able to enhance classification over consideration of each set independently. The individual performance of HI and FWL features for distinguishing different classes can vary significantly; e.g., it is difficult to distinguish surfaces made of different material with similar structure using FWL alone, whereas HI has a diminished classification accuracy of targets shaded by nearby taller objects. Given these differences, an optimal classification technique for fused features from HI and FWL must be able to choose subsets of features or even different classifiers, for different target classes. A practical

solution is to utilize the pairwise classification framework, which discriminates target classes in pairs. Since the scope of each individual classification is reduced, the optimal feature subset become easier to determine for each class pair, with the assistance of appropriate feature selection tools.

1.3 Objective and Research Contributions

The intent of this dissertation is to build and evaluate methods to integrate FWL data and HI, and to improve performance of land-cover classification using the merged data set. The combination of FWL and passive remote sensing techniques leverages detailed target geometric and spectral target properties for enhanced classification accuracies, with potential benefit to land use evaluation, natural resources management, target detection and urban planning applications. The proposed methods for FWL reconstruction and information retrieval will broaden the application of FWL data and can be extended to the fusion of FWL data with other raster format remote sensing data, e.g., orthophotos.

Since a broad variety of classification algorithms can be applied to the fused FWL and HI data, the performance of different classification algorithms are evaluated in supervised land use classification. The evaluation criteria takes both overall accuracy and consistency among different classes into consideration, along with an analysis of how the classifiers interact with different input data sets, i.e., different kinds of single data sets and fused data sets, in order to reveal the contribution of adding FWL features to the classification.

Specifically, the novel contributions of this research are:

(1) Voxelization of FWL data by adapting multiple waveforms into one common column of voxels and the subsequent reconstruction of the voxelized waveforms into single vertical waveforms at the same resolution level as original waveform.

(3) Development of a vertical energy distribution coefficient (VEDC) as a novel FWL feature.

(4) Use of a pairwise classification technique to improve classification performance of the fused data set.

(5) Comparison and quantitative analysis of different classification methods operating under the pairwise classifier framework.

1.4 Thesis Organization

In Chapter 2, the background of LiDAR and FWL systems is presented, together with a summary of waveform processing techniques for feature extraction. Chapter 3 introduces hyperspectral sensors and imaging geometrics, analyzes the characteristics of HI, and then covers the development of feature extraction and classification techniques for HI. Review of land cover classification is introduced in Chapter 4, unsupervised classification, supervised classification algorithms and classification accuracy assessment are covered in detail. The proposed methodology is presented in Chapter 5, including voxelization and synthesizing of FWL, feature extraction of FWL and HI, and classification based on the framework of a pairwise classifier. The choice of individual classifier to implement the pairwise classification is also discussed. Chapter 6 presents detailed introduction to the two data sets used in the dissertation. The first data set contains different kinds of both artificial and natural objects, while the second data set

features a well labeled vegetation training set which is based on multiple remote and field observation with professional interpretation. Sensor system characteristics, acquisition parameters, analysis of the land cover classes of the selected areas, and the methods of obtaining ground reference data are discussed for both data sets. Experimental results and analysis are presented in Chapter 7, proposed methods are first evaluated using the first data set with different possible combinations of algorithms, and the performance of the best combination is applied to the second data set to show the capability of separating different classes with considerable similarities. Finally, a summary of the dissertation and discussion of future work in the topic can be found in Chapter 8.

Chapter 2 Full Waveform LiDAR

2.1 Airborne Laser Scanning

Airborne laser scanning (ALS) systems are based on light detection and ranging (LiDAR) techniques, where the source of light is a laser (Shan and Toth, 2009). Since its inception in the early 1990s, ALS has been widely deployed with both academic and commercially available sensors. The most common way of utilizing laser ranging is to measure the time-of-flight of a brief laser pulse traveling from the sensor to an object and returning to the detector. With the speed of light represented by *c* and the measured elapsed time by *t*, the range *R* between the laser and object can be determined as

$$R = c \cdot t/2. \tag{2-1}$$

The accuracy of the range measurement is dependent upon the timing accuracy and resolution, as well as the speed of light in the transmitting medium. The speed of light is effected by the laser wavelength and the mean atmospheric temperature and pressure; the latter two are measured each time for speed of light adjustment. In order to measure the travelling time of outgoing pulse accurately. A very stable oscillator is used to control a high-speed counter used as the timer. A small part of the outgoing pulse is redirected to a photodiode in order to trigger the timer, which occurs when a certain power threshold is detected. Similarly, after the reflected pulse is collected by the receiving optics, a photodiode is again used to convert the received energy to an output voltage. When the output voltage climbs over a pre-defined threshold, the time counter stops and the time-of-flight is obtained, upon which the range between the laser source and object can be computed. The simplified process is illustrated in Figure 2-1.

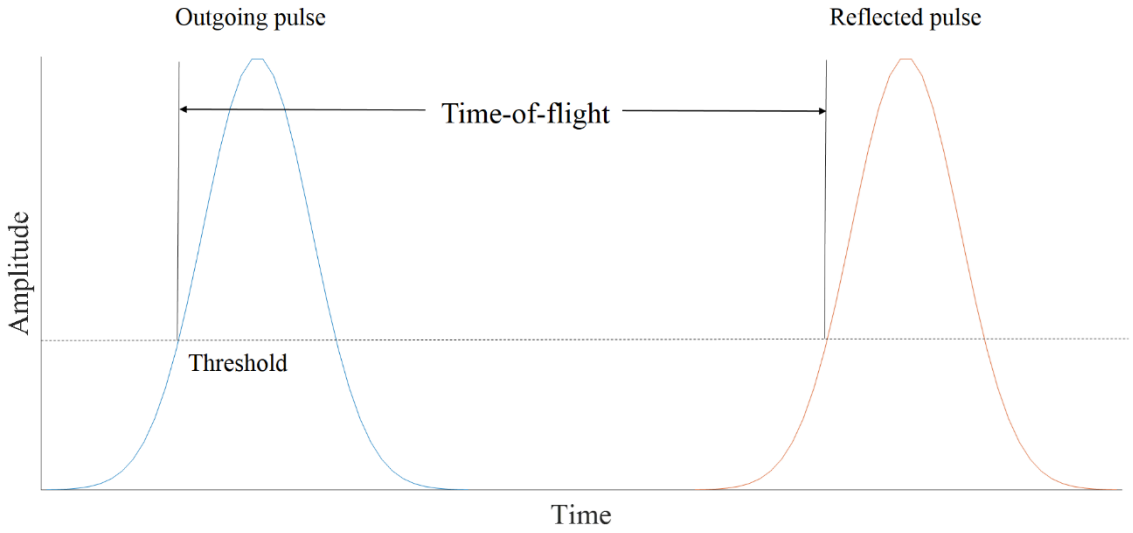


Figure 2-1 Laser ranging time counting

In order to derive 3D positions from LiDAR ranges, two degrees of controlled motion are needed. By mounting the laser ranger on a moving airborne platform, continuous range measurements are obtained along a ground profile coinciding with the movement of the platform (Shan and Toth, 2009). If the pulses are not always fired vertically down to the ground, but in a systematically varying across-track direction, enabled by a scanning mechanism, then a swath of coverage is obtained with the movement of the platform. Typical ALS scanning mechanisms include oscillating and nutating mirrors, rotating polygons, and fiber scanners (Wehr & Lohr, 1999). A typical ALS system consists of the following important components: scanner unit, position and orientation system (POS), and hardware and software systems serving as control, processing and storage agents for the entire system.

The scanning unit is comprised of the laser generator, transmitting and receiving optics, and scanning mechanism. The POS system is a key component of an ALS (Wehr

& Lohr, 1999), and consists of an integrated GNSS (Global Navigation Satellite System) and INS (Inertial Navigation System). The GNSS provides location and velocity information whereas the INS provides the attitude of the platform body. Because of its high data rate, an INS is able to provide accurate position and velocity information between the GNSS measurement update intervals; each GNSS update, in turn, is used to rectify the residual error of the INS accelerometer and gyro sensors (Schwarz et al., 1993). Such integration enables the kinematic positioning of ALS measurements. With the help of a POS, the ALS data can be referenced to a global geographical coordinate frame (Zhang & Shen, 2013), allowing the analysis of ALS data as an independent data source. The hardware system also contains a data processer, encoder, and a storage device for the collected data (laser scanning data and POS data).

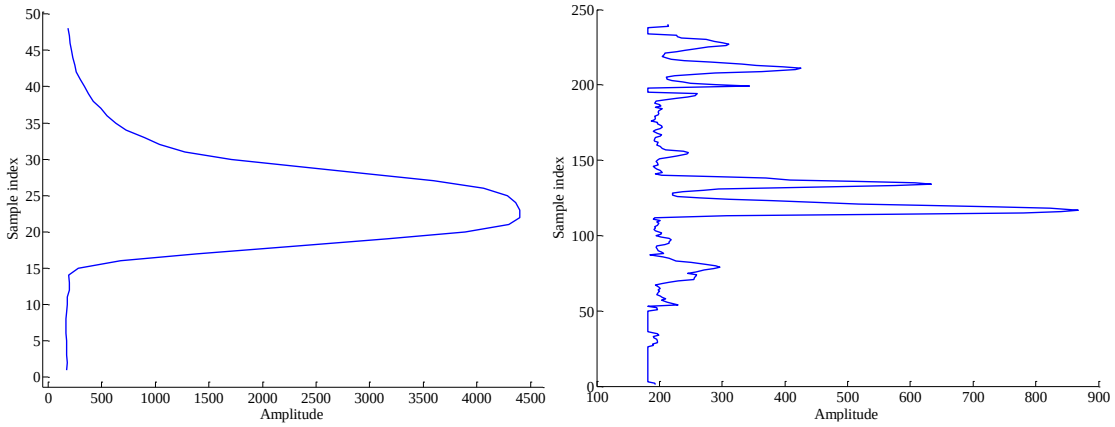
The most common approach for extracting object information from reflected laser pulses is analog detection, which operates directly on the electrical voltage output from the photodiode. Traditional ALS systems detect distinct peaks from the voltage time series using a constant fraction discriminator (Wagner et al., 2004), and extract them as discrete returns reflected from objects. In such systems, the waveform is inverted and delayed by a fixed time and added to the outgoing pulse. The stacked signal yields a constant fraction waveform and the peaks can be detected at the zero-crossing points of the waveform (Shan and Toth, 2009). Also the maximum voltage value of the detected peak is recorded as a descriptor for return energy. A single record containing the location and amplitude of a distinct peak is defined as a discrete return. The discrete returns collected by an ALS system over a period of time are combined into a single point cloud data product (Korhonen et al., 2011).

2.2 Full Waveform LiDAR

The backscattering properties of illuminated objects are correlated with recorded LiDAR return intensities (Wagner et al., 2006). The energy reflected to the receiver is converted by an analog detector into a voltage, yielding a time series of signal strength. Digitization of the entire backscattered pulse time series using analog-to-digital converters generates a full waveform recording of the reflected energy within the laser pulse footprint (Mallet & Bretar, 2009). The time series are samples with a fixed interval, typically 1-2 ns, leading to approximately 15-30cm distance between adjacent waveform samples. Waveform amplitudes are typically quantized at 8-12 bits. Compared with discrete return systems, which can record up to 4 returns per emitted laser pulse, full waveform data has a much higher density, since a waveform typically contains hundreds of samples (Anderson et al., 2016). Although not all samples in waveform correspond to a detected return echo as with discrete data, users still can expect consistent spatial sampling of the observed space with constant range resolution.

The majority of FWL systems sample and record both the transmitted and return pulse. By comparing the pulse shape before and after being reflected by the object, it is possible to characterize the vertical structure and surface roughness of the reflecting object (Reitberger et al., 2008). For example, for a flat and solid surface like bare ground, the reflected waveform should have a shape similar to the outgoing pulse with only a scale difference (Pirotti, 2011). However, when the pulse is reflected from a very rough or slanted surface, the waveform can be significantly broadened from the original shape. For distributed targets like vegetation canopies, a single laser pulse could travel through the top layer of leaves, branches, and stems as well as intercept the underlying ground

shadowed by the canopy (Sun & Ranson, 2000). For such targets multiple peaks in the return waveform are possible, an example is shown in Figure 2-2. Since the energy contained in outgoing pulse is almost constant, the pattern of amplitude changes of the return waveform can reveal how the laser energy was reflected at different level of the targets, which is useful for understanding the vertical structure under the top layer. Therefore, in comparison to discrete point clouds, post processed waveform analysis is expected to reveal target structure in complex environments where waveform energy profiles lack separable peaks due to the distributed nature of reflecting targets (Hofton et al., 2000).



(a) (b)

Figure 2-2 Waveform example: (a) Outgoing pulse; (b) Return waveform from tall vegetation

2.3 Waveform Data Processing Techniques

2.3.1 Waveform Decomposition

Appropriate post-processing is needed to extract information from waveform data. A common approach is to detect distinct peaks from each waveform and then analyze the echoes at the peak locations. Because waveforms are digitized from converted return energy that is a result of the interaction between the transmitted pulse and the target

backscattering characteristics, a basic assumption regarding the response function of a target, or cluster of scattering targets, is essential. A scatter cluster can be modeled by a Gaussian response function (Wagner et al., 2006)

$$\sigma_i'(t) = \widehat{\sigma_i} e^{-\frac{(t-t_i)^2}{2S_i^2}}, \tag{2-2}$$

where $\widehat{\sigma_i}$ and S_i are the amplitude and standard deviation of cluster i, and t_i represents the location of the cluster. The waveform received by the detector can then be modeled as a combination of individual functions of scattered energy in the laser pulse footprint (Wagner et al., 2006),

$$P_r(t) = \sum_{i=1}^{N} \frac{D_r^2}{4\pi R_i^4 \beta_t^2} P_t(t) * \sigma_i'(t). \tag{2-3}$$

In this equation, P_r is the received waveform, P_t is the transmitted waveform, D_r is the aperture diameter of the receiver, R_i is the range to the scatterer, β_t is the transmitted beam width, and σ_i' is the target scatter response function. In most FWL systems the transmitted waveform is either analytically known or recorded during transmission, and the received waveform can therefore be modeled based on the shape of the outgoing pulse. If the outgoing pulse of the LiDAR system can also be approximated as a Gaussian function, then the waveform received by the detector can be modeled as a combination of Gaussian functions (Hofton et al., 2000; Wagner et al., 2006). This assumption is normally valid because existing ALS systems typically have Gaussian –like outgoing pulse shapes (Kirchhof et al., 2008).

To separate individual return echoes from a complex waveform containing a mixture of Gaussian returns and study the properties of each return echo, a peak detection

algorithm is required (Pan et al., 2015). A basic approach is to use numerical derivatives to find the local maxima of the waveform. First order derivatives can be used to directly find the stationary points at the waveform maxima. An alternative is to use second order derivative to search for inflection points, and then determine the peak location using the symmetry property of the function. Derivatives based methods are affected by noise in the waveform data, therefore, threshold based filtering of detected peaks are usually applied based on knowledge of waveform shape (e.g., approximate noise distribution). After peaks are identified using these approaches, curve fitting algorithms can be employed to find the parameters of individual echo functions that approximate the return waveform (Tolt & Larsson, 2007; Qin et al., 2012; Adams et al., 2012).

Another class of waveform peak detection and decomposition algorithms is direct deconvolution. Notable algorithms include the Wiener filter method (Jutzi & Stilla, 2006), expectation maximization (Parrish & Nowak, 2009), and B-spline convolution (Roncat et al., 2011). These algorithms directly solve for the surface response function, which is independent of the FWL system, thus the analytical form of outgoing pulse function is not required. Comparison between waveform modeling based decomposition and direct deconvolution is not sufficiently addressed in the literature, except in (Neuenschwander 2008) where the direct deconvolution is reported to improve range determination and extract more structural information for canopy over modeling decomposition when both were applied on Geoscience Laser Altimeter System (GLAS) data. In general, waveform modeling based decomposition is more extensively used than direct deconvolution, because it divides the waveform into echo functions with analytical form, and the parameters of echo functions can be associated with physical properties - thus there are

normally more convenient for further analysis and classification. In addition, the waveform modeling based methods do not require estimate of noise spectrum of waveform.

2.3.2 Waveform Spatial Characteristics and Reconstruction

Similar to discrete return point cloud data, once the waveform peaks are identified, the 3-D peak locations can be determined with the help of the onboard POS system. However, to better utilize the waveform data, the entire return waveform can be considered instead of just the peak locations. An important spatial characteristics of ALS return waveform profiles is the off nadir angles. This is due to the scanning mechanisms, which steer the laser beam via a mirror through a field of view of up to 60 degrees (±30 degree from nadir). As a result, a majority of the FWL return energy profiles are at off nadir angles. The scan angle histogram of raw waveforms from a FWL survey acquired with an oscillating mirror with a scan angle range of ±20 degrees is shown in Figure 2-3, with an illustration of the can pattern on the ground shown in Figure 2-4.

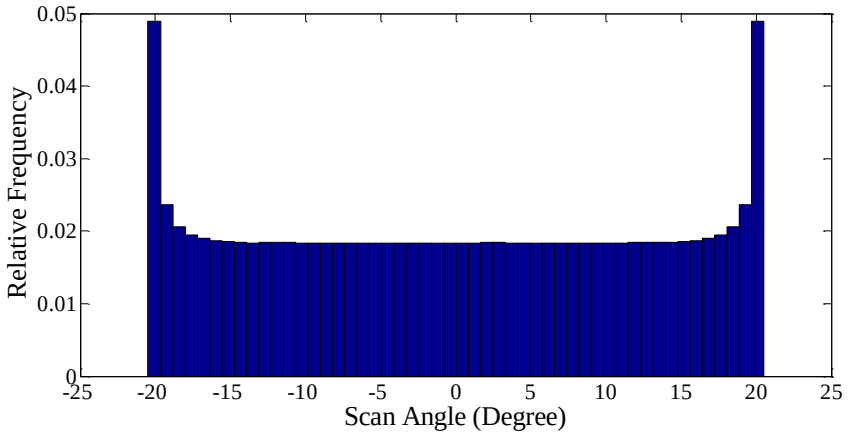


Figure 2-3 Scan angle histogram for raw waveforms

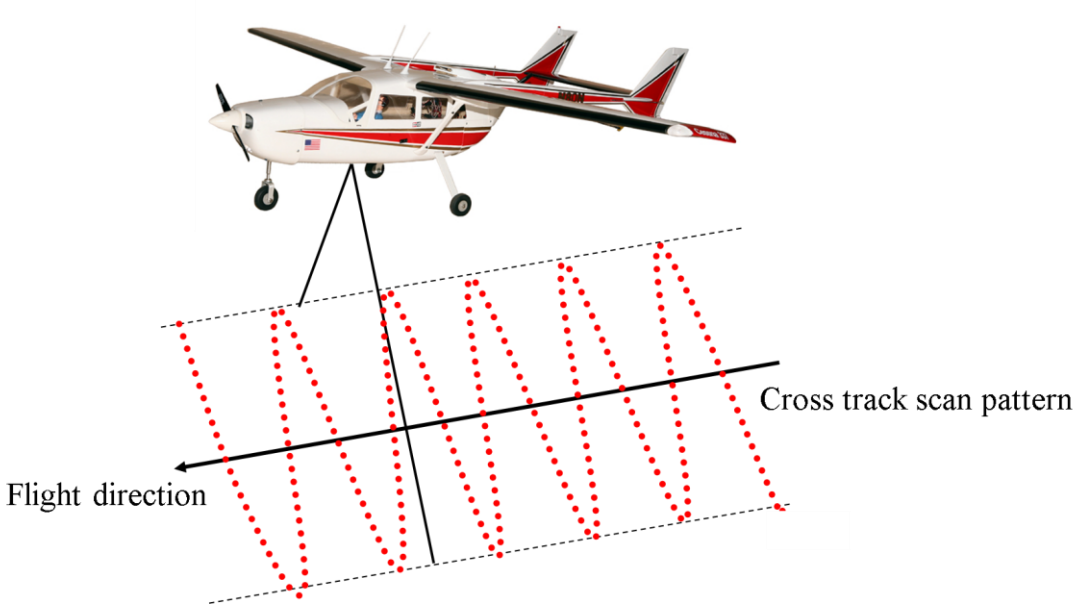


Figure 2-4 Scan pattern on ground with an oscillating scanning mirror

For the purpose of utilizing FWL data in classification applications as well as in combination with other remote sensing imagery, FWL data needs to be converted to 2D raster data or reconstructed into a vertical structure in a uniformly sampled 3D raster space. Several previous studies have attempted to address this problem and can be broken into two categories: generation of a raster format model directly from the FWL data (Lefsky et al., 2007), or decomposition of the FWL data followed by georeferencing of individual decomposed echoes. Raster models are normally built by assigning waveform peaks according to their ground coordinates, and usually only location information of the peaks is reserved and further utilized (Hollaus & Höfle, 2010). Decomposition based methods first determine the 3-D coordinates of each waveform peak location, followed by reconstruction of new waveforms by stacking of all the individual echoes located in a common voxel (Jung 2011).

2.3.3 Waveform Feature Extraction

Feature extraction for FWL data or "pseudo-waveforms" generated from discrete return LiDAR point clouds have been discussed in some previous studies, for example (Mallet et al., 2011; Parrish et al., 2014). These waveform features include height and intensity based metrics computed from the original waveforms, and parameters of the Gaussian components obtained from waveform decomposition. An evaluation of some waveform feature metrics can be found in (Parrish et al., 2014). Height and intensity based FWL metrics, which can only be retrieved once the waveforms are georeferenced, are widely used in canopy characterization applications. The most commonly used metrics include: height of last return (HLR), LiDAR canopy height (LHT) or penetration depth (PD), height of median energy (HOME), maximum amplitude of waveform (MA), height/median ration (HTRT), and simple ground return ratio (GRND) (Drake et al., 2002). LHT is computed as the distance between the locations of the first sample in a waveform above a certain threshold and the center of the last return (HLR). This first sample is considered as canopy top if the illuminated area is forested. For other types of targets, the LHT can be determined the same way and used as the target top layer height above ground. HOME is determined by finding the median intensity value of the whole waveform, and HTRT is the ratio of HOME to LHT. GRND works as an indicator of the penetration of the laser pulse; it is determined as the total intensity of samples around the ground level divided by the total of other samples' intensities. GRND can be extended to compute the total energy of the waveform (usually approximated by area under curve), canopy energy (or target energy, area under curve corresponding to LHT), and the ratio of these two values (ER). Rise time (RT) and fall time (FT) are sometimes used in

association with the above features: RT represents the duration for energy to climb from 10% to 90% of the amplitude of the first return, FT represents the duration for energy to drop from 90% to 10% of the last return amplitude for the trailing edge. Skewness (SW) is also used as a measure of symmetry of the waveform shape, which provide rough intensity distribution information. The HLR, MA, PD and SW features are also used in this dissertation as part of waveform data features.

Waveform decomposition based features are obtained in a similar manner as discussed in section 2.3.2. The peak locations of the echoes retrieved are used for georeferencing and then the other properties, e.g., amplitude, width and rising time can be used as features for classification attributes (Jung 2011; Guo et al. 2011). Peak locations of echoes can be used to generate height based metrics similar to the discussion in the previous paragraph; the amplitude of echoes can be used as indicators of target backscattering properties; width and rising time of echoes reflect geometric shape of the object. Decomposition based features have been successfully used in urban (Guo et al. 2011) and forested areas (Hovi et al., 2016).

Chapter 3 Hyperspectral Imagery

3.1 Hyperspectral Remote Sensing

The term “hyperspectral imaging” was first introduced in the field of imaging spectrometry in 1985 (Goetz et al., 1985). This passive remote sensing technique was first defined as “the acquisition of images in hundreds of contiguous, registered, spectral bands such that for each pixel a radiance spectrum can be derived.” Hyperspectral spectra may be measured for all spectral regions, i.e., VIS (visible, wavelength 400-700 nm), NIR (near infrared, 700-1400 nm), SWIR (shortwave infrared, 1.4-3 μm), MWIR (midwave infrared, 3-8 μm), and LWIR (longwave infrared, 8-15 μm). However most hyperspectral imagers operate within the spectrum from 0.35 to 2.5 μm, with usually 100 to 288 bands. Hyperspectral imagers have been extensively utilized on satellite, airborne, and terrestrial platforms (Shippert, 2003).

Four essential components comprise hyperspectral remote sensing system: the radiation source, atmospheric conditions, the hyperspectral sensor and the observed surface (Manolakis et al., 2003). In most remote sensing applications, sunlight is the source of radiation. The solar energy passes through atmosphere before interacting with the imaged surface, is reflected by the imaged surface, and finally received by the sensor. The received spectrum is actually the solar spectrum modulated by atmospheric effects and the imaged surface reflectance. The reflectance spectrum, defined as a wavelength (λ) dependent quantity, can be expressed as (Manolakis et al., 2003)

$$reflectance\ spectrum(\lambda) = \frac{reflectance\ radiation\ (\lambda)}{incident\ radiation\ (\lambda)}. \quad (3\text{-}1)$$

Atmospheric effects (e.g., absorption, scattering) need to be compensated before the study of the reflectance spectrum. This is normally corrected by using a radiative transfer model, the first algorithm proposed was Atmosphere Removal Algorithm (ATREM), where atmospheric gas transmission spectrum is estimated for modeling absorption, and the scattering is modeled due to the atmospheric molecules and aerosols (Gao et al., 1993). While the ATREM keeps evolving (Thompson et al., 2015), other models such as FLAASH (Matthew et al., 2002) and HATCH (Qu et al., 2003) have also been developed with enhancements to certain aspects of ATREM. After the radiometric compensation is applied, the reflectance spectrum of surface is obtained.

Given the broad range of wavelengths, the recorded spectrum is an excellent source for discriminating surface material, an example is shown in Figure 3-1. The set of all unique spectras in a given scene represent spectra endmembers. Since the spatial resolution of hyperspectral sensors is limited, the spectrum of each pixel can be regarded as a mixture of the endmembers (Keshava & Mustard, 2002). The contiguous sampling properties of a hyperspectral sensor make it possible to examine the correlation of collected pixel spectra with spectral data bases to improve the apparent SNR of the collected data. Also with knowledge of spectrum range and interval, it is possible to use statistics-based unmixing techniques on pixel spectra to extract parameters regarding surface material components (Bioucas-Dias et al., 2012). Spectral resolution also limits the efficacy of acquired imagery, with different applications imposing varying spectral resolution (i.e., bandwidth) requirements. For geological applications, 10nm spectral

resolution is satisfactory; whereas new applications of HI such as the study of vegetation fluorescence requires a bandwidth of less than 1 nm (Guanter et al., 2007).

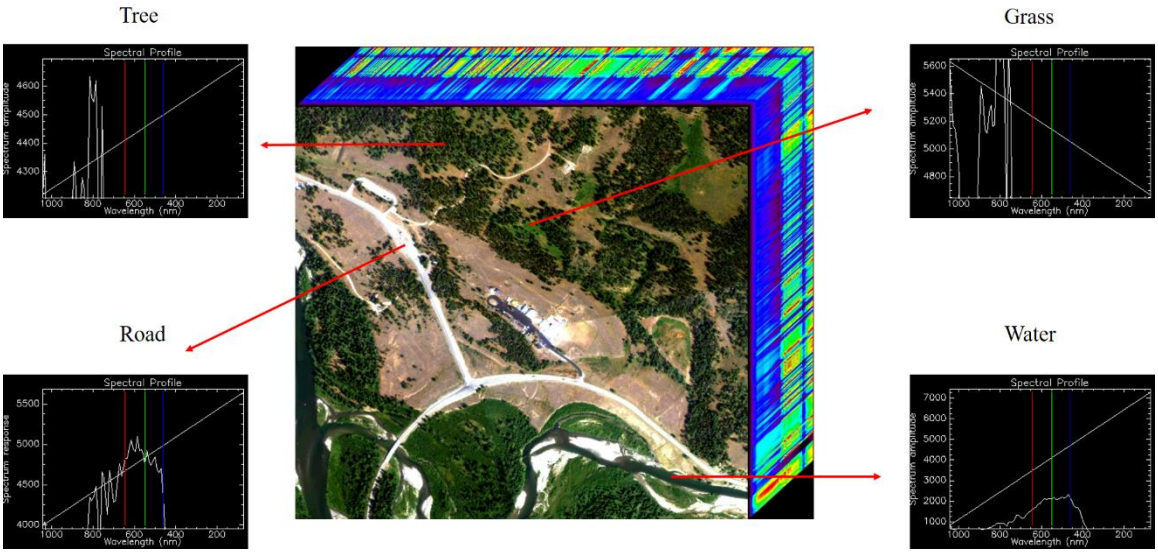


Figure 3-1 Spectral content of HI

Hyperspectral imaging systems acquire imagery using a number of detector elements, and each detector element is usually dedicated to one pixel. The ground area coverage by each pixel (i.e., spatial resolution) is determined by the instantaneous field of view (IFOV) of the system. The IFOV is determined by the altitude of the platform, the size of the detector array and sensor optics. Sensor spatial resolution limits the minimum size of the object that can distinctly detected from its surroundings in the imagery.

In order to collect 2D spatial information utilizing a 1D spatial detector array, scanning methods are required. Three kind of scanning approaches are used for 2D HI collection. The first approach is whiskbroom scanning, which was initially used by NASA on its Airborne Visible/Infrared Imaging Spectrometer (AVRIS) system (Mouroulis & Green, 2003). Whiskbroom sensors measure the spectrum of one pixel at a time, by moving the mirror, which reflects the light to a single detector, two

dimensionally in the observed area to build up a 2D image. Since the spectrum of all pixels in the 2D grid are collected one-by-one, this scanning method requires a longer imaging time. Its advantage is that the light traveling path to the sensor optics is the same for all pixels. The second approach is pushbroom scanning (Lawrence et al., 2003), which records an entire line of an image at the same time instead of just one pixel. This method requires the array of detector elements to be moving along one dimension of the 2D grid during image acquisition (the geometry is illustrated in Figure 3-2). This method is usually employed in an environment where either the imaging platform or the observed object is moving. Pushbroom scanning is extensively used because of the fast scanning time and the capability of being adapted to different platforms; the airborne HI sensor used in this dissertation is also a pushbroom sensor. The last approach, tunable filter imaging, is different from the first two as it acquires the spectral information iteratively. The tunable filter approach is an area scanning method, which collects one spectral band after another for the entire scene (Gat, 2000). Neither the scene nor the HI unit is moving in this approach, and it is only efficient when the number of acquired bands is low.

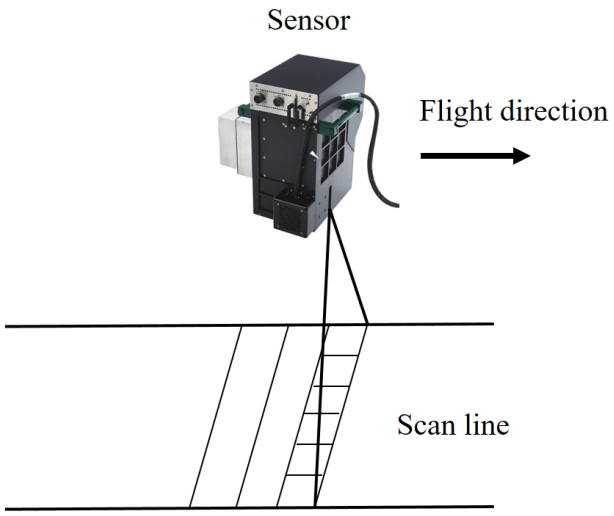


Figure 3-2 Airborne pushbroom scanning to acquire 2D HI

3.2 Hyperspectral Imagery Processing

In the Earth science domain, HI has been applied to many subjects such as vegetation studies, soil science, geology, biochemical studies, water quality monitoring, and atmospheric characterization (Manolakis et al., 2003). HI has also played a very important role in land cover classification applications using both supervised clustering and unsupervised pixel clustering, because of its capability to reveal the composition of object materials (Kerekes & Baum, 2002) with ample spectral information. When the objects of interest do not have a predefined spectrum shape, or when the relationship between received spectrum and the "real" spectrum of the material is unknown, i.e. the material composition is not accurately recognizable by examining a single spectrum, clustering of pixels is an efficient way to help interpret the imagery and control the overall error when the occasional misclassification of a single pixel is not significant.

To efficiently utilize HI in classification applications, appropriate data processing techniques are needed to extract discriminative information from the contiguous and correlated HI bands. Processing of HI for classification usually starts with dimension reduction or a feature extraction step. Because a large number of bands are available for each pixel, feature reduction will diminish the amount of redundant information (Bioucas-Dias & Nascimento, 2008). Commonly used methods for feature reduction include principle component analysis (PCA) (Bateson & Curtiss, 1996), singular value decomposition (SVD) (Herries et al., 1996), maximum noise fraction (MNF) (Green et al., 1988), independent component analysis (ICA) (Lennon et al., 2001) and others. These methods seek to transform the original data given certain optimal conditions. For example, PCA achieves an optimal transformation in a least squares sense by searching

for significant eigenvectors of the data covariance matrix. Singular value decomposition use the eigenvectors of the positive definite matrix which is the product of the data matrix and its transpose to form basis vectors, and then uses these vectors to represent the data with a lower rank. The MNF transform is similar to PCA, but adds a noise-whitened process to PCA to ensure the "important" components picked by the algorithm have higher signal-to-noise ratio than the rest. Independent component analysis "consists in finding a linear decomposition of the observed data into statistically independent components" (Nascimento & Dias, 2005), by finding the mixing and separating matrix of the "independent sources". Dimension reduction is achieved by ICA by finding projection directions where maximum statistical independence exists and transforming the original data along these directions to a lower dimensional space.

After transformation of HI to a new space where features can be efficiently extracted, a broad variety of machine learning algorithms are used for feature learning and labeling of the HI pixels. Classification based on feature reduction or discriminant analysis algorithms can be developed with orthogonal subspace projection (OSP) (Harsanyi & Chang, 1994) and linear discriminant analysis (LDA) (Du & Chang, 2001); these methods are mostly employed for unsupervised classification. Other common unsupervised classification (clustering) algorithms that have been applied to HI include k-means (Funk et al., 2001), Iterative Self Organizing Data Analysis Technique (ISODATA) (Liew et al., 2002) and hierarchical clustering (Kumar et al., 2002).

Supervised classification of HI using a maximum likelihood classifier (MLC) on AVRIS data was performed as early as 1994 by Jia and Richards (1994), where MLC was applied to a dataset after unsupervised discrimination analysis to classify 7 land cover

types. However MLC requires a feature selection method to work efficiently and requires data to satisfy a certain distribution. To tackle the large number of spectral bands in HI and the often limited number of labeled training samples, kernel based methods have attained great popularity for HI classification (Camps-Valls & Bruzzone, 2005). Support vector machines (SVM) have been shown to be a very effective tool in HI classification (Gualtieri & Cromp, 1998; Melgani & Bruzzone, 2004), especially with the non-linear learning capability made available by a suitable kernel definition. In addition to spectral information, composite kernels can utilize both spatial and spectral information from HI (Camps-Valls et al., 2007). Other studies have used kernel learning in combination with other clustering methods, e.g., (Kwon & Nasrabadi, 2005; Baofeng et al., 2008; Ma et al., 2010).

Finally, ensemble learning uses many classifiers to deal with assigned tasks from the overall classification problem, and their decisions are weighted and combined to overcome the performance limitation of the original classifiers (Chan & Paelinckx, 2008). Popular ensemble learning algorithms include Adaboost (Freund & Schapire, 1996; Kawaguchi & Nishii, 2007) and random forest (Ham et al., 2005).

Recent research into HI classification has focused on spectral-spatial joint classification development (Huang & Zhang, 2013; Lunga et al., 2014; Q. Zhang et al., 2015), with the purpose of maximizing the utility of manifold data structure from HI.

3.3 Complementary Characteristics of FWL and HI

The latest HI classification research attempts to exploit both the spatial and spectral information for HI for optimal discrimination. Limited by the nature of HI,

image information only provides a two-dimensional description of the spatial neighborhood. On the other hand, FWL offers detailed three-dimensional spatial information, therefore allowing geometric properties to be extracted which are not limited to the imaged surface. For example, in a forested area, FWL records canopy cover, interlayer canopy, shape and height of vegetation, and ground topography at the same time (Asner et al., 2007), while HI can only detect canopy cover spectral information.

HI captures object surface reflectance characteristics as a sampled spectrum, and FWL records the object backscattered energy as intensity values associated with each sample. Surface reflectance is a very good discriminator of the composition of an object, and the backscattering properties of the object can reveal both material information (although not as discriminatory as a HI spectrum) and surface roughness. Furthermore, although the apparent backscattering properties contained in FWL do reflect material composition, they are also significantly affected by the acquisition geometry (e.g. incidence angle) when the waveform was recorded. Whereas for HI, although the absolute scale of HI spectrum is also affected by the position of the sensor and imaged surface, a contiguous sampling of a broad spectrum range is available, which can normalize the scale differences by studying the relationship between different bands.

Finally, the spatial resolution of airborne HI is generally at the meter level, whereas airborne FWL can have more than ten waveforms penetrating the voxel defined by the same size footprint (i.e., per square meter). With such dense spatial sampling, adding FWL to HI classification is expected to improve classification accuracy and potentially reduce the required number of ground training samples.

Chapter 4 Remote Sensing Data Classification

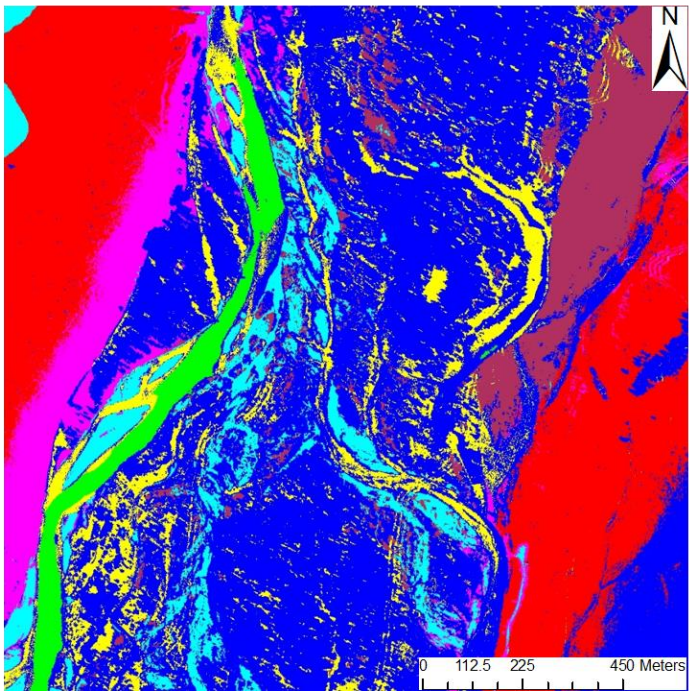
4.1 Background

Land cover and land use information is crucial for decision making in global change studies and environmental applications (Sellers et al., 1995). Remote sensing data is an important source for providing such information, because of the capability to provide measurements of a broad range of object characteristics at a large scale and fine resolution. The term classification refers to the extraction of the desired information from the remote sensing data. The classification problem can be defined as "given a set of training data points along with associated training labels, determine the class label for an unlabeled test instance" (Aggarwal, 2015). Remote sensing data classification techniques group pixels (or the minimal data size unit for non-rasterized data) with common properties to represent a variety of land cover features. Land cover feature examples include forested, urban, or agricultural areas. An example of remote sensing data classification is illustrated in Figure 4-1, where HI of a 1200m x 1200m area is classified into seven land cover classes with training data selected through manual interpretation of the HI.

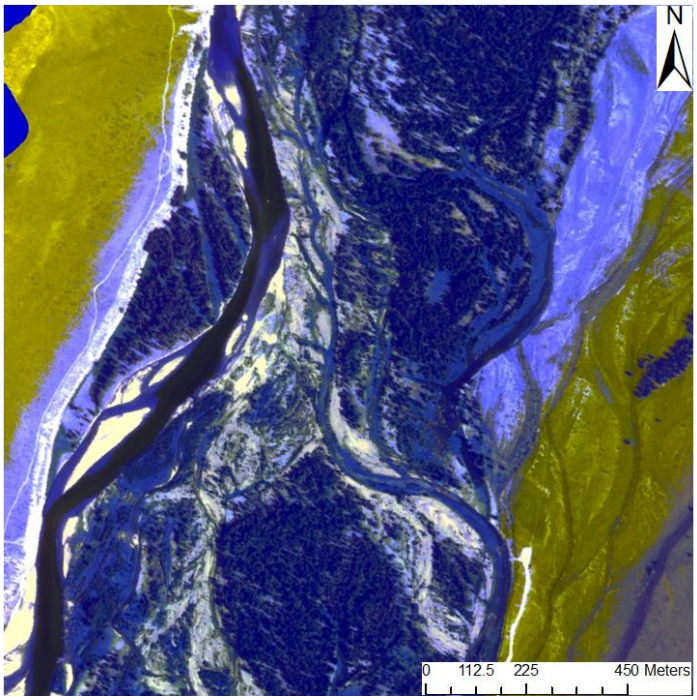
Applications of classification techniques for distinguishing land cover types developed rapidly with the increasing availability of remote sensing data sets beginning in 1970's with the satellite-borne multispectral data from the Earth Resources Technology Satellite (Sinnock et al., 1974; Ulaby & McNaughton, 1975). More modern sensors include the Landsat multispectral scanner and thematic mapper system (Shlien & Smith, 1975; Nelson & Hoffer, 1981; Cibula & Nyquist, 1987; Ormsby, 1982) and,

SPOT (Satellite for observation of Earth) data (Shimoda et al., 1988; Franklin & Peddle, 1990),both of which are used for automatic large scale classification of land in both forestry and hydrology applications. Synthetic aperture radar (SAR) has also been employed for classification since the launch of European Remote Sensing (ERS) satellites. Unlike the spectral information collected by Landsat or SPOT, SAR is an active microwave based remote sensing technique which measures the energy backscattered from illuminated targets. Because of this property, SAR can be used to infer terrain attributes and vegetation structure information, and has proven to be effective in distinguishing woody vegetation with different trunk characteristics (Craig et al., 1995). Airborne hyperspectral data merged as a data source for classification with the development of the airborne visible/infrared imaging spectrometer (AVIRIS) (Roger, 1996; Harsanyi & Chang, 1994). The rich spectral information contained in HI provides a potential solid observation basis for separating a large number of ground cover classes in complex scenes. The potential use of LiDAR data for classification was first examined in (Jensen et al., 1987), where the LiDAR data was registered to a classification map generated from airborne multispectral imagery, and the class assignments showed strong correlation with a LiDAR derived height model. Early classification of LiDAR data used intensity and height information and was mainly applied to applications where the number of land cover classes was quite limited (Weed et al., 2002). Waveform LiDAR has also been applied to land cover classification, with a variety of structural features extracted mainly for forestry applications (Neuenschwander et al., 2009).

(a)



(b)



(c)

Figure 4-1 Remote sensing imagery and classified land cover map: (a) RGB band from HI; (b) Classification map; (c) Classification result with texture from HI band for better visualization

4.2 Land Cover Classification Algorithms

Classification algorithm can be broadly divided into two categories: supervised and unsupervised. For supervised classification, the patterns of data class dependence on the features hidden in the training samples are revealed and represented using a classifier model (Kotsiantis, 2007). The model is then applied to assign class labels to test samples. This process is conceptually illustrated in Figure 4-2.Unsupervised classification refers to methods employed when no training samples with class labels are available and the data can only be clustered by similarity, and the class labels are assigned by the user based on knowledge regarding land cover features in the scene. Unsupervised classification is automatic and usually can be achieved with simpler algorithms compared to supervised classification, but the accuracy of classification is difficult to evaluate and requires additional interpretation by the user.

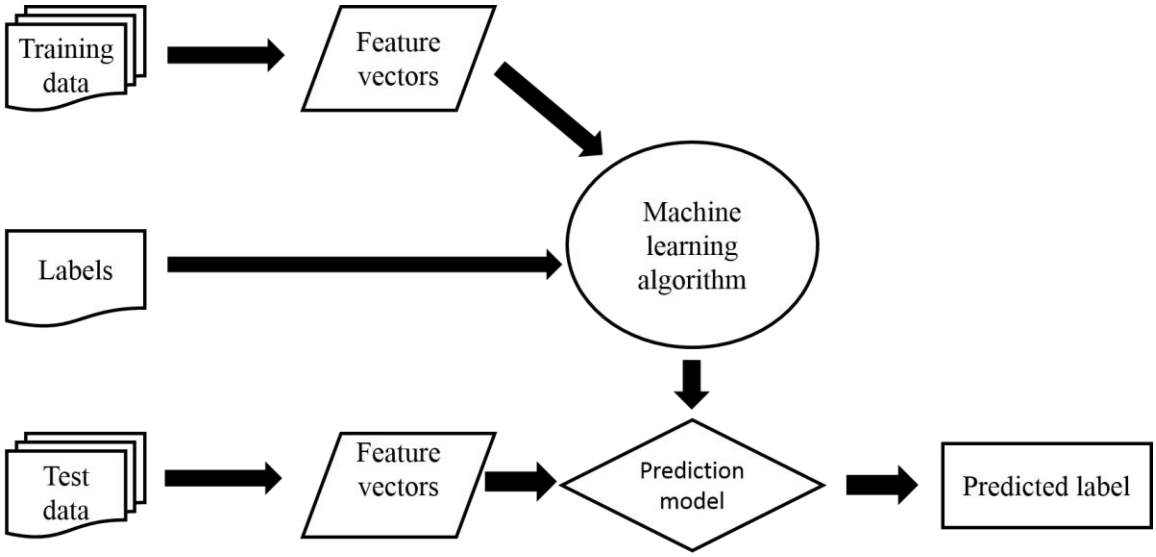


Figure 4-2 Supervised classification flowchart

4.2.1 Unsupervised Classification Algorithms

The basic premise of unsupervised classification is the selection of a set of seed points around which clusters (groups of similar pixels) are built, where the membership of each cluster can be changed to achieve an optimal partition. A number of representative unsupervised classification techniques are briefly described in the following, but the discussion is by no means exhaustive.

Iterative Self-Organizing Data Analysis Technique (ISODATA) is a popular method for unsupervised classification for remote sensing data, and was originally proposed for use in multispectral data classification (Gowda, 1984). The method is based on the K-means algorithm (MacQueen, 1967), K-means and ISODATA methods both use spectral distance as a similarity measure for assigning candidate pixels to clusters. Working in an iterative manner, ISODATA estimates the center of each cluster by moving each pixel from one cluster to another to reduce the quadratic spectral distance.

This procedure is applied to every candidate sample and the cluster centers are recalculated after assignment of each new sample.

Neural networks are also used for unsupervised classification when the algorithms for individual "neurons" work in an unsupervised manner (Hara et al., 1994). The strength of neural networks is their ability to determine an appropriate similarity metric in the training process, and adjust the weighting of the entire system according to data reliability. Neural networks have been applied to both SAR data (Hara et al., 1994), multispectral data (Baraldi & Parmiggiani, 1995) and HI (Goel et al., 2003). Markov random fields (MRF) have been used for unsupervised parametric classification with the help of a K-means classifier for an initial coarse classification (Yamazaki & Gingras, 1999). Discrimination analysis tools such as linear constrained distance-based discriminant analysis (LCDA) can also employ K-means or ISODATA as an initial coarse classifier (Du & Chang, 2001).

Recent developments in unsupervised classification are mostly improvements to the methods above. The introduction of fuzzy sets theory to clustering techniques is an important development and is widely used (Tran et al., 2005). Fuzzy clustering methods change the method of determining cluster variance and cluster size to a "soft' deterministic manner (Duda & Canty, 2002). The strength of fuzzy clustering methods is that data samples located in an area of overlapping clusters have little impact on the cluster parameters, which mitigates the influence of outliers, noise or points with high uncertainty.

4.2.2 Supervised Classification Algorithms

In contrast to unsupervised methods, supervised classification algorithms require training data to adjust algorithm parameters in order to achieve efficient partitioning of the test feature space. As with the prior discussion of unsupervised classification methods, we briefly review several representative supervised classification methods in this section.

The maximum likelihood classifier (MLC) is a supervised classification method that has been proven effective for a variety of remote sensing data sets, such as multispectral data (Strahler, 1980), HI (Jia & Richards, 1999), and SAR images (Kuplich et al., 2000). Another extensively used supervised classification method is the support vector machine (SVM) technique. According to (Huang, Davis, and Townshend 2002), SVM separates different classes by searching for optimal boundaries using optimization algorithms . Support vector machines are capable of adapting different learning functions as a "kernel" to meet the requirements of the training data (Melgani & Bruzzone, 2004). Both MLC and SVM are used in this dissertation and further details of these two classification methods will be given in Chapter 5.

Besides MLC and SVM, a broad variety of machine learning techniques have been introduced to solve classification problems in a supervised manner. Decision tree classification iteratively divides the data set into smaller partitions using a set of tests defined at each branch of the tree. Decision trees are nonparametric and thus do not place assumptions on the distribution of data (Friedl & Brodley, 1997). This allows decision trees to learn nonlinear pattern between features and class labels, as well as employ mixing type of features (Pal & Mather, 2003). Ensemble learning is a family of methods that train a number of simple classifiers such as decision trees, and combine the decisions

through a weighted voting process. Adaptive boosting (Freund & Schapire, 1996) and random forests (Gislason et al., 2006) are important implementations of ensemble learning. Adaptive boosting is a boosting ensemble learning method which calculates the output using several different models and then combines the results using a weighted average approach. Random forest is a bagging ensemble learning method which decreases the prediction variance by incorporating random feature set selection, so a bootstrap sample of observations is available for each tree (Ham et al., 2005).

4.2.3 Accuracy Assessment of Land Cover Classification

The validity of classification results must be evaluated by an accuracy assessment process. In the context of using machine learning methods for supervised classification, quantitative analysis can be performed by comparing the predicted labels with the actual class label of reference samples (Foody, 2002). The confusion matrix is the most extensively used measure of classification accuracy in the remote sensing field. A confusion matrix is a cross-tabulation of the predicted label by the classifier against the reference data (Canters, 1997), that makes it possible to infer whether additional discrimination information is needed to separate correlated classes.

Several measures are typically generated from a confusion matrix to assist in understanding overall system performance and better interpret the confusion matrix (Foody, 2002). An overall accuracy can be computed which quantifies the percentage of all correctly labeled instances. When the accuracy of individual classes is the focus, there are two ways methods of evaluation: producer's accuracy and user's accuracy (Congalton, 1991). Producer's accuracy is the fraction of correctly classified samples with regard to all samples of that class in the reference data. User's accuracy is the fraction of correctly

classified samples with regard to the number of all samples assigned to the particular class by the classifier, and is therefore referred to as the reliability of the classification.

Several measures derived from combinations of the producer's accuracy and user's accuracy have also been proposed with the goal of evaluating the classification performance from the view of both producer's and user's accuracies (Liu et al., 2007). Conditional kappa accuracy, or Cohen's kappa coefficient, was originally introduced to quantify the agreement between two raters (Cohen, 1960), and is computed as the ratio of the actual relative agreement percentage against the probability of random agreement of two raters. In supervised land cover classification, the agreement between pre-assigned reference labels and the classification results of the reference data is used for evaluating kappa accuracy (Smits et al., 1999).

4.3 Classification of Fused FWL and HI

Early works addressing the issue of fusing FWL with passive remote sensing techniques have predominantly been limited to using FWL data that has been converted to a DSM and canopy parameter raster model for tree type recognition applications (Anderson et al., 2008; Asner et al., 2007; Paris & Bruzzone, 2015). The FWL data can also be processed in a similar way for HI feature extraction (Wu and Prasad 2013; M. J. D. Sarrazin et al. 2011), where feature extraction has been done by performing feature reduction techniques on raw FWL data, first and second derivatives, then the reduced feature sets are merged with HI feature sets. Fusion of FWL and HI can also be achieved by first decomposing the LiDAR waveforms and then registering the echo functions to HI, based on the process shown in Jung, 2011. The method of utilizing the whole waveform without fitting a model function has also been investigated in (Buddenbaum et al., 2013),

where the waveforms were reconstructed and processed at a coarse spatial level (50cm voxel height), and were only used for separating pre-selected trees-only training samples, without evaluating improvements to overall fused land cover classification.

To summarize, classification of combined FWL/HI datasets have been mainly limited to certain types of targets (e.g., tree type classification), or land cover classification using similar procedures for feature generation from FWL and HI without thoroughly considering the characteristics of the different data sources. Furthermore, 3D information contained in FWL is often compromised in order to achieve co-registration with HI.

Chapter 5 Proposed Methods

In this chapter, the proposed method for fusing and classifying features extracted from FWL and HI is presented. This includes georeferencing and voxelization of the FWL data, feature extraction and combination of the FWL and HI, and the classification scheme for the fused FWL/HI data set. Figure 5-1 contains a flowchart of the entire process.

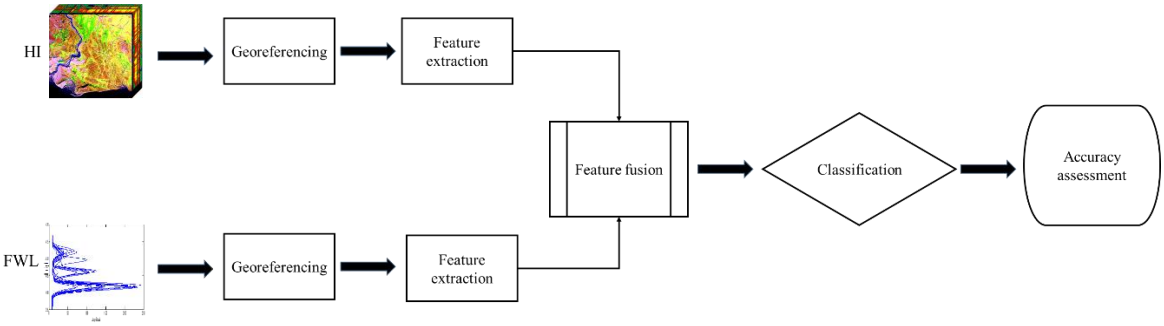


Figure 5-1 Flowchart of FWL/HI processing, fusion, and classification

5.1 Waveform Data Processing

5.1.1 Direct Waveform Georeferencing

As discussed in Chapter 1, it would be beneficial for applications using FWL data to fix the location of each complete raw waveform in 3-D space and use the georeferenced waveform data to generate a vertical superposition response for targets within each raster cell, in order to have a one-to-one correspondence between HI pixels and FWL records. To achieve this, we adapt the existing georeferencing method for LiDAR discrete returns and apply them to waveform positioning.

FWL returns are usually sampled at even time intervals ranging from 1 to 2 nanoseconds, yielding a range resolution between waveform samples of 0.15 to 0.30 m.

Combined with the aircraft trajectory, the waveform sample ranges are used to locate the waveforms as 3D arrays in a global reference frame. First, the equivalent slant range from the sensor to each sample in the return waveform is computed from the time elapsed since laser pulse emission,

$$R = \frac{c}{2} * \left(interval + T_{peak} - T_{peak}^{0} + n \cdot \Delta T \right), \tag{5-1}$$

where R is the slant range, c is the speed of light, *interval* is the time from the first sample on the outgoing pulse to the first sample on the return pulse, T_{peak} is the peak time of the return pulse, T_{peak}^{0} is the peak time of the outgoing pulse (trigger start time), n is the sample location on the return waveform, and ΔT is the timing interval of the waveform digitizer.

With the slant ranges calculated, each waveform sample can be georeferenced via a coordinate system transformation using data from the ALS POS (position and orientation) system. The transformed 3-D coordinates are given as (Glennie, 2007)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{G}^{l} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}_{GPS}^{l} + R_b^l(\omega\varphi\kappa) \cdot \left(R_s^b(d\omega d\varphi d\kappa) \cdot r^s(\alpha R) - \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix}^{b} \right), \tag{5-2}$$

where the output vector components X, Y and Z are the coordinates of a single waveform sample in a global coordinate frame, and X_0, Y_0, and Z_0 are the coordinates at the origin of the navigation sensor (normally the INS center). The angles ω,φ,κ are the roll, pitch and yaw of the sensor with respect to the local level frame (as determined by the INS). The angles $d\omega,d\varphi,d\kappa$ are the boresight angles that correct for the difference in alignment between the laser scanner and the INS measurement frame determined by boresight

calibration (Skaloud & Lichti, 2006). The lever arm offsets l_x, l_y, and l_z are the physical offsets between the INS origin and the measurement origin of the laser scanner assembly. In this equation, the sensor measurement vector $r^s(\alpha, R)$ is given as

$$r^s(\alpha\ R) = \begin{bmatrix} \sin(\alpha)\cdot R \\ 0 \\ -\cos(\alpha)\cdot R \end{bmatrix} = \begin{bmatrix} \sin(\alpha)\cdot\frac{c}{2}*\left(\text{interval}+T_{peak}-T_{peak}^{0}+n\cdot\Delta T\right) \\ 0 \\ -\cos(\alpha)\cdot\frac{c}{2}*\left(\text{interval}+T_{peak}-T_{peak}^{0}+n\cdot\Delta T\right) \end{bmatrix}, \quad (5\text{-}3)$$

where the scan angle α and range R are measured and returned by the laser scanner. According to equation 5-2 and 5-3, the georeferencing accuracy of each waveform sample is affected by INS altitude errors, boresight errors, laser scanner errors, lever-arm offset errors, and GPS positioning errors, as well as the data acquisition parameters (flight height, scan angle range); detailed discussions of error analysis for kinematic laser scanning can be found in (Schaer et al., 2007; Skaloud & Lichti, 2006; Glennie, 2007).

Using the formulas detailed above each FWL sample can be georeferenced and transformed from the time domain to the spatial domain. A georeferenced waveform can be viewed as a 3D array representing the laser pulse energy distribution along the laser travel path, where the energy distribution is represented by the amplitude of each sample. Note that the irregular spatial distribution of georeferenced FWL returns needs to be accounted for. In order to co-register the FWL with raster based remote sensing data. The spatially irregular georeferenced FWL samples need to be located in regular intervals aligned with the raster grid; this is accomplished via a 3-D voxel-based reference frame.

5.1.2 Waveform Voxelization and Synthesis

5.1.2.1 Waveform Voxelization

To build a 3-D voxel frame, the area illuminated by the laser scanner is first divided into a regular 2-D grid. The 2-D grid is then vertically extended into 3-D at regular vertical intervals, thereby forming vertical columns of voxels, where each voxel can be thought of as a rectangular prism. The georeferenced waveforms are then intersected with the columns to determine the waveform signature within each column. Note that waveforms can be divided and assigned to multiple columns. The process is illustrated in Figure 5-2.

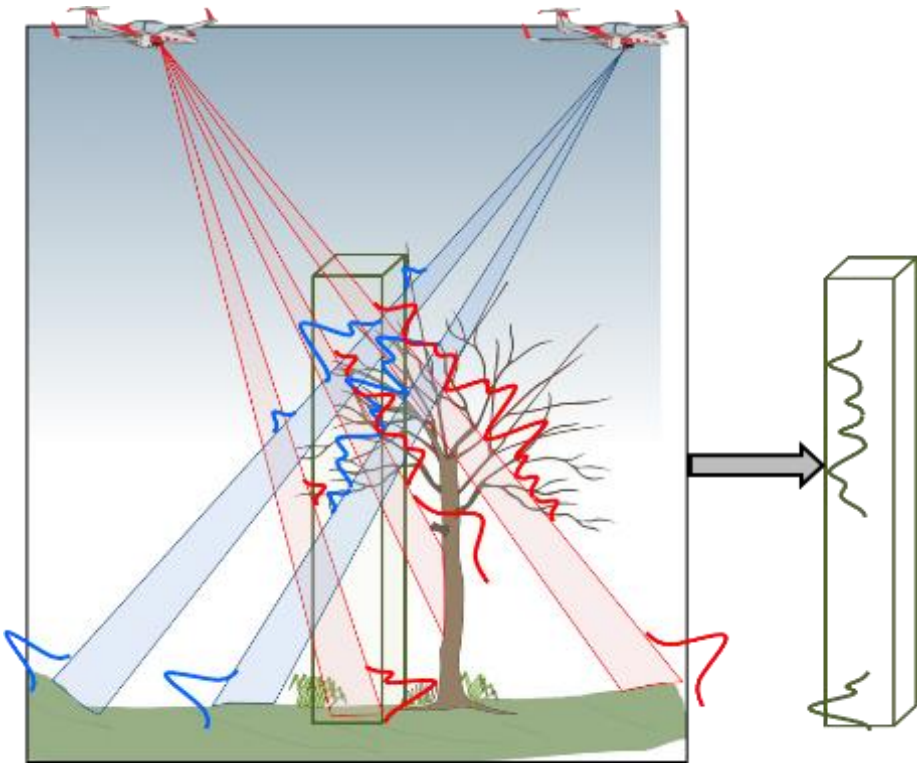


Figure 5-2 Waveform stacking for vertical FWL return profile

All of the waveform samples that intersect a voxel column of interest can thus be indexed with the column signature, and combined using a waveform synthesizing

algorithm, discussed in the following section, to generate a single representative waveform for the column. Note that with this methodology all samples in a waveform are utilized instead of only retaining the locations of return echoes detected by Gaussian decomposition or other deconvolution techniques. Retaining the entire waveform is important for the following reasons: 1) return echo shapes may deviate from the assumed model function when the target surface is not flat, 2) waveform segments may contain return energy that is below the threshold for echo detection, and 3) superposition of waveform amplitudes using information from multiple scans may accentuate small echoes.

It is essential to define an appropriate voxel size that takes into consideration the pixel size of the HI imagery (horizontal resolution) and also matches the sampling interval of the LiDAR waveform digitizer (vertical resolution). In this dissertation, the size of each voxel was chosen to match the size of the HI pixel and the raw waveform digitizing interval.

5.1.2.2 Weighted Summation Methods for Waveform Synthesizing

The optimal relationship when fusing LiDAR waveforms with HI pixels is a 1:1 relationship, i.e., each pixel corresponds with one waveform. In order to achieve this, a method to synthesize all waveform profiles that intersect a single column of voxels is required.

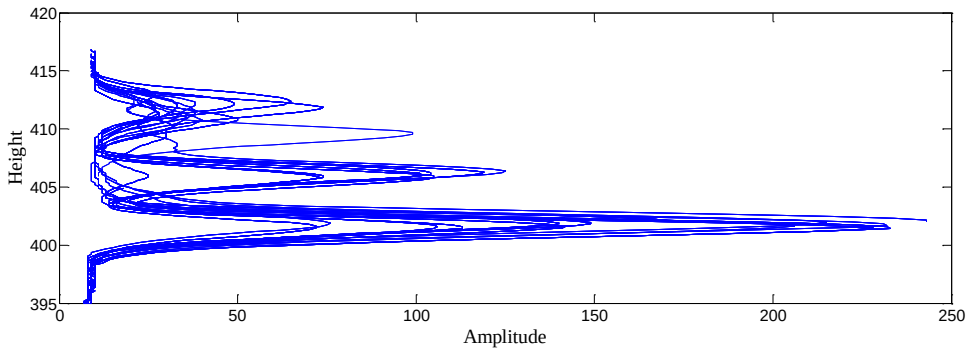
Several numerical waveform synthesizing methods were tested in our early work (Wang et al., 2013). We begin by examining empirical weighted interpolation methods.

For each voxel containing waveform samples, the interpolated intensity can be represented as

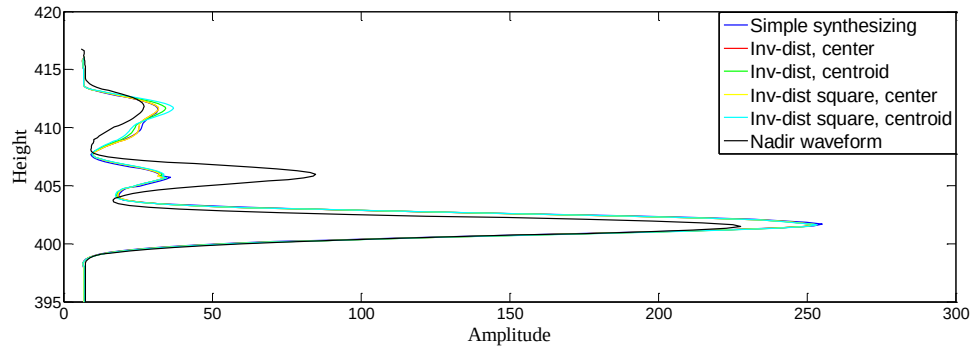
$$I_{Voxel} = \sum_{i=1}^{n} \omega_i I_i \Big/ \sum_{i=1}^{n} \omega_i \,, \tag{5-4}$$

where n is the number of waveform samples in a particular voxel, I_i is the intensity of each waveform sample in the voxel, ω_i is the weighting coefficient for each sample, and I_{voxel} is the weighted intensity which is retained for the voxel. In this way, a single response value is obtained for a single voxel no matter how many waveforms penetrate the voxel and how many waveform samples are allocated.

Four different weighing methods were evaluated, including inverse distance from the geometrical center, inverse squared distance from the geometrical center, inverse distance from the centroid, and inverse squared distance from the centroid of the voxel. An example comparing the weighted interpolation methods for synthesizing waveforms is given in Figure 5-3. The waveforms were obtained over a column of voxels containing high vegetation. An original waveform having a look angle close to nadir was chosen as a reference for the reconstructed waveforms. It can be observed that the positions and shapes of the interpolated waveforms are close to the nadir waveform, despite some amplitude deviations. Furthermore, the different weighting methods do not affect the synthesized waveform shape significantly, which is because the waveform samples in a voxel tend to have a compact spatial distribution.



(a)



(b)

Figure 5-3 Waveform synthesis example: (a) All raw waveforms intersecting a voxel of interest, (b) reconstructed waveform using four different weighted interpolation methods

To quantitatively evaluate the shape deviation between the synthesized waveforms and reference waveform (close to nadir original waveform), a waveform modeling and decomposition approach was used to describe the waveform shape using parameters of decomposed echo functions. Both the original and synthesized waveforms were decomposed into individual echo functions, and the parameters of the functions were compared. The experiment was carried on a data set of mixed urban and forested area. Since the system had a Gaussian-shaped transmitted pulse (Wang et al., 2013), both waveforms were decomposed into Gaussian echo functions. The parameters used for similarity comparison were the position, amplitude and width for the Gaussian

components, with the similarity quantified using a Canberra metric(Lance & Williams, 1966),

$$p = 1 - \frac{1}{N}\left[\sum_{i=1}^{N}\left(\frac{\left|Padameter_{truth} - Parameter_{synthesis}\right|}{Padameter_{truth} + Parameter_{synthesis}}\right)\right], \quad (5\text{-}5)$$

where N represents the number of voxels, and p is the similarity measurement. Experimental results are shown in Table 5-1; inverse square distance to center method achieved the highest agreement for all three parameters among all synthesizing methods.

Table 5-1 Parameter agreement between reference and synthesized waveforms for all methods

Fusion method/ Consistency	*Inverse distance to center*	*Inverse square distance to center*	*Inverse distance to centroid*	*Inverse square distance to centroid*
Echo position	99.35%	99.46%	99.46%	99.46%
Amplitude	85.05%	90.05%	82.41%	84.76%
Width	72.27%	75.56%	72.41%	74.54%

5.1.2.3 Maximum Amplitude Method for Waveform Synthesizing

Although the waveform synthesizing interpolation methods are able to retain the waveform shape to a certain level, in the case where a return echo (potential target) is only captured by one waveform, the amplitude will be significantly reduced after the interpolation. However, we are trying to preserve all the possible waveform features. A further disadvantage is the requirement for interpolation within each voxel, which is computationally heavy given that each vertical waveform is generated from hundreds of voxels and numerous individual waveform returns.

Considering the drawbacks of the weighted interpolation methods, we have instead chosen a method that retains the maximum amplitude in each individual voxel (Hermosilla et al., 2014) to derive a synthesized waveform (SWF), which can be expressed as

$$I_{Voxel} = \max\left(\left[I_i\right]_{i=1,2,\ldots,n}\right). \tag{5-6}$$

The advantage of this approach is that the maximum amplitude contained in each voxel is likely to be a better estimate of the response of the target when hit by the laser pulse from the direction that maximizes the backscattering intersection, without being compromised by energy loss caused by reflections from other targets in the laser cone of diffraction. This approach also has the benefit of being computationally efficient and produces a smooth voxel waveform which provides superior fused classification accuracy; results are shown in Chapter 7.

5.1.3 Vertical Energy Distribution Coefficient

The goal of voxelizing and synthesizing the original waveforms is to preserve location and intensity information from the raw data without fitting the original waveforms to assumed model functions, e.g., Gaussian decomposition. Decomposition methods applied to the SWF for feature extraction are also not desirable given the increased complexity in the SWF generated from multiple overlapping scattering surfaces. Under these conditions the efficacy of using Gaussian parameter features will be impacted by the divergence between the SWF and fitted Gaussian models due to the increased SWF complexity.

Therefore, in order to discriminate SWF for different land cover classes, features describing the distribution of return energy along each waveform profile, i.e., metrics for both intensity and height information contained in the SWF, are desired. We propose to use vertical energy distribution coefficients (VEDC) as features to encapsulate the integrated intensity-height metrics of the SWF. The SWF of a column of voxels represents the reconstructed backscattered energy profile of targets in the vertical direction. By examining the ratio of return energy contained in defined segments of the SWF to the total SWF energy, an estimate of the energy distribution can be achieved. The VEDC metric is computed as follows:

1. Each SWF is segmented according to the height distribution of samples using a constant interval. Details regarding the interval selection are given in Chapter 7.

2. The backscattered energy contained in each SWF segment is estimated by computing the area under the SWF curve.

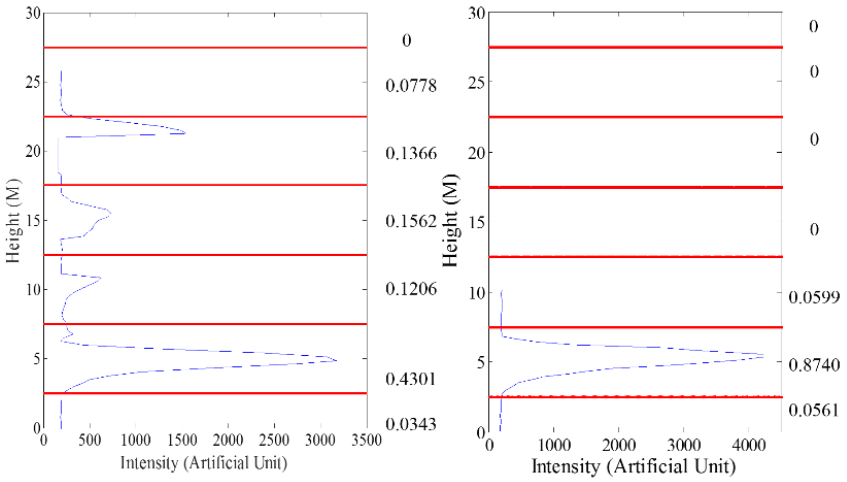
3. The ratio of return energy for each segment with respect to the total energy of the SWF is computed, i.e., the VEDC is obtained as

$$VEDC_i = \frac{\int_{h_i}^{h_{i+1}} Intensity(h)dh}{\int Intensity(h)dh}, \tag{5-7}$$

where h_i represents the starting height of the i th interval of VEDC, and h_{i+1} the starting height of the next VEDC interval. The function *Intensity (h*) is the intensity of the entire SWF.

An example of VEDC for both a tree and ground SWF is shown in Figure 5-4. The tree SWF penetrates a thin canopy and also contains a ground return even though a

significant portion of the outgoing energy was reflected by the canopy. On the other hand, the energy of the ground SWF is mostly concentrated near the ground return (last return), which contains 87.40% of the total energy. The VEDC of these two SWF clearly show the differences between these two classes.



(a) (b)

Figure 5-4 Example of VEDC of waveforms from different classes: (a) tree, (b) ground.

5.1.4 VEDC Feature Generation

VEDC is a data-dependent feature of the SWF because knowledge of the elevation distribution of the SWF samples is necessary to determine both the number of VEDC segments and the span of each segment. The height range of the SWF in the targeted area is determined by examining the SWF of all voxels. Once the height range for the VEDC is fixed, only the number of VEDC components is required, since the height interval of each component is uniform.

5.1.4.1 Supervised VEDC Feature Generation

The dimension of VEDC can be determined using a supervised method: a portion of the SWF training samples can be classified using VEDC with varying dimensionality.

The classification accuracy can then be used as a metric to estimate the optimal VEDC dimension with respect to the inter-class discrimination provided by VEDC. Since VEDC does not require a large number of segments, i.e., number of features, this supervised VEDC generation method does not add significant computational burden.

5.1.4.2 Unsupervised VEDC Feature Generation

Supervised feature generation requires an appropriate classification technique for the accuracy evaluation. However, different classifiers may yield different optimal dimensions for VEDC. Therefore, an unsupervised feature generation method is desirable.

According to equation 5-6, VEDC represents the ratio of the energy contained in each segment with respect to the total energy of the SWF. Therefore the sum of all components is always equal to 1, with individual component values determined by the segmentation. Thus, if we assume the probability that a single return in a waveform is reflected from a target located in the height interval of $[h_i, h_{i+1}]$ is P_i, then we have $P_i = VEDC_i$, the *i*th component of the waveform VEDC. If the dimension of VEDC is fixed at N, then based on the definition of VEDC, it is obvious that

$$\sum_{i=1}^{N} P_i = 1. \tag{5-8}$$

Based on these assumptions, the Shannon entropy of the VEDC feature can be computed to evaluate the unpredictability of the feature, serving as an efficacy metric for feature generation. Because VEDC is generated for classification, it is preferable that waveforms from different land cover classes be used for the entropy evaluation. Suppose the number of classes to be labeled is L and number of samples in class j is M_j; then the entropy of the *i*th component of the VEDC is

$$E_i = \sum_{j=1}^{L} \sum_{m=1}^{M_j} (VEDC_i)_{jm} log_n [(VEDC_i)_{jm}]. \quad (5\text{-}9)$$

The entropy value of each VEDC component can be summed to measure the total information content preserved by the segmentation of VEDC and, therefore, can be used to adjust the VEDC dimension. Furthermore, the entropy value of each VEDC component can be used to evaluate the importance of each component. Height intervals with significantly low entropy values can be combined with adjacent intervals to reduce the redundancy in the VEDC feature set. An illustration of the relationship between the entropy value and VEDC feature efficiency is given in Figure 5-5. With increasing entropy, the information contained in the generated VEDC features are more efficient for data separation, i.e. the samples from two classes show better in-class convergence and between class discrimination with a larger entropy value (8 band VEDC). This demonstrates that entropy can be used as a measure of VEDC feature information content. Details of using the unsupervised VEDC feature generation algorithm on actual datasets are given in Chapter 7.

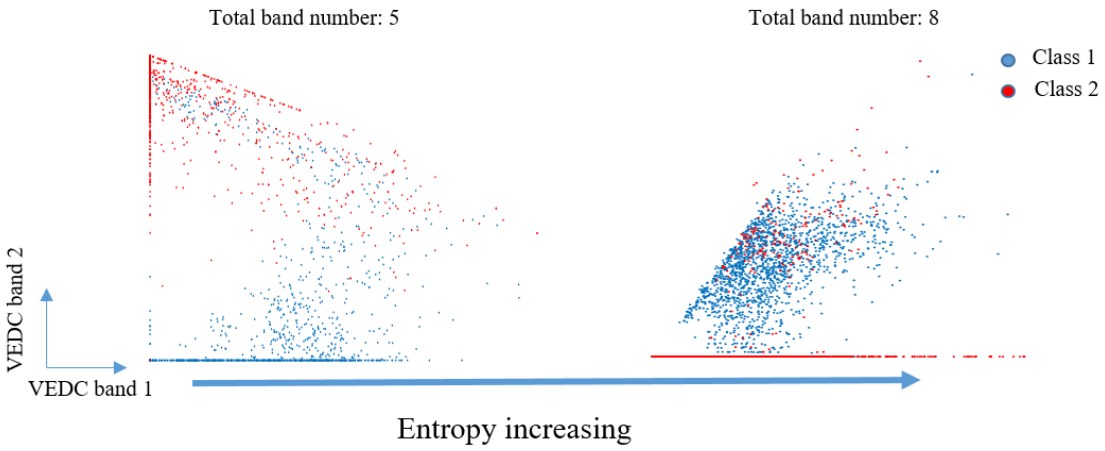


Figure 5-5 Entropy and VEDC information content

5.2 Feature Extraction for HI

Feature extraction techniques for HI and HI classification methods are well developed, and the methodology was reviewed in Chapter 3. Principal components analysis (PCA) has long served as an efficient dimension reduction and feature extraction tool for better classification performance (Smith et al., 1985). Principal component analysis is by definition a particular kind of orthogonal linear transformation. Since HI bands are close in frequency and sensor spectral resolution is not infinitely fine, correlation may exist between HI bands. The PCA orthogonally transforms HI bands into a new space where the correlation among transformed bands is minimized. Furthermore, PCA preserves data variance within a smaller number of bands, making class discrimination in PCA transformed data easier to utilize.

Covariance analysis of the input data is the first step when performing PCA. If we consider the input data, i.e., an HI data cube, as an $m \times n$ matrix **X**, where column number n denotes the number of pixels (the 2D imagery is reshaped as a 1D array because spatial information is not considered) and row number m represents the number of spectral bands, the PCA method studies the pattern of the data matrix **X** by calculating its covariance matrix

$$S = \sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T, \tag{5-10}$$

where $\bar{x}$ denotes the mean spectrum vector computed from **X**, and x_i is the spectrum vector of each pixel. Once the covariance matrix is obtained, the eigenvectors are determined by finding vectors λ satisfying the condition

$$|S - \lambda I| = 0, \tag{5-11}$$

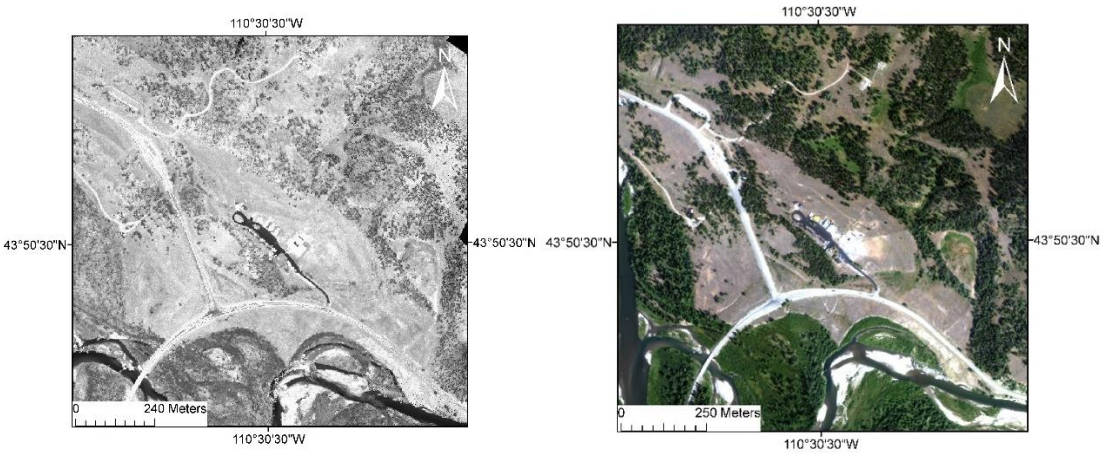
where I is the identity matrix with dimension m. These mutually orthogonal eigenvectors represent patterns within the input data. The eigenvalues associated with the vectors represent the corresponding variance; therefore, eigenvectors with larger eigenvalues are considered the principal components of the data. By ignoring eigenvectors with smaller eigenvalues, the dimension of the data can be reduced with the majority of HI variance information preserved. Typically a very high percentage of variance of the input data is preserved with a relative small number of PCA bands. The final data after PCA transformation can be expressed as (Smith, 2002)

$$FinalData = FeatureVectorMatrix \times AdjustedData, \tag{5-11}$$

where *FeatureVectorMatrix* is the transformation matrix composed of the selected principal components, and *AdjustedData* is the original data with the mean subtracted and transposed. After PCA, each vector in the final data is a linear combination of the HI bands in the original data.

5.3 Fusion of FWL and HI

If we build the voxels for FWL using the HI pixel locations, then each SWF is aligned with a corresponding HI pixel at its spatial center. In this research, the HI and FWL data were acquired using the same onboard POS; therefore, the HI and FWL data are already located in the same 3-D coordinate frame after georeferencing. An example of the registered SWF and HI data is shown in Figure 5-6. The intensity image is generated by assigning the value of maximum intensity from the SWF to each voxel to a pixel. Once the SWF data and HI are registered, an integrated feature space can be generated by stacking the HI and SWF features.



(a) (b)

Figure 5-6 Registered FWL and HI: (a) Intensity image of FWL; (b) HI of the same area.

5.4 Classification Method

5.4.1 Pairwise Classifiers

A stacked feature set comprised of both HI and SWF features offers information on spectral reflectance as well as target backscattering geometric properties and is therefore expected to enhance classification performance compared to using each dataset independently. An important issue to be considered in the process of designing a classifier for FWL/HI data is that the performance of HI and SWF features for distinguishing different classes can vary significantly. For instance, it is hard to separate asphalt covered roads and concrete covered roads using only SWF, while HI has difficulty correctly classifying ground cover that is shadowed by a high vegetation canopy.

With these considerations, an optimal classification technique for the fused features of HI and SWF should have the ability to choose subsets of features or even different classifiers for different target classes. The proposed solution is based on the framework of pairwise classifiers. Pairwise coupling for classification is an alternative to

the problem of multiclass polychotomous labeling (Friedman, 1996; Hastie & Tibshirani, 1998). A pairwise classifier decomposes a multiclass discrimination problem into a series of two-class classification problems to simplify the process.

The pairwise classifier framework has been utilized for the classification of remote sensing data such as Synthetic Aperture Radar (SAR) (Crawford, 1999), and HI (Kumar et al., 2001), and has been shown to improve classification given a relatively large number of input classes. There has also been some preliminary work using a Bayesian based pairwise classifier for FWL (Neuenschwander et al., 2009). Although a pairwise classifier shows potential advantages for classification of combined datasets from different sources, this method has been primarily limited to classification of a single data source or similar data sources (e.g., Polarimetric and Interferometric SAR), and has not been considered for classification of spatially and spectrally different modalities such as HI and FWL together.

The basic concept of pairwise classification is shown in Figure 5-7. First, if we assume the number of classes is *N*, the classification problem can be divided into C_N^2 two-class problems (Friedman, 1996). For each pair of two-class problems, a different classifier can be trained to give optimized discrimination between the two classes. Since classification problems are usually solved by optimizing the value of a mis-classification risk function, two-class classification can be achieved with target function in a significantly simpler form. After each individual classifier is trained, using training features from the two classes, test data is fed to each classifier to get an estimate of likelihood or a vote for the class label. The final label of a test feature vector is assigned by applying appropriate decision rules based on the votes from all individual classifiers.

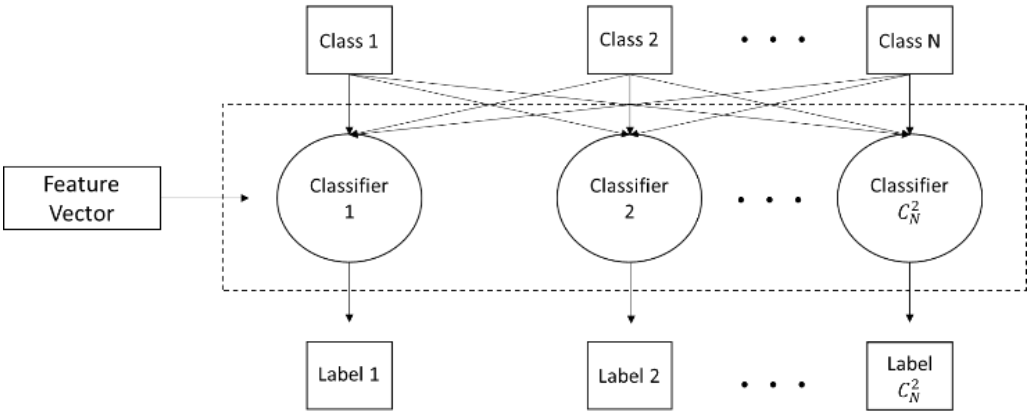


Figure 5-7 Architecture of a pairwise classifier

Commonly used decision rules used for assigning votes from all individual classifiers are maximum vote number and maximum posterior probability rules. The maximum vote number decision rule looks for the class label with the highest frequency in the output of all individual classifiers for one input feature vector, and assigns that label as the final class. The maximum posterior probability rule considers the posterior probability output by the classifier, and assigns the label with maximum probability value as the final class.

Pairwise classification simplifies the multiple classes labeling problem, and more importantly provides the flexibility to use different types of classifiers and subsets of feature spaces for each class pair. Such flexibility can be quite helpful for SWF/HI fused data due to the reasons discussed earlier in this section. In this dissertation, both a maximum likelihood classifier (MLC) and a support vector machine (SVM) are tested as candidates for each individual pairwise classifier.

5.4.2 Maximum Likelihood and SVM Classifier

The application of a pairwise classifier relies on the appropriate selection of each individual classifier. Herein, both a maximum likelihood classifier (MLC) and a support

vector machine (SVM) are evaluated because they are the most commonly used linear (MLC) and non-linear (SVM) classification techniques, and have been shown in the literature to be efficient for classification of both HI (Jia & Richards, 1994; Gualtieri & Cromp, 1998) and FWL data (Neuenschwander et al., 2009; Jung et al., 2014).

Maximum likelihood classification assumes that the probability distribution of a target class is determined by a multivariate normal distribution with the same dimension as the input feature vector (Richards & Jia, 2006). Under this assumption, the discriminant function for the maximum likelihood function can be written as

$$g_i(x) = -ln|\Sigma_i| - (x - m_i)^t \Sigma_i^{-1}(x - m_i), i = 1,2,\cdots,N, \quad (5\text{-}12)$$

where $\boldsymbol{x}$ is a feature vector, m_i is the mean feature vector obtained from class i training samples, Σ_i is the covariance matrix of class i samples, and N is the number of classes. The label for a test feature vector $\boldsymbol{x}$ is assigned by finding the largest $g_i(x)$ among all the N classes. Because of the nature of MLC (it relies on a multivariate distribution estimate), it is very sensitive to the ratio of training sample number to its dimension. In cases when the number of training sample is limited, an efficient feature selection algorithm is required to reduce the dimensionality.

The SVM has been found to be advantageous in many applications for remote sensing data classification (Melgani & Bruzzone, 2004). Linear SVM aims to find an optimal hyperplane that separates two classes of samples, and makes the distance from the closest sample point to the hyperplane a minimum (Burges, 1998). To describe this problem, we begin with the assumption that we want to assign a class label $y_i \in \{-1, +1\}$ to each feature vector x in a d-dimensional feature space R^d to separate the

feature vectors with a hyperplane. We then assume there exists a plane with normal vector *w* and bias *b* that separates the two classes. We can then build a discriminant function as:

$$f(x) = w \cdot x + b, \tag{5-13}$$

and the label is assigned to each *x* using decision function $sign[f(x)]$. The hyperplane can thus be fixed by estimating *w* and *b* by:

$$y_i * f(x) > 0, i = 1,2,\cdots,L, \tag{5-14}$$

where *L* is the number of training samples. With the goal of optimizing the hyperplane by maximizing the geometrical margin between two classes, which can be represented as $\frac{2}{\|w\|}$, this problem can be translated into

$$minimize: \|w\|^2/2$$

$$subject\ to: y_i * f(x) \geq 1. \tag{5-15}$$

This is an optimization problem with linear constraints that can be solved by converting it to a dual problem using a Lagrangian formulation which is given as

$$maximize: \sum_{i=1}^{L} \alpha_i - \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$subject\ to: \sum_{i=1}^{L} \alpha_i y_i = 0\ and\ \alpha_i \geq 0. \tag{5-16}$$

The most interesting characteristic of SVM is its capability to address nonlinear classification with the help of a kernel function. In cases where the decision surface

function is a nonlinear function, a similar optimization process of separating the two classes can be achieved if we can find a nonlinear transformation Φ that maps the feature vectors to a higher-dimensional space where they are linearly separable again. We can see from equation 5-16 that the solution for the hyperplane involves computation of the inner product $(x_i \cdot x_j)$; therefore, solving this problem in a transformed feature space requires computation of a transformed inner product $[\Phi(x_i) \cdot \Phi(x_j)]$. If a "kernel function" *K* can be found to satisfy the following condition: $K(x_i, x_j) = [\Phi(x_i) \cdot \Phi(x_j)]$, then it is possible to avoid solving for Φ and computing $[\Phi(x_i) \cdot \Phi(x_j)]$; we can use $K(x_i, x_j)$ instead. Therefore equation (4-5) is transformed to

$$maximize: \sum_{i=1}^{L} \alpha_i - \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j)$$

$$subject\ to: \sum_{i=1}^{L} \alpha_i y_i = 0\ and\ 0 \le \alpha_i \le C, \qquad (5\text{-}17)$$

where *C* is a constant introduced in the nonlinear case in order to control the error (Burges 1998). The exact implementation of nonlinear SVM depends on the form of the kernel function. Common kernel functions include linear, polynomial and Gaussian radial basis function (RBF) kernels. Under the condition that the number of features is not extremely large, RBF kernels shows advantages over the other two including: capability for handling nonlinear cases, requiring less parameters to be determined compared to the polynomial kernel, and maintaining stable performance with a broad range of parameters, i.e., the numerical risks for RBF is relatively low (Chang et al., 2010). Therefore, RBF is employed in this research. The form of RBF can be expressed as

$$K(x_i, x_j) = exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \quad (5\text{-}18)$$

where Υ is a tunable parameter that defines the range one single training sample could reach. With the introduction of a kernel function, SVM can effectively implement nonlinear learning, and this is proven to be very advantageous in some cases when MLC fails to deliver sufficient accuracy (Melgani & Bruzzone, 2004).

Using SVM for multi-class classification involves the selection of an appropriate implementation strategy (one-against-all, one-against-one), since it is capable of only two-class separation. However when employed in a pairwise classifier frame, each individual SVM classifier faces only one two-class labeling problem and therefore is not necessary to adapt it for the multi-class case.

5.4.3 Feature Selection Algorithms

To optimize the performance of each pairwise classifier, appropriate feature selection methods are also required. Feature selection is widely used prior to applying a classification method; because, in many cases, a reduced feature space is beneficial for pattern learning, especially for HI (Serpico & Bruzzone, 2001). Two key components of feature selection are the search strategy and the criteria for feature importance. Search strategies can be separated into optimal and suboptimal searches. Optimal feature searches aim to find the best subset of features subject to the evaluation criteria, while the suboptimal search finds a good solution but is not guaranteed to be the best subset. Because an exhaustive search is difficult to avoid for an optimal search, suboptimal searches are usually more computationally effective. In this dissertation, because feature

selection needs to be applied to each individual classifier for feature subset optimization, a suboptimal search is the more realistic solution (Serpico & Bruzzone, 2001).

Sequential Floating Forward Selection (SFFS) is an effective tool for suboptimal feature selection, and works well with a large number of input features (Ververidis & Kotropoulos, 2008). It utilizes the correct classification rate (CCR) of a classifier as the criterion for feature ranking in a wrapper-based framework. Sequential searching means the SFFS algorithm adds or removes features one at a time into the subset for a wrapper. Starting from an empty set and incrementally adding features is called forward selection, whereas starting from the whole set and deleting one feature at a time is called backward selection. The SFFS consists of a forward step (insertion) and a conditional backward (deletion) step, which assists in avoiding convergence to a local maxima of the criterion function. The SFFS search is optimized by running a series of statistical tests whose accuracy relies on estimation of the variance of the CCR during cross-validation repetitions. SFFS can be used with any classifier in the context of any wrapper and therefore is used as the feature selection tool for pairwise classifiers with both MLC and SVM in this dissertation.

Chapter 6 Data Sets Description

Two data sets with simultaneous FWL and HI acquisition are analyzed in this dissertation. The first data set is located in Moran, Wyoming, close to the Snake River. The second data set was acquired over the Eel River Critical Zone research site named "Rivendell" located in Mendocino, California. The data sets will be referred to as Moran and Rivendell hereafter. Both data sets were collected by the National Center for Airborne Laser Mapping (NCALM). System characteristics and configuration, data set details, and methods for determining a reference classification sample for each data set are discussed in this chapter.

6.1 Moran Data Set

6.1.1 System Characteristics and Acquisition Details

Full waveform LiDAR, HI, and high resolution orthophotography were collected by NCALM in August 2012, with all sensors mounted on the same airborne platform. The FWL data set was collected with an Optech Aquarius LiDAR system with a green (532 nm) laser scanner using a 12-bit full-waveform digitizer. The system produced both discrete return and FWL data simultaneously. The HI was acquired with a CASI-1500, a visible and near-infrared wide array hyperspectral imager. Seventy-two spectral bands were recorded across the working spectrum of the CASI-1500, with a ground pixel spatial resolution of 1.2 m. The HI imagery was processed and georeferenced using data from the same POS system as the FWL. High-resolution aerial photography was also flown for the area using a DIMAC 60 Mpixel digital camera. The imagery was orthorectified with a final ground pixel size of 5 cm, which is significantly higher spatial resolution than the

HI. System and acquisition parameters for the FWL, HI and orthophotography are given in Tables 6-1, 6-2, and 6-3.

Table 6-1 LiDAR system specification and acquisition parameters

Specifications	
Parameter (Unit)	**Value**
Laser wavelength (nm)	532
Range capture	Up to 4 range measurements
Pulse width (FWHM in ns)	8.3
Digitization interval (ns)	1
Analog-to-digital quantization	4096 (12-bit)
Scan field of view (degree)	0 - 25
Beam divergence (mrad)	1
Pulse repetition rate (KHz)	33, 50, 70
Flight height (Above ground level, m)	300 - 600
Acquisition parameters	
Pulse repetition rate (KHz)	33
Flight height (Above ground level, m)	510
Shot density (1/m2)	4.2
Scan field of view (degree)	20

Table 6-2 CASI-1500 (Hyperspectral) specification and acquisition parameters

Specifications	
Parameter (Unit)	**Designed Value**
Spectral rage (continuous coverage) (nm)	380 - 1050
Number of spectral channels	Up to 288
Number of spatial pixels	1500
Total field of view (degree)	40
Instantaneous field of view (IFOV, mrad)	0.49
Focal ratio	f/3.5
Spectral resolution (FWHM) (nm)	< 3.5
Pixel size (μm)	20 X 20
Acquisition parameters	
Spectral rage (continuous coverage) (nm)	366.4 - 1043.7
Number of spectral channels	72
Flight height (Above ground level, m)	2016.6
Pixel size on the ground (m)	1.2

Table 6-3 DIMAC specification and acquisition parameters

Specifications	
Parameter (Unit)	**Designed Value**
Sensor size (mm)	53.9 X 40.4
Pixel size (μm)	6 X 6
Lens focal length (mm)	70
Shutter (sec)	1/125 to 1/500
Image output pixels	8984 X 6732 (60MP)

Acquisition parameters

Flight height (Above ground level, m)	510
Pixel size on the ground (cm)	5

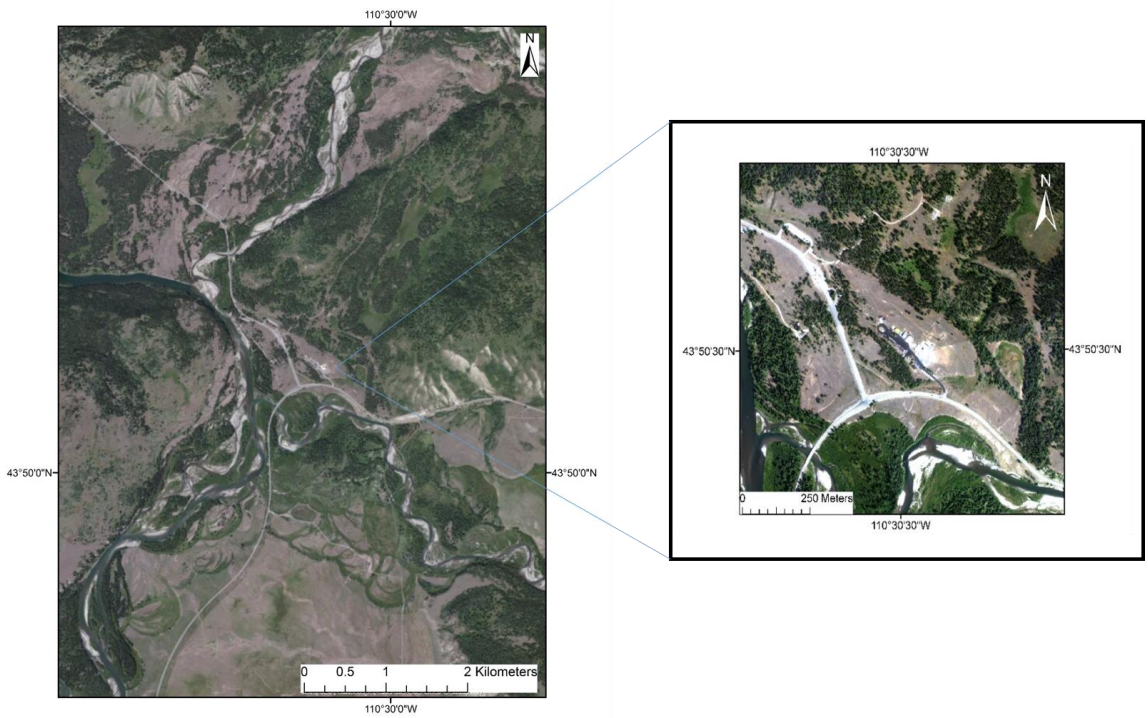


Figure 6-1 Moran area overview and experimental site in RGB (643 nm, 548 nm, and 462 nm) bands.

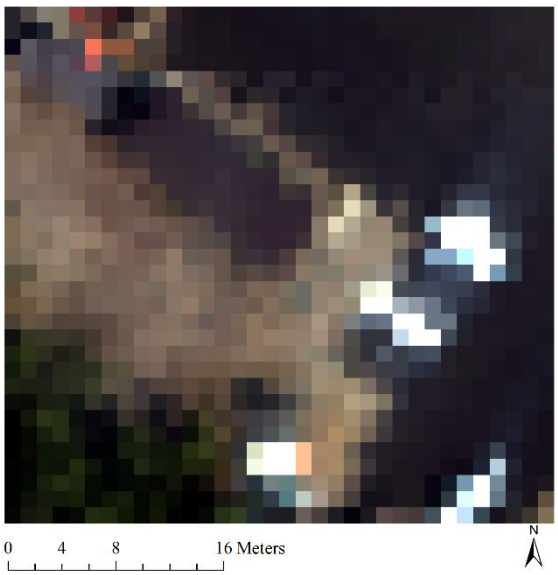
The selected study area is a complex close to the Snake River, with a size of 1200m × 1200m. It consists of a built up area in the center, a surrounding forested area, and portion of the Sane River river bank on the south and west sides, as shown in Figure 6-1.

6.1.2 Ground Truth Labeling

Ground truth labeling was accomplished using manual interpretation of the high-resolution orthophotos and HI. Given the different ground pixel sizes between the orphophoto and HI (see prior section), one HI pixel corresponds to 576 orthophoto pixels.

A comparison of the HI and orthophoto resolution is shown in Figure 6-2. In order to avoid zooming issues, the region of interest (ROI) for each class was first chosen in the HI. Since the HI and aerial photographs were obtained on the same platform, and georeferenced by the same GNSS/INS integration system, the ROIs could be reconciled via a common map projection using the ENVI (Environment for the Visualization of Images, version 4.8, Exelis Visual Information Solutions, Boulder, Colorado) ROI tool. Once the HI identified ROIs were viewed on the orthophoto, pixels of classes other than the one of interest were trimmed from the ROIs. The final ROIs were used as high-resolution orthophoto validated ground truth for the classification experiments.

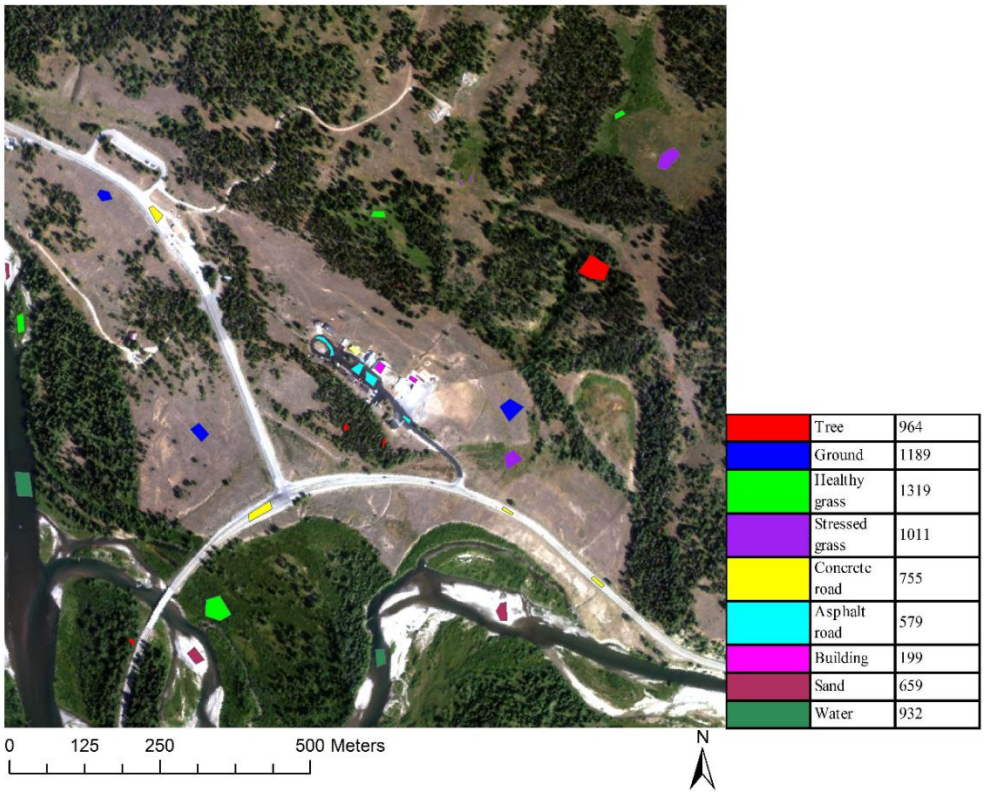
Nine natural and synthetic target classes were identified for the study area: trees, bare ground, healthy grass, stressed grass, concrete road, asphalt road, building, sand, and water. The training and validation sample ROIs for each class are shown in Figure 6-3. Note that the ratio of training to validation samples is approximately 2:1. The training samples were selected such that samples from each class are spatially distributed (i.e., not clustered in a single portion of the image) when they exist in multiple locations of the study area. The validation samples were selected such that they are spatially separated from the training samples whenever possible, enabling a location based stratified sampling pattern.



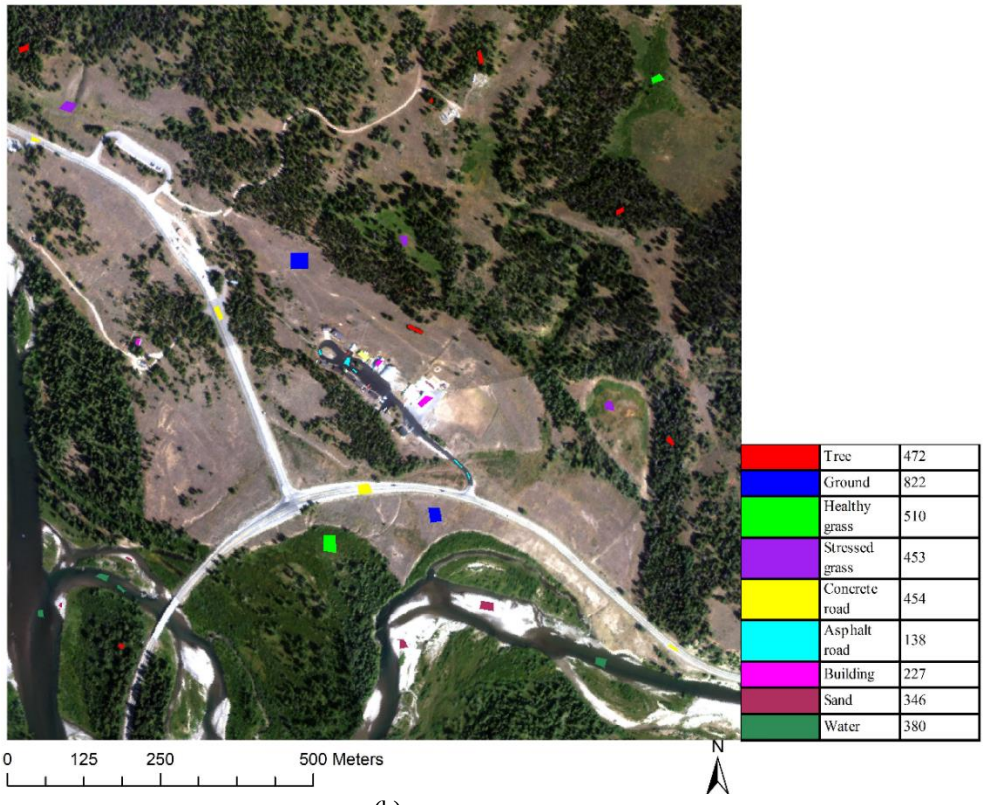
(a)

(b)

Figure 6-2 A comparison of HI and aerial photograph level of detail for a common area: (a) HI level of detail sample; (b) Orthophoto level of detail sample. Pixel sizes are 1.2 m and 0.05 m, respectively.



(a)



(b)

Figure 6-3 Color coded images of training and validation samples: (a) Training samples; (b) Validation samples.

6.2 Rivendell Data Set

6.2.1 System Characteristics and Acquisition Details

The Rivendell data set was collected in July 2014 and contains the South Fork Eel River on the west side going from north to south and Elder Creek in the center going from east to west. The data was acquired with an approximate coverage of 4000m (east-west) by 5000m (north-south). Full waveform LiDAR data and HI were collected by NCALM using the same systems introduced for the Moran data set (see tables 6-1 through 6-3).

Table 6-4 Acquisition parameters for the Rivendell data set

LiDAR	
Parameter (Unit)	**Value**
Pulse repetition rate (KHz)	33
Flight height (Above ground level, m)	600
Shot density (1/m2)	3.57
Scan field of view (degree)	20
Hyperspectral Sensor	
Spectral rage (continuous coverage) (nm)	368.8 - 1041.3
Number of spectral channels	48
Flight height (Above ground level, m)	435.8
Pixel size on the ground (m)	1

Relevant differences in the data acquisition parameters are summarized in Table 6-4. The Rivendell data set is introduced as the second test data set because of it contains

several vegetation types with similar spectral signatures but different height ranges and canopy shapes; therefore, it can be used for testing the capability of the proposed methods for separating similar targets based on vertical structure. Also, the training sample set was created by expert with in-depth knowledge regarding vegetation grown in this area, with the help of multiple kinds of remote sensing data and ground observation, thus the quality of the training sample will benefit the assessment of classification accuracy.

6.2.2 Ground Truth Labeling

Ground truth labeling was obtained in a different manner than for the Moran data set. Six types of ground covers were identified by interpreting high-resolution imagery and a 1 m resolution LiDAR-generated DEM (this work was performed by Collin Bode at the Department of Integrative Biology, University of California, Berkeley) and are listed as follows: building, riparian, meadow, chaparral, broad leaf vegetation, and conifer vegetation. An overview of the area with the distribution of identified targets is shown in Figure 6-4.

The ratio of training and testing data is 1:2, with the class-specific details listed in Table 6-5. This ratio is chosen because more reference samples are available for the Rivendell data set than Moran, so a smaller percentage could be kept as training samples and still provide enough instances to train the classifiers. Also this change of the ratio tests the performance of the proposed classification framework under different working conditions. All classes except the building have 2000 training samples and 4000 validation samples due to the limited amount of buildings in the study area. Coverage for the building class is quite limited in the study area; only 135 voxels were picked for this

class. Half of the training data were used to determine VEDC dimension for the SWF and classification parameters; the remainder were used to train the classifiers.

Table 6-5 Size of training and testing data by class (in number of HI pixels)

Class	***Number of training sample***	***Number of testing sample***
Broad leaf	2000	4000
Building	45	90
Chaparral	2000	4000
Conifer	2000	4000
Meadow	2000	4000
Riparian	2000	4000

To evaluate the performance of the proposed methods for separating targets with structural complexity and spectral similarity, the broad leaf and conifer vegetation classes were further identified at the species level. Madrone and oak trees were identified in the broadleaf class, and douglas-fir and redwood trees in the conifer class. Note that not all samples in the broad leaf and conifer vegetation class can be positively identified as one of the individual species; thus, the samples for the four specific tree types are limited, with 6000 samples identified for the madrone and douglas-fir species and 1500 samples for the oak and redwood, respectively.

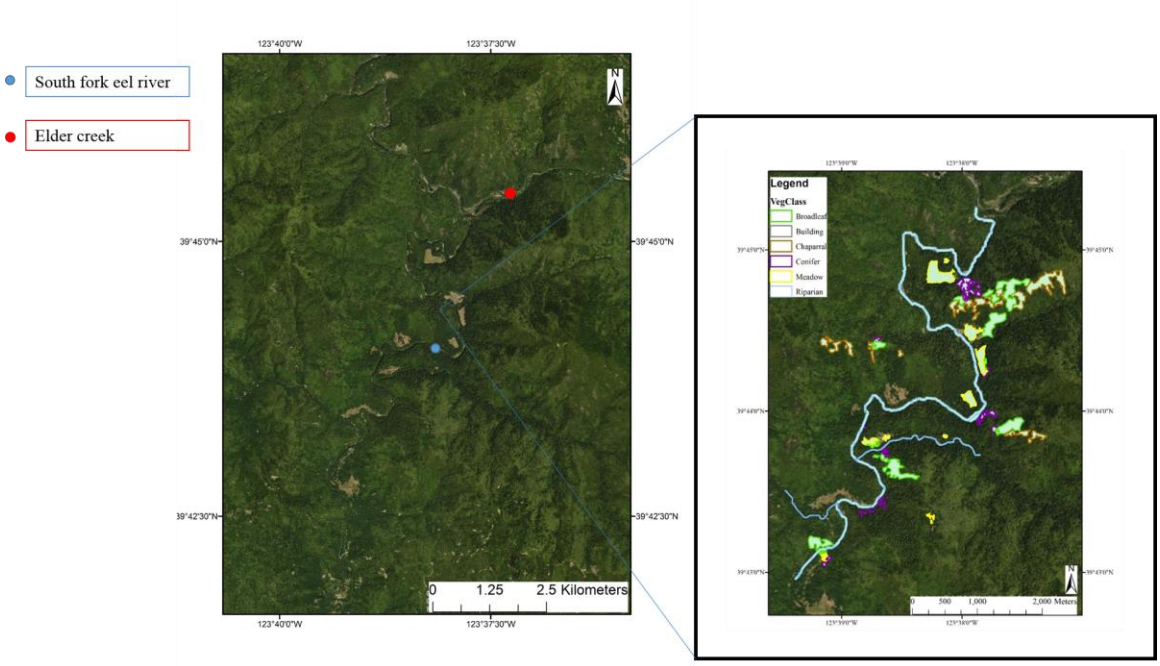


Figure 6-4 Rivendell data set with ROI

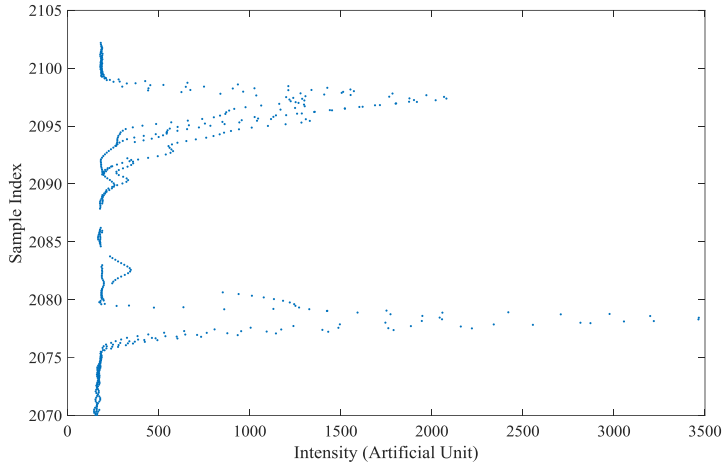
Chapter 7 Experimental Results and Analysis

In this chapter, results generated by the proposed FWL/HI classification methods are presented. Feature selection from the FWL/HI feature space is performed, and the pairwise classifier framework is applied using two widely used classifiers, MLC and SVM. Classification accuracy is evaluated using FWL and HI individually for comparison, with the results indicating classification improvement when using the combined FWL/HI data sets. Since a discrete LiDAR point cloud (LPC) was collected simultaneously with the FWL data, it is also evaluated both alone and in conjunction with the HI for classification tasks. The FWL and LPC data is further compared to evaluate the effectiveness of waveform data for classification performance.

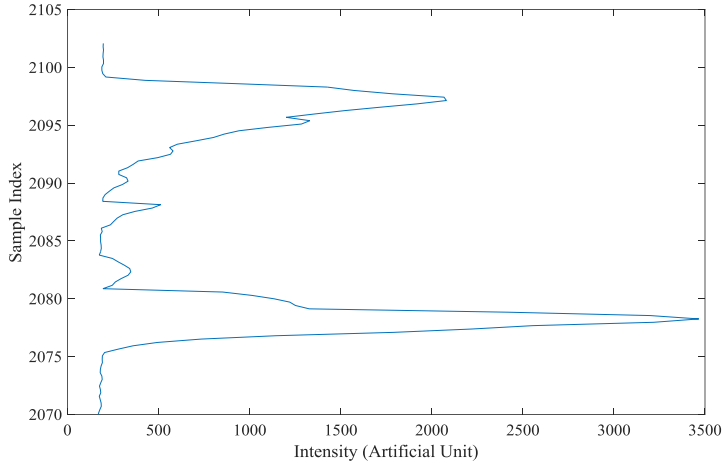
7.1 Results and Analysis of Moran Data Set

7.1.1 Waveform Voxelization

The voxel reference frame was defined to match the georeferenced HI pixel locations. The ground footprint of each voxel was a 1.2m × 1.2m square corresponding to the HI pixel size. A total of one million voxel columns were constructed in the experimental area. A small percentage (about 2%) of voxel columns contained a very small number of return samples and were thus removed from further processing. The height of each voxel was set to be 15 cm, which corresponds with the FWL digitizer interval of 1 ns. An example of a SWF generated from all samples located in one column of voxels is shown in Figure 7-1.



(a)



(b)

Figure 7-1 Waveform synthesis example: (a) All waveform samples Intersecting a Voxel Column (containing 240 voxels), (b) Synthesized waveform.

A total of 10,209,526 raw waveforms from 13 LiDAR flight lines with partial or full coverage over the experimental area were used to generate the SWFs. With such a high density of original waveforms, each SWF was generated from an average of approximately 10 original waveforms. A waveform density map of the experimental area is shown in Figure 7-2.

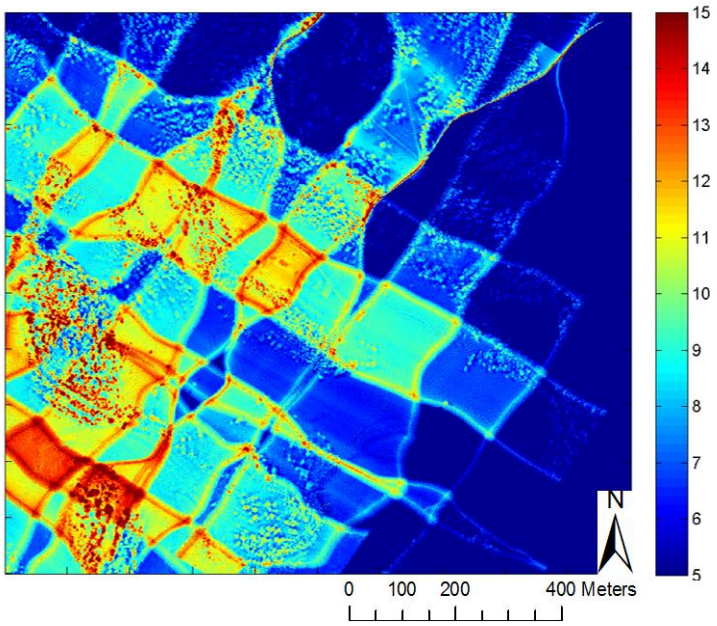


Figure 7-2 Waveform density map

7.1.2 Disjoint Data Classification

7.1.2.1 VEDC Feature Generation

The dimension of the VEDC was determined using both supervised and unsupervised methods. This was done because the optimal dimension and height segmentation of VEDC may vary for different data sets. The unsupervised approach determines VEDC dimension based on the accuracy of VEDC only classification. However, in instances where multiple classifiers are used, an unsupervised approach may be preferred. Therefore, an unsupervised feature generation method is designed attempt to maximize the entropy of VEDC features using a relatively low number of bands.

The supervised VEDC feature generation was first conducted by examining the relationship of classification accuracy using VEDC features only versus the number of VEDC components. The training SWFs were aligned by height, and the height range of interest for VEDC chosen as the range that 80% of the return samples with an intensity

higher than 20% of the maximum return intensity fell into. The VEDCs were then computed based on the voxel components falling within this height interval. The relationship between classification accuracy using VEDC only versus the dimensionality of VEDC is shown in Figure 7-3. The best performance for SVM classification was achieved with 7 VEDC bands (65.47%), with the classification accuracy decreasing to approximately 60% when the dimensionality (number of bands) is increased beyond 8. The supervised VEDC dimension was therefore set to be 7 for this data set. As previously mentioned, the height of the last return, penetration depth, maximum amplitude, and skewness of the SWFs were also calculated as additional vertical waveform features, producing a feature vector with 11 components describing each SWF. For the HI, 15 principal components were retained for each HI pixel after PCA, representing over 99% of the cumulative variance. Feature extraction for the synthetic waveforms generated from the LPC was implemented in the same manner as for SWF.

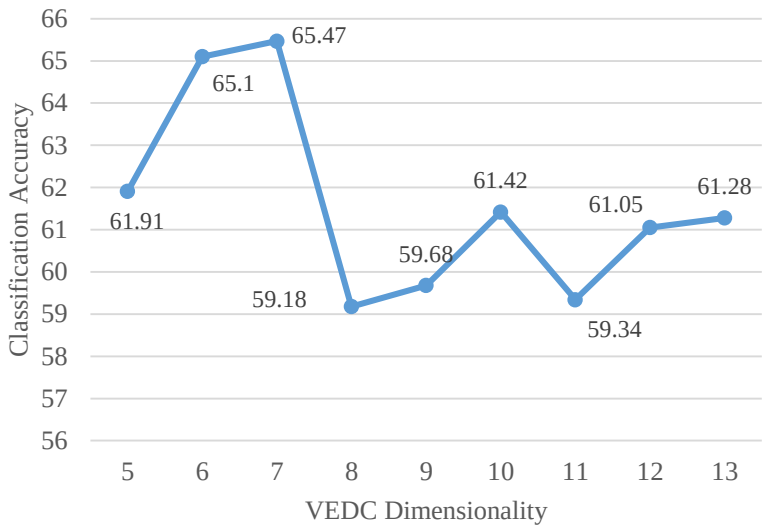


Figure 7-3 Classification accuracy using SVM for VEDC only

As reviewed in Chapter 5, VEDC feature generation via unsupervised methods is more flexible, particularly when using multiple classifiers, and is thus more appropriate for many applications. Based on the method introduced in Chapter 5, the number of VEDC bands is determined using the entropy sum for all bands.

If we assume the VEDC band number is *N*, then the entire height range *H* can be divided into *N* parts, with the starting height h_i of the *i*th part given as

$$h_i = (i-1) * \frac{H}{N}. \tag{7-1}$$

Thus, we can evaluate the overall entropy against varying VEDC dimension for the training data. It can be inferred from the supervised VEDC dimension determination reviewed in the prior paragraph that VEDC features are able to preserve useful waveform information with relatively low dimensionality. Using half of the training data, the overall entropy as a function of band numbers, ranging from 1 to 30 is shown in Figure 7-4. Note that the band number describes the total number of bands of VEDC over the entire vertical observation range, and doesn't refer to a band index. The actual relationship of summed entropy value versus the unsupervised VEDC band number can be fitted by a two factor exponential function. Although the entropy shows constant grow with an increasing number of bands, the pace slows after a certain point. This point can be found be examining the curvature of the fitted function curve, the closest integer to the peak of curvature ranging from 1 to 30 is 10. Therefore, the 10 bands are selected as our initial VEDC dimension.

The entropy of each VEDC band is displayed in Figure 7-5 along with the entropy contributed by each land cover class. According to the plots, bands 2 and 3 contribute

relatively little information, and a majority of the land cover classes do not contain significant information in bands 2 and 3. Therefore, we combined band 2 with band 1, and band 3 with band 4, and reduce the overall VEDC band number to 8. Classification results of using both VEDC generation methods (supervised and unsupervised) will be compared in Section 7.1.3.2 to validate the unsupervised method.

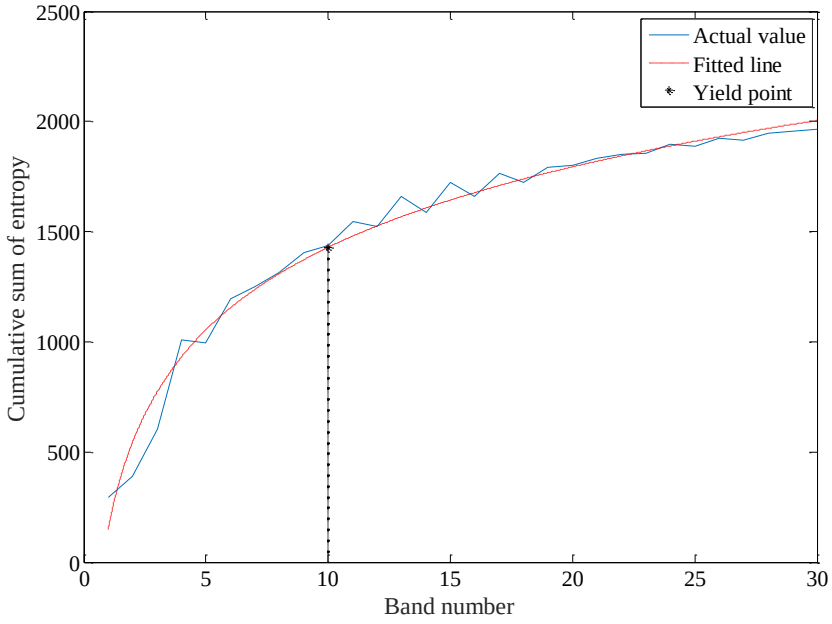
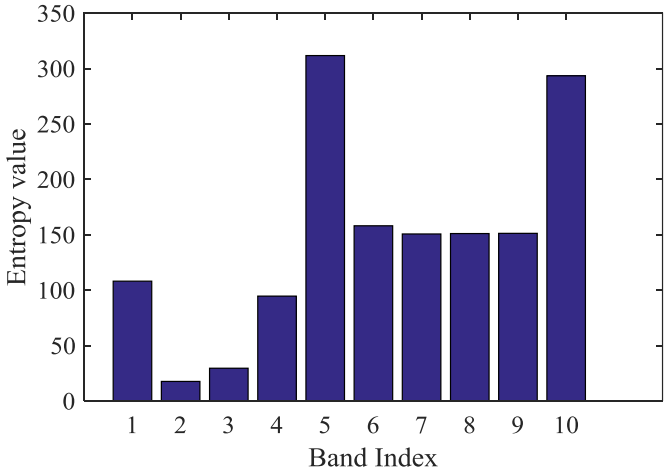
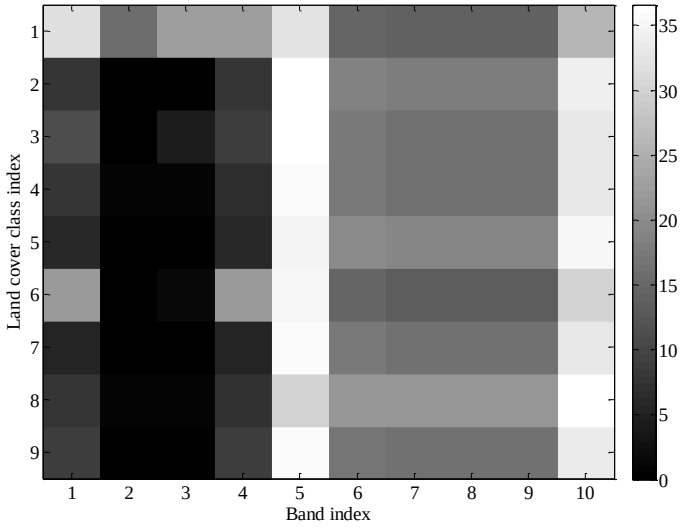


Figure 7-4 Entropy value versus VEDC band number



(a)



(b)

Figure 7-5 Component-specific entropy: (a) Entropy of each VEDC component; (b) Entropy of each land cover class

7.1.2.2 Classification Accuracy

Classification results using solely SWF, HI and synthetic waveform generated from LPC (still denoted by LPC in the table) are shown in Table 7-1, the overall accuracy of each method is listed and κ denotes the kappa agreement coefficient. Using HI data with the SVM pairwise classifier achieved the best overall performance whereas the LPC showed the poorest performance for both classifiers. The SWF with MLC achieved an acceptable performance with the 11 extracted features, demonstrating that the VEDC features are effective metrics for quantifying the intensity-height distribution information from the SWF, and they also provide a significant advantage over just using metrics derived from the LPC.

Table 7-1 Classification accuracy and kappa agreement of using FWL, HI and the discrete LiDAR point cloud

Data type	***Overall accuracy (%)***		***Kappa agreement (%)***	
	MLC	SVM	MLC	SVM
SWF	79.83	71.30	75.48	65.68
HI	78.01	85.82	74.76	82.12
LPC	65.66	60.17	50.88	53.72

7.1.3 Fused Data Classification

7.1.3.1 Classification using the entire SWF

The SWFs created using the method described in Chapter 5 can be directly used for classification without performing the VEDC feature extraction. Combining each waveform's 240 samples with the 15 PCA bands from HI produces a feature vector with 255 dimensions. Classification with these fused feature vectors using MLC produces a low overall accuracy of 43.7%, potentially due to the large dimension of the input feature vectors. However, an acceptable overall accuracy of 82.0% was achieved using SVM, classification details are presented in Table 7-2.

Table 7-2 Confusion matrices of SVM classification for direct integration of entire SWF and HI

Predicted / ***Actual***	**Tree**	**Healthy grass**	**Ground**	**Road**	**Pitch road**	**Building**	**Sand**	**Water**	**Stressed grass**	**Total**
Tree	458	6	4	0	0	0	0	0	4	472
Healthy grass	7	478	0	0	0	0	0	0	25	510
Ground	79	1	698	0	0	0	0	0	44	822
Road	0	0	7	381	0	1	65	0	0	454
Pitch road	0	1	2	16	119	0	0	0	0	138
Building	1	7	13	39	14	42	111	0	0	227
Sand	0	0	1	71	3	4	265	2	0	346
Water	4	3	0	0	1	0	0	372	0	380
Stressed grass	4	22	121	0	0	0	0	0	306	453
Total	553	518	846	507	137	47	441	374	379	0.820

Another approach to prepare waveforms for classification is to transform the SWFs using PCA, similar to the approach commonly applied to HI. By examining the covariance structure of the waveforms, 35 bands were extracted from each synthesized waveform, representing over 99% of the cumulative variance. Therefore the input feature vector has a reduced dimension of 50 (compared with directly using all waveform samples). The accuracy of MLC classification increases to 69.3% (an improvement of 25.6%), which is most likely due to the significant reduction in the feature vector dimension. However, the accuracy of SVM classification dropped to 56.0%, indicating that the transformed waveform intensity may not sufficiently preserve discriminatory information from the SWFs. Possible reasons why PCA was not able to effectively extract information from SWF include: 1) spatial information is disregarded in the PCA transformation, and therefore the discriminatory power of SWFs is reduced by only examining the intensity information; 2) certain components of SWFs may contain noise which would be considered as large variance bands by PCA.

In general, the performance of direct fusion of SWF or PCA transformed SWF with HI falls below standalone classification of HI only, which suggests an effective feature extraction method designed for SWF is required. Therefore, a classification experiment is carried out after VEDC feature extraction in the following section.

7.1.3.2 Classification using proposed methods

Proposed methods utilize a pairwise classification framework with MLC and SVM used as the individual pairwise classifiers. The features used for each classifier and each pair of classes were selected using a SFFS approach. Feature selection and classification results are shown in this section.

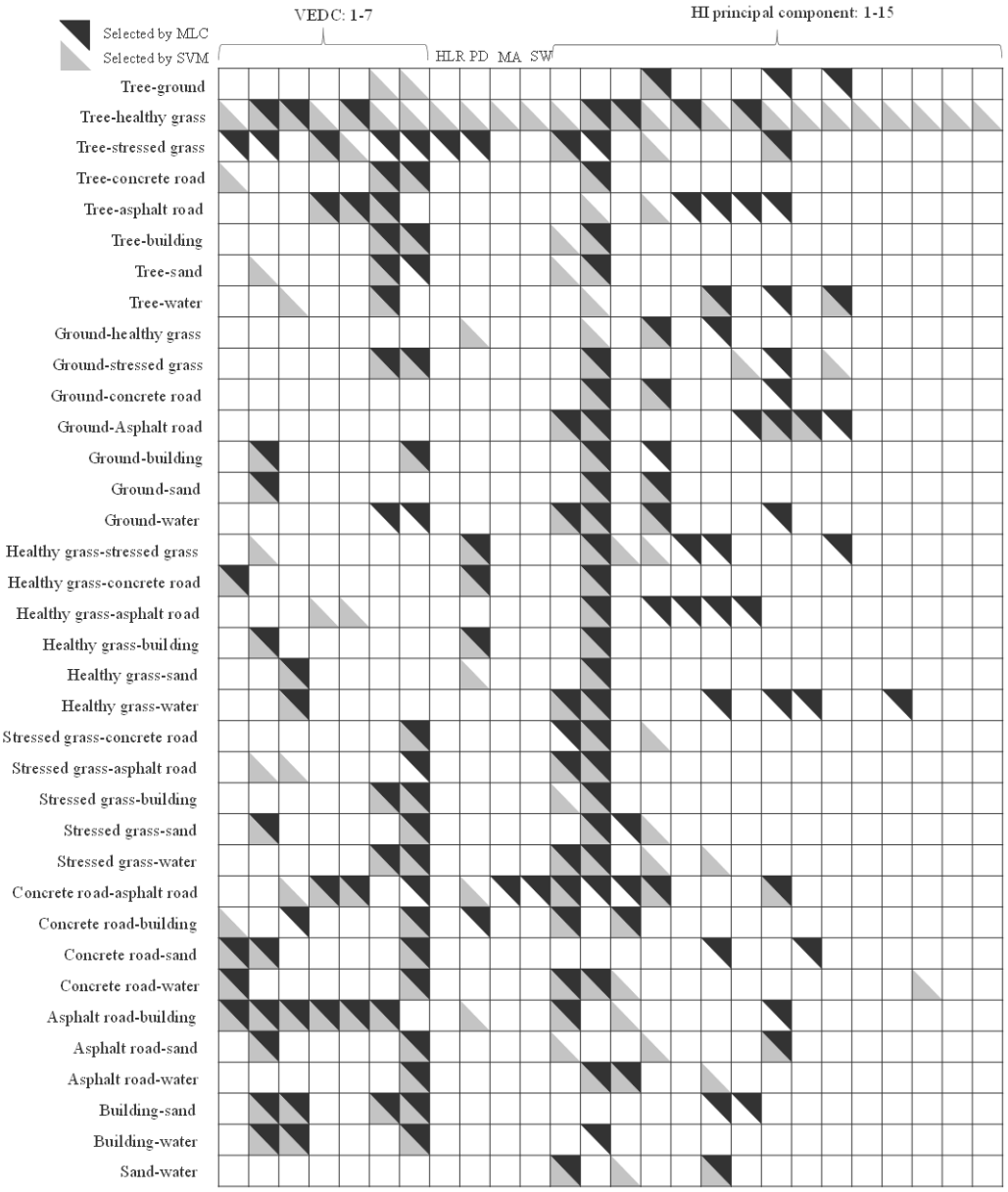
(1) Feature selection results

Classification of the fused SWF and HI data set using the proposed methods required feature selection for each individual classifier by exploiting a portion of the training data. We used 20% of the training samples to perform a feature importance estimation for each method (MLC or SVM). SWF and HI features were ranked together, features are selected using SFFS method based on classification accuracy; the selected features for each class are shown in Table 7-3. The numbers above the brackets are the index of the features in the supervised VEDC or HI principal component bands. For unsupervised VEDC, HI features selection remain the same, SWF features selection shifted bands in some cases due to the different height segmentation, but the total number of class pairs choosing SWF features remains the same. Acronyms were used for the other four LiDAR derived features: HLR for height of last return, PD for penetration depth, MA for maximum amplitude, and SW for SWF skewness.

The important features for MLC and SVM may differ for each pair of classes, but generally they contain many common features. It can be easily observed in Table 7-3 that the VEDC components were the more important SWF features in most cases. The penetration depth feature was also chosen in several pairs, but the height of last return, maximum amplitude, and skewness features were seldom chosen in any class pair. The chosen HI components for the pairwise classifiers were mostly from the first 10 components of PCA, indicating that the number of preserved principal components was adequate for classification. When using MLC for feature ranking, HI features were picked for all class pairs (36 total), and SWF features were picked for 30 class pairs. SVM feature ranking reserved HI features for 34 pairs and SWF features for 32 pairs. In

most cases, when SWF features were less important than HI, the ground class was involved. Both the MLC and SVM classifiers determined that only HI features were necessary for separating ground from concrete roads or asphalt roads. This is a natural complement to SWF because reflected waveforms from ground and the two type of roads are quite similar.

Table 7-3 Selected features for each type of classifier and each pair of classes; HLR stands for height of last return, PD for penetration depth, MA for maximum amplitude, and SW for SWF skewness.



(2) Classification accuracy evaluation

Classification accuracies of supervised and unsupervised VEDC feature fused with HI are compared in Table 7-4. We can see that the VEDC features generated by the unsupervised method have a slightly lower MLC accuracy (by 2.9%) and higher SVM accuracy (by 2.6%). Because in most cases an unsupervised method is optimal for VEDC feature generation, and the unsupervised VEDC achieved the best accuracy out of all four combinations, we will use unsupervised VEDC for all further experiments.

Classification accuracy for each class using the chosen features, as well as the overall accuracy is shown in Table 7-5. It is obvious that the LPC/HI was outperformed by SWF /HI in all cases, indicating that the point cloud is not as effective as SWF when used in combination with HI. Considering that a similar conclusion was reached earlier, we can safely infer that using the proposed features extracted from a LPC is not able to offer the same level of classification information content as what is provided by SWF.

For fused SWF and HI, the classification accuracy with either type of classifier exceeded the accuracy of using either data set alone. The best performance individually was achieved by HI alone with a SVM classifier, this accuracy was exceeded by 0.3% with MLC and 9.4% with SVM using the fused SWF/HI dataset. The gain in kappa agreement was also more significant with the fused data. Therefore the conclusion can be drawn that the fusion of SWF and HI provides a significant improvement in classification accuracy.

Table 7-4 Classification accuracy comparison between two types of VEDC

	Supervised VEDC+HI	***Unsupervised VEDC+HI***
MLC	89.0	86.1
SVM	92.6	95.2

Table 7-5 Classification accuracy using MLC and SVM with fused SWF/HI and LPC/HI

Class	***User's Accuracy (%)***				***Producer's accuracy***			
	SWF+HI		**LPC+HI**		**SWF+HI**		**LPC+HI**	
	MLC	**SVM**	**MLC**	**SVM**	**MLC**	**SVM**	**MLC**	**SVM**
Tree	86.2	95.6	86.2	88.1	75.5	98.0	65.2	64.0
Ground	85.9	100	98.7	94.5	99.3	96.8	96.3	99.4
Healthy grass	90.6	96.9	83.5	95.5	75.4	92.9	84.4	82.4
Stressed grass	85.7	92.1	79.5	79.7	84.3	97.4	85.5	96.0
Concrete road	89.2	99.8	92.9	98.5	79.9	85.2	82.4	95.7
Asphalt road	43.5	90.6	47.1	78.3	100	87.4	36.5	100
Building	96.9	60.8	43.2	95.6	77.7	99.3	43.0	81.3
Sand	87.6	98.6	87.0	93.1	98.7	100	100	100
Water	84.7	99.5	46.3	59.5	100	100	98.9	100
Overall Accuracy								
Accuracy	86.1	95.2	80.9	88.7	86.1	95.2	80.9	88.7
κ	84.0	94.5	78.0	87.0	84.0	94.5	78.0	87.0

As shown in Table 7-5, the performance of SVM exceeded MLC for most classes, and the difference in overall accuracy was 9.1% for SWF/HI. A possible explanation for the difference between classifiers is that the distribution of features from the two data sources may deviate from the MLC assumption that all features are associated with a common Gaussian distribution. The only exception was for buildings, where MLC showed better accuracy than SVM. By examining the confusion matrices given in Table 7-6, we can see that SVM confuses these classes with similar classes more frequently than MLC, i.e., the concrete road and asphalt road classes are confused with the building class, and healthy grass, bare ground and tree classes are confused with the stressed grass

class. It is possible that even after transforming input features to a higher dimensional space, the samples are still difficult to separate via the SVM method because of their high homogeneity. However, the higher MLC user accuracies for the building classes does not necessarily indicate a stronger discrimination ability, but more likely that the distribution model obtained inside MLC was over-estimating the probability of presence of these two classes, at the cost of under-estimating the probability of similar classes. Furthermore, SVM exhibits a more stable performance among classes (12.4% user's accuracies standard deviation) than MLC (15.4%), which is an important characteristic that supports the use of SVM in this application.

Table 7-6 Confusion matrices of MLC and SVM classification for the fused data: (a) MLC results, (b) SVM results

(a)

Predicted / ***Actual***	**Tree**	**Healthy grass**	**Ground**	**Concrete Road**	**Asphalt road**	**Building**	**Sand**	**Water**	**Stressed grass**	**Total**
Tree	407	16	4	0	0	11	0	0	34	472
Healthy grass	17	462	0	0	0	0	0	0	31	510
Ground	63	33	706	0	0	19	0	0	1	822
Concrete Road	10	4	0	405	0	31	4	0	0	454
Asphalt road	0	16	0	61	60	0	0	0	1	138
Building	0	0	0	7	0	220	0	0	0	227
Sand	10	2	0	24	0	2	303	0	5	346
Water	16	32	0	10	0	0	0	322	0	380
Stressed grass	16	48	1	0	0	0	0	0	388	453
Total	539	613	711	507	60	283	307	322	460	0.861

(b)

Predicted / *Actual*	**Tree**	**Healthy grass**	**Ground**	**Concrete Road**	**Asphalt road**	**Building**	**Sand**	**Water**	**Stressed grass**	**Total**
Tree	451	13	6	0	2	0	0	0	0	472
Healthy grass	5	494	0	0	0	0	0	0	11	510
Ground	0	0	822	0	0	0	0	0	0	822
Concrete Road	0	0	0	453	0	1	0	0	0	454
Asphalt road	0	0	1	12	125	0	0	0	0	138
Building	0	1	10	62	16	138	0	0	0	227
Sand	0	0	0	5	0	0	341	0	0	346
Water	2	0	0	0	0	0	0	378	0	380
Stressed grass	2	24	10	0	0	0	0	0	417	453
Total	460	532	849	532	143	139	341	378	428	0.952

7.2 Results and Analysis of the Rivendell Data Set

Based on the results of the Moran data set, the efficacy of the proposed method was validated, and the fusion of unsupervised VEDC and HI with the SVM classifier exhibited the best performance among all possible candidates. To further evaluate the performance of the proposed method in an environment comprised of different target classes with similar vertical structures, the Rivendell data set is introduced for additional experiments.

7.2.1 Waveform Voxelization and Feature Extraction

As before, the reference frame of the georeferenced HI is used for voxel construction, with the voxel footprint size set to 1 m × 1 m (same as HI pixel size) and voxel height set to 15 cm to match the FWL digitizing frequency. Since the experimental area for Rivendell is much larger than the Moran data set (almost 13 times larger), and the number of class is smaller (6 general classes for Rivendell versus 9 for Moran), not all possible voxels were built for waveform voxelization and synthesizing in the experimental area. Only the training and validation voxels necessary for the analysis herein were built and the intersecting waveforms processed. Five partially overlapping FWL scans were flown over the area. Each column of voxels had approximately four intersecting waveforms from which to reconstruct the SWFs.

After the SWFs were obtained over each column of voxels, both the supervised and unsupervised VEDC approach were used for waveform feature extraction in the same manner as described in Chapter 5 and implemented for the Moran data set. Using half of the training data, the SVM classification accuracies with changing VEDC dimensions are shown in Figure 7-6 and the overall entropy as a function of number of bands, ranging from 1 to 30, is shown in Figure 7-7. The supervised approach suggests 8 bands provides the optimal option. For the unsupervised approach, although the entropy shows constant growth with an increasing number of bands, in agreement with the Moran entropy values, the pace begins to slow after a point of maximum curvature, which is also 8 bands in this case.

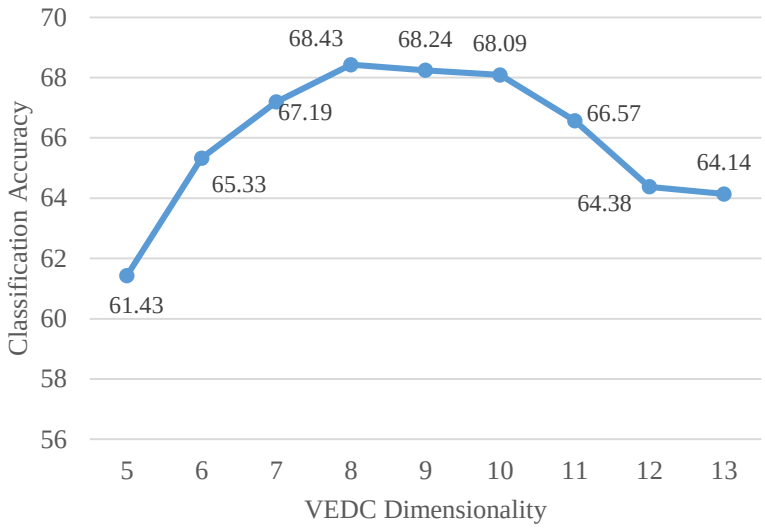


Figure 7-6 Classification accuracy using SVM for VEDC only

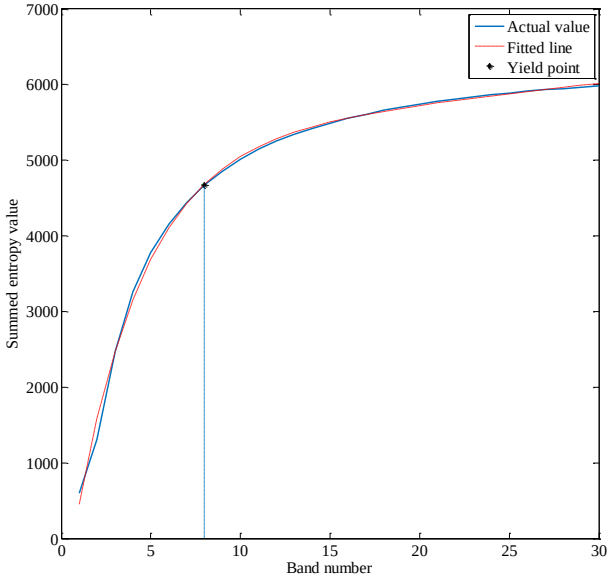
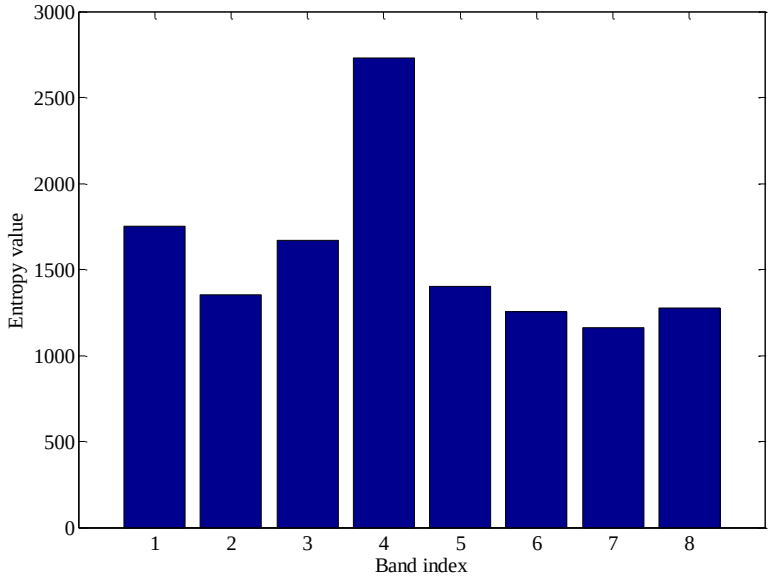
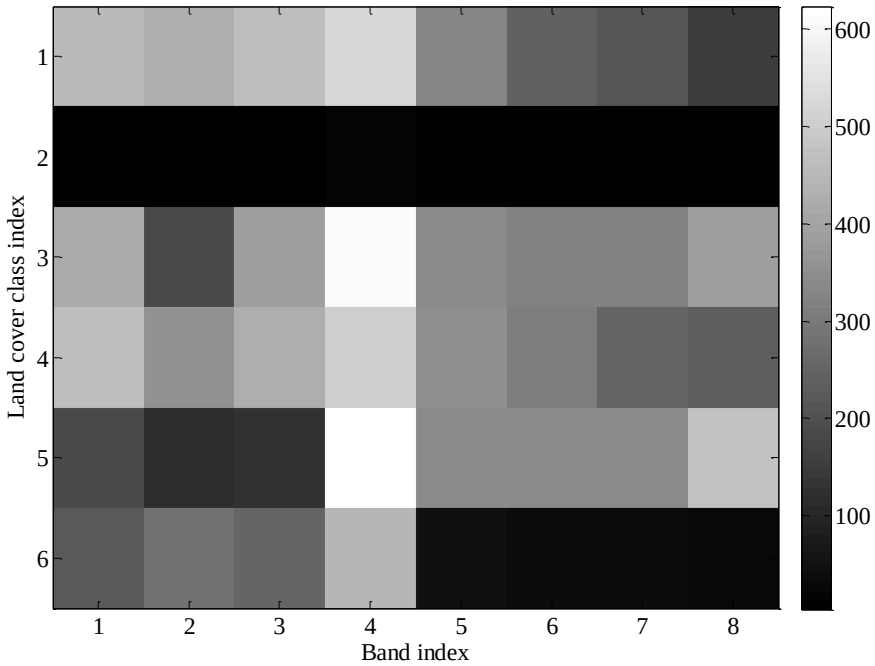


Figure 7-7 Entropy value verses VEDC band number for the Rivendell data set



(a)



(b)

Figure 7-8 Component-specific entropy for the Rivendell data set: (a) Entropy of each VEDC component; (b) Entropy contribution for each land cover class

The entropy of each VEDC feature is compared with class specific details in Figure 7-8. Note that the values for class 2 are significantly smaller because the small

size of class 2 training data. Unlike the Moran data set results, there are no VEDC bands that provide significantly less information than the other bands; therefore, all the 8 bands of VEDC are used as waveform features for fusion and classification. Since supervised approach also suggests an 8 band VEDC and none of the bands were merged for the unsupervised approach, both the VEDC methods will generate the same VEDC feature set for Rivendell data set. Finally, PCA of the HI indicated that 12 PCA bands are sufficient to encapsulate more than 99% of the data variance. Combined with the four additional SWF shape features, the same as used in the Moran data set (HLR, PD, MA, and SW), a feature vector with 24 bands is created for each voxel column.

7.2.2 Disjoint Data Classification

Classification results from SWF, and HI alone as well as the fused SWF/HI feature vectors using MLC and SVM are shown in Table 7-7. For SWF classification, SVM classification outperforms MLC significantly; whereas the two classifiers yield similar results for HI. Among the single data set classification combinations, HI outperforms SWF with both classifiers, and SVM for HI achieves the best overall accuracy of 89.72%. The confusion matrices generated from the SVM results for SWF and HI are shown in

Table 7-8 where it can be seen that although the overall performance is similar, SWF outperforms HI in chaparral, meadow, and riparian classes, whereas HI outperforms SWF in the other three classes.

Table 7-7 Classification accuracies of standalone data set and fused data

Data type	***Overall accuracy (%)***		***κ (%)***	
	MLC	**SVM**	**MLC**	**SVM**
SWF	64.67	84.14	55.92	81.15

HI	83.02	89.72	78.82	87.18
SWF+HI	84.98	97.13	81.27	96.43

Table 7-8 Confusion matrices: (a) VEDC classification using SVM; (b) HI classification using SVM.

(a)

Predicted / ***Actual***	**Broad leaf**	**Building**	**Chaparral**	**Conifer**	**Meadow**	**Riparian**	**Total**
Broad leaf	2820	0	307	789	7	77	4000
Building	1	38	13	12	0	26	90
Chaparral	156	1	3655	188	0	0	4000
Conifer	669	0	304	2977	0	50	4000
Meadow	2	0	0	15	3785	198	4000
Riparian	60	0	0	182	130	3628	4000
Total	3708	39	4279	4163	3922	3979	0.841

(b)

Predicted / ***Actual***	**Broad leaf**	**Building**	**Chaparral**	**Conifer**	**Meadow**	**Riparian**	**Total**
Broad leaf	3210	0	373	306	25	86	4000
Building	0	75	9	0	0	6	90
Chaparral	406	0	3408	137	0	19	4000
Conifer	265	0	64	3577	0	94	4000
Meadow	1	0	2	0	3996	1	4000
Riparian	50	2	15	166	8	3759	4000
Total	3932	77	3871	4216	4029	3965	0.897

7.2.3 Fused Data Classification

7.2.3.1 Six Class Land Cover Classification

To select the optimal feature subset for the fused data set, 20% of the training samples were used to perform a feature importance estimation for each classification method (MLC and SVM). The SWF and HI features were ranked together, features are selected using SFFS method based on classification accuracy; the selected features for each class are shown in Table 7-9. Features selected for MLC and SVM may differ for each pair of classes, but similarities can be observed. It is shown in Table 7-9 that the VEDC components were the more important SWF features in most cases. The other four waveform features are only selected in a few instances. When using MLC for feature ranking, HI features are picked for 14 class pairs (15 total), and SWF features are picked for 14 class pairs. Feature ranking for SVM reserved HI features for 14 pairs and SWF features for 14 pairs. For the only one classifier pair waveform features are not selected, the two target classes are meadow and riparian, both of which contain little vertical structural information.

Per Table 7-7, the fused data set has the best overall performance, exceeding the standalone HI accuracy for both classification methods, e.g., a 7.41% improvement for the SVM. The gain in kappa agreement is also significant for the fused data compared to the standalone SWF or HI. This observation agrees with the Moran data set in that the fusion of SWF and HI provides a significant improvement in classification accuracy. A 12.15% accuracy difference is also observed between SVM and MLC for fused SWF and HI, which shows a similar trend with the Moran data set. Confusion matrices of MLC and SVM classification for the fused Rivendell data set are shown in Table 7-10 where it can

be observed that most MLC misclassification occurs in the broad leaf, chaparral and conifer vegetation classes. This may be due to their overlapping height ranges and spectral signatures. However, with SVM, these classes are still well separated.

Table 7-9 Selected features for data set 2

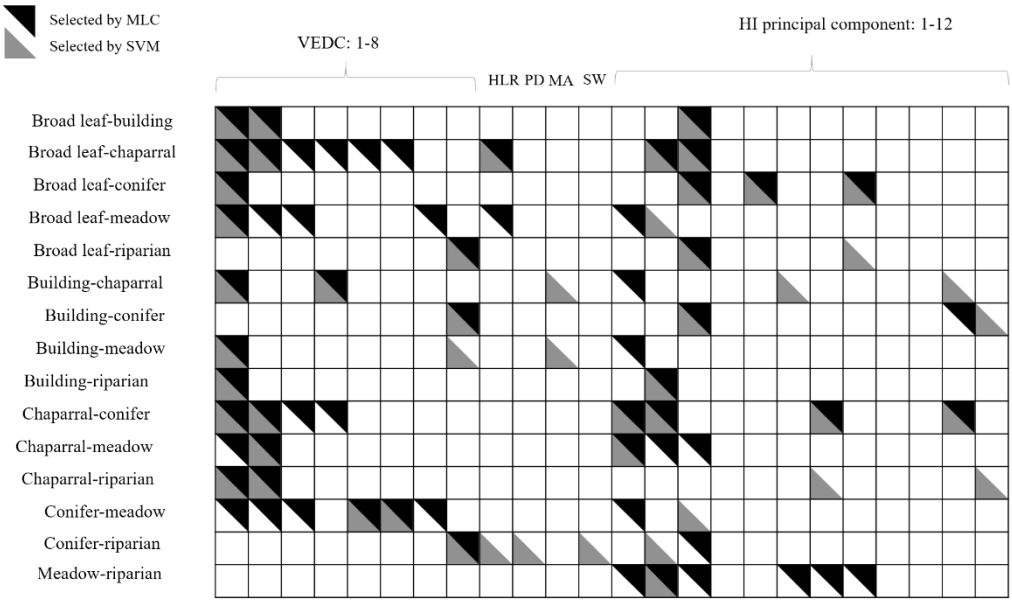


Table 7-10 SVM and MLC results for fused data set: (a) MLC results; (b) SVM results.

(a)

Predicted / ***Actual***	**Broad leaf**	**Building**	**Chaparral**	**Conifer**	**Meadow**	**Riparian**	**Total**
Broad leaf	3027	0	397	441	52	83	4000
Building	23	39	0	0	2	26	90
Chaparral	357	0	3364	279	0	0	4000
Conifer	481	0	205	3241	6	67	4000
Meadow	14	0	305	0	3611	70	4000
Riparian	87	28	10	46	38	3791	4000
Total	3989	67	4281	4007	3709	4037	0.850

(b)

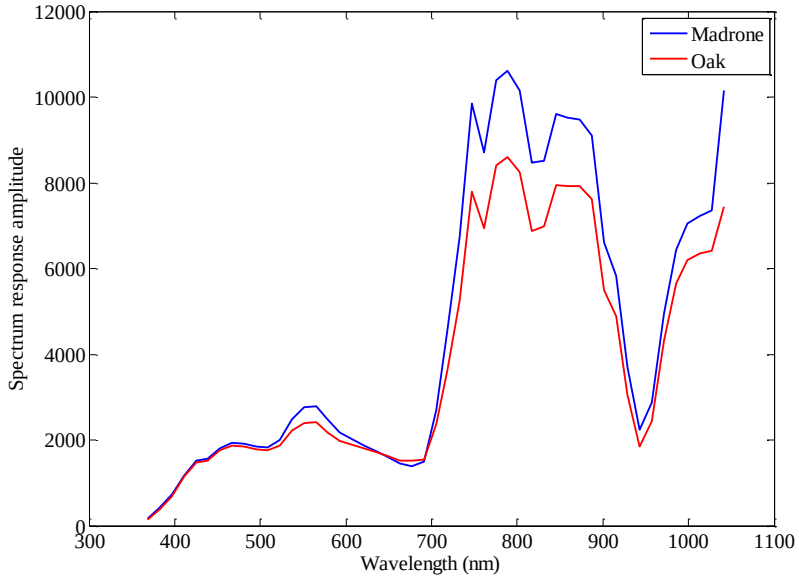
Predicted / *Actual*	**Broad leaf**	**Building**	**Chaparral**	**Conifer**	**Meadow**	**Riparian**	**Total**
Broad leaf	3838	0	50	94	17	1	4000
Building	1	75	0	0	4	10	90
Chaparral	88	0	3831	81	0	0	4000
Conifer	122	0	81	3789	0	8	4000
Meadow	4	0	0	1	3988	7	4000
Riparian	1	0	0	2	4	3993	4000
Total	4054	75	3962	3967	4013	4019	0.971

7.2.3.2 Vegetation Species Classification

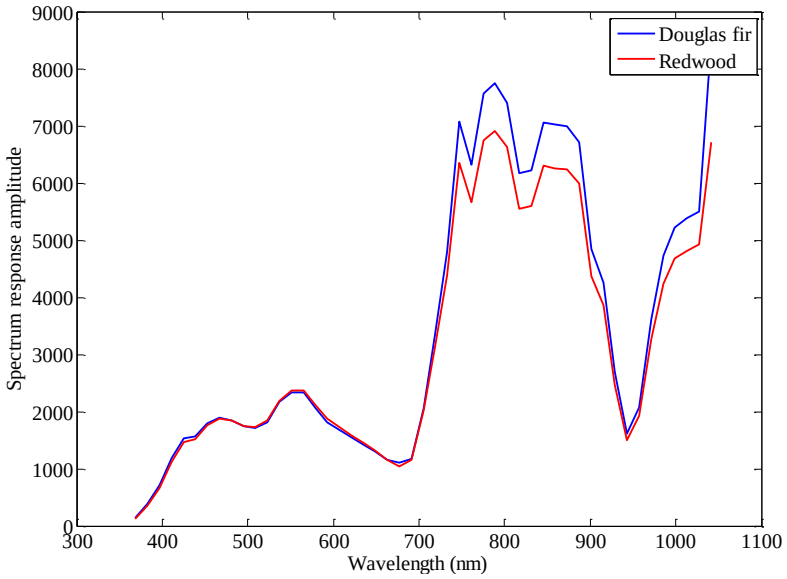
An overall SVM classification accuracy exceeding 97% was achieved using the proposed methods for separating the six land cover classes for Rivendell. To further evaluate the performance of the fused data set for separating objects with similar properties (spectral or structural), a deeper classification is performed where multiple trees species were identified within the broadleaf and conifer vegetation classes. Madrone and oak trees are discriminated from within the broadleaf class, and douglas-fir and redwood trees from within the conifer class.

The HI spectrum of the different tree specifies in a common class are quite similar, as shown in Figure 7-9, with distinguishable scale differences only in certain bands (25 – 50). In this situation, the VEDC feature may be useful if the different species within one general class have different canopy structure, as shown in Figure 7-10. The VEDC features are presented in Figure 7-11 where it is clear that the oak species (e.g., live oak,

black oak and tan oak) have a higher first VEDC value than madrone. Since the madrones found in this area are mostly shade-intolerant, they have more open canopy then the oaks (Rivendell training data set description, Bode 2016). The individual madrone tree crowns do not develop into a continuous layer, making the canopy more easily penetrated by laser light, this is visual illustrated in the photographs of the canopy shown in Figure 7-11 (a) and (b). These characteristics cause the value of first VEDC band to be higher and the last lower, for madrones as compared to oaks. Similarly for the conifer class, douglas-firs tend to have conical tops, and redwoods broader, umbrella shaped tops (Berrill et al., 2013) (Figure 7-11 (c) and (d)). The differences in shape are reflected in the VEDC profiles shown in Figure 7-10. These observations are also supported by the feature selection results shown in Table 7-11, where classifiers dealing with madrone versus oak and, douglas-fir versus redwood, select the first VEDC band, indicating that structural differences between these similar trees exist in the tree canopies.

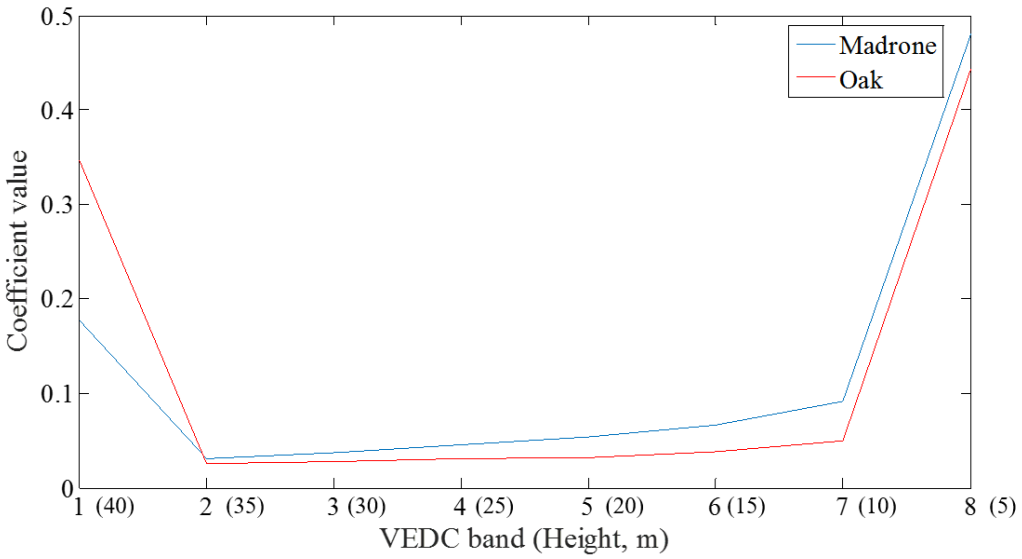


(a)

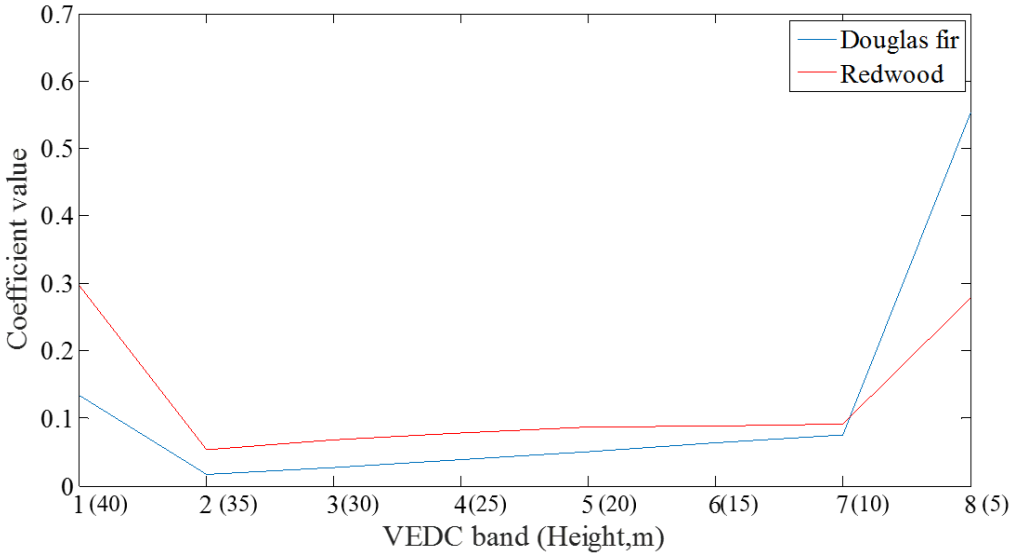


(b)

Figure 7-9 (a) Mean spectrum of madrone and oak (both broadleaf) from a training sample; (b) Mean spectrum of douglas-fir and redwood (both conifer) from a training sample



(a)

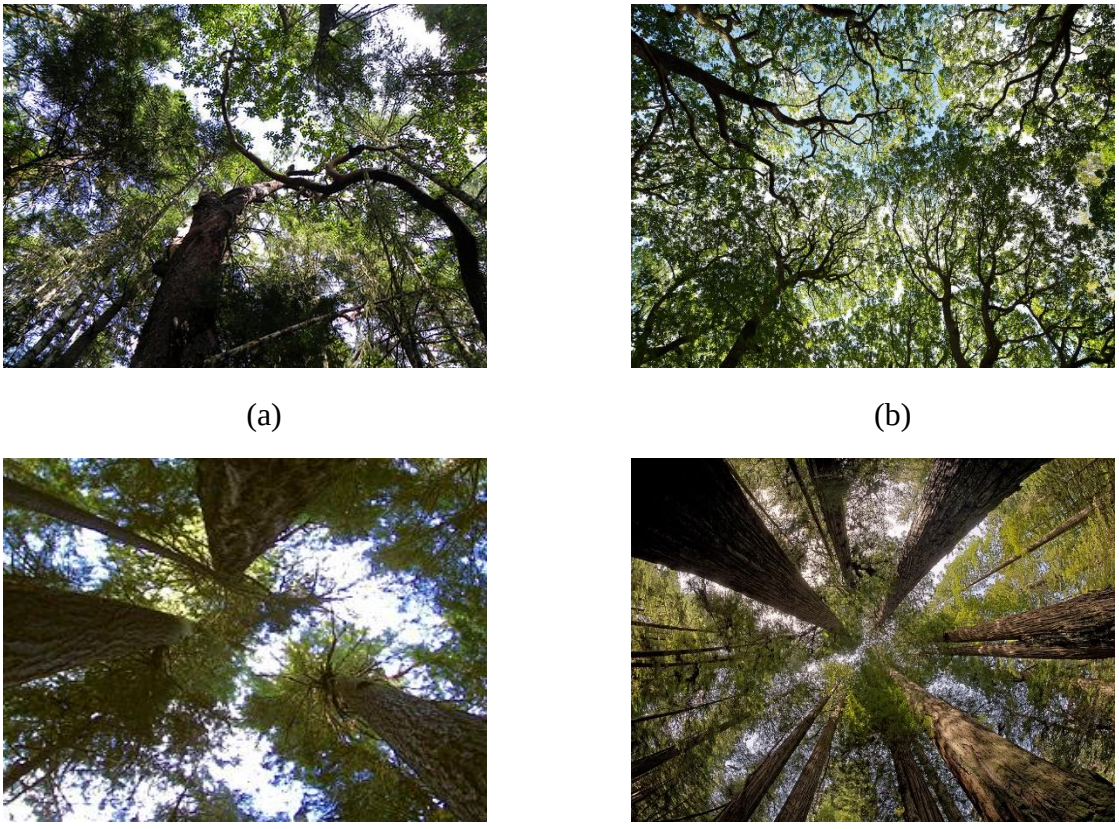


(b)

Figure 7-10 (a) Mean VEDC of madrone and oak (both broadleaf) from training sample; (b) Mean VEDC of douglas-fir and redwood (both conifer) from training sample

Since the broad leaf and conifer classes have been subdivided into more specific species classes, the number of land cover classes is increased from six to eight. The number of training samples for the madrone was 2000 (4000 validation samples) and 500 (1000 validation samples) for the oaks. The same training and validation sample sizes apply to the douglas-fir and redwood species, respectively. Classification results using the stand-alone SWF and HI data sets and the fused data set are shown in Table 7-12. With the six general classes separated into eight species, we can see that the classification accuracy improves for all cases even though the overall class number increases. The classification accuracy achieved with SWF (92.5%) and HI (93.3%) is very similar. However, an interesting observation is that SWF performs better for broad leaf trees, whereas HI is better at separating conifer trees. A possible explanation is that the redwoods may show a red tone as compared to the blue-green color of douglas-firs, and this difference may be captured by HI. Fusion of the SWF and HI provides both structural

and spectral information for the target classes, and yields a high overall accuracy of 98.6%, with no strong correlation between different classes evident. The overall accuracy of the eight class fused fata set is also 1.5% higher than the six class fused data set. This may indicate that treating different tree species as a single class increases the difficulty for a classifier to learn a common pattern within the data.



(a) (b)

(c) (d)

Figure 7-11 Canopy structure: (a) Madrone; (b) Oak; (c) Douglas-fir; (d) Redwood.

Table 7-11 Selected features for 8 classes

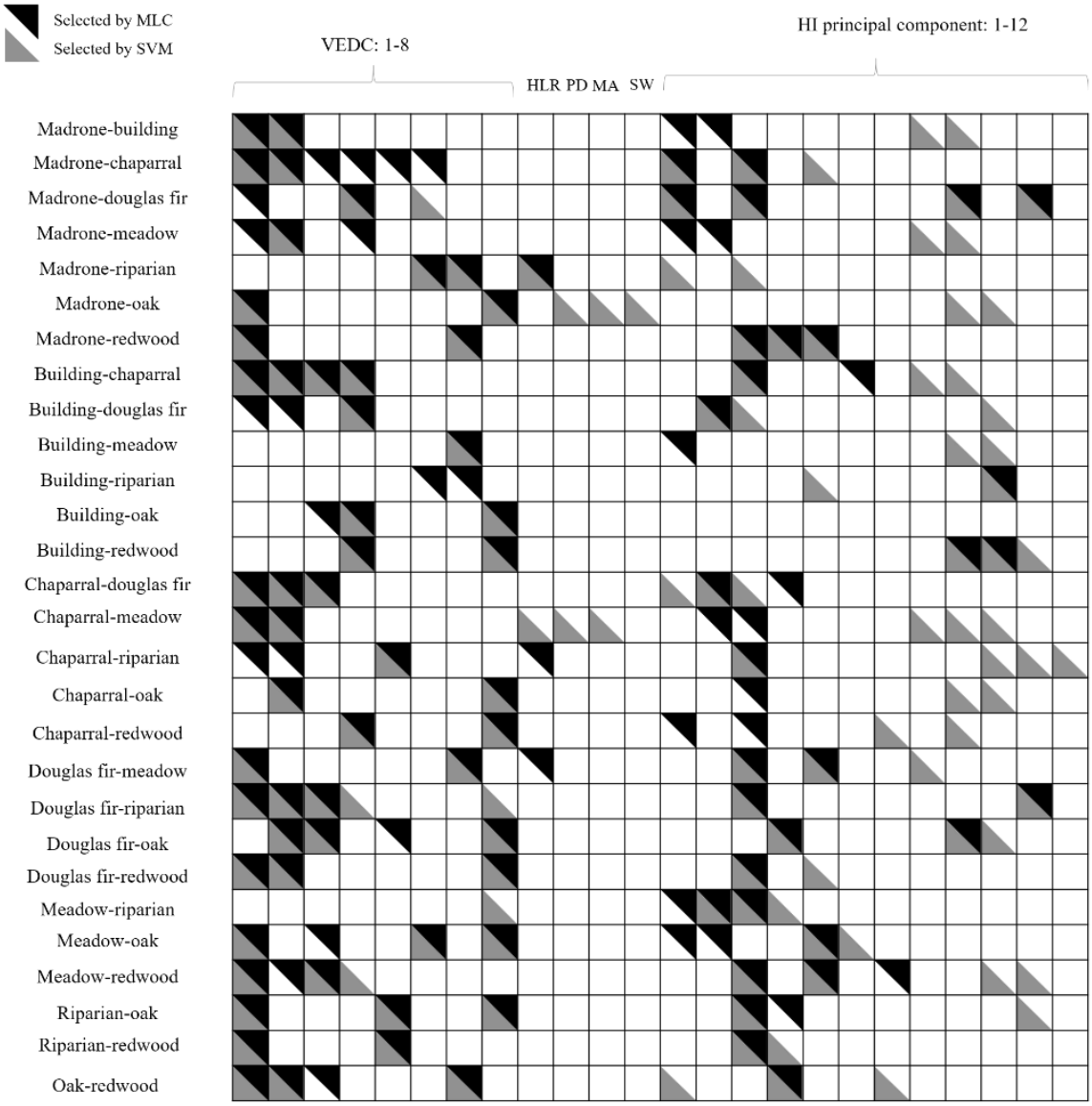


Table 7-12 (a) Confusion matrix for SWF classification using SVM; (b) Confusion matrix of HI classification using SVM; (c) Confusion matrix for fused data classification using SVM.

(a)

Predicted ***Actual***	**Broad leaf-Madrone**	**Building**	**Chaparral**	**Conifer-Douglas fir**	**Meadow**	**Riparian**	**Broad leaf-Oak**	**Conifer-Redwood**	**Total**
Broad leaf-Madrone	3836	0	8	71	0	0	0	85	4000
Building	0	68	7	0	0	15	0	0	90
Chaparral	17	11	3779	192	0	0	1	0	4000
Conifer-Douglas fir	151	3	381	3391	0	47	6	21	4000
Meadow	2	1	0	2	3850	145	0	0	4000
Riparian	2	25	0	82	155	3734	2	0	4000
Broad leaf-Oak	1	0	0	1	0	0	990	8	1000
Conifer-Redwood	142	0	1	34	0	0	42	781	1000
Total	4151	108	4176	3773	4005	3941	1040	896	0.925

(b)

Predicted ***Actual***	**Broad leaf-Madrone**	**Building**	**Chaparral**	**Conifer-Douglas fir**	**Meadow**	**Riparian**	**Broad leaf-Oak**	**Conifer-Redwood**	**Total**
Broad leaf-Madrone	3519	0	179	238	19	38	3	4	4000
Building	0	75	9	0	0	6	0	0	90
Chaparral	253	0	3664	76	1	5	1	0	4000
Conifer-Douglas fir	190	0	73	3633	0	71	8	25	4000
Meadow	2	0	6	0	3992	0	0	0	4000
Riparian	101	0	25	106	11	3757	0	0	4000
Broad leaf-Oak	4	0	5	6	4	0	981	0	1000
Conifer-Redwood	1	0	1	5	0	0	2	991	1000
Total	4070	75	3962	4064	4027	3877	995	1020	0.933

(c)

Predicted ***Actual***	**Broad leaf-Madrone**	**Building**	**Chaparral**	**Conifer-Douglas fir**	**Meadow**	**Riparian**	**Broad leaf-Oak**	**Conifer-Redwood**	**Total**
Broad leaf-Madrone	3891	0	18	89	0	0	0	2	4000
Building	0	79	0	2	0	9	0	0	90
Chaparral	7	0	3890	103	0	0	0	0	4000
Conifer-Douglas fir	23	0	44	3931	0	0	0	2	4000
Meadow	2	0	0	0	3998	0	0	0	4000
Riparian	1	0	0	0	4	3995	0	0	4000
Broad leaf-Oak	0	0	0	0	0	0	1000	0	1000
Conifer-Redwood	4	0	0	4	0	0	1	991	1000
Total	3928	79	3952	4129	4002	4004	1001	995	0.986

With the proposed methodology, a vegetation species and ground cover classification map can be obtained, as shown in Figure 7-12. These results indicate that the improved classification enabled by the proposed fusion methods could be used for voxel-sized natural resources evaluation applications.

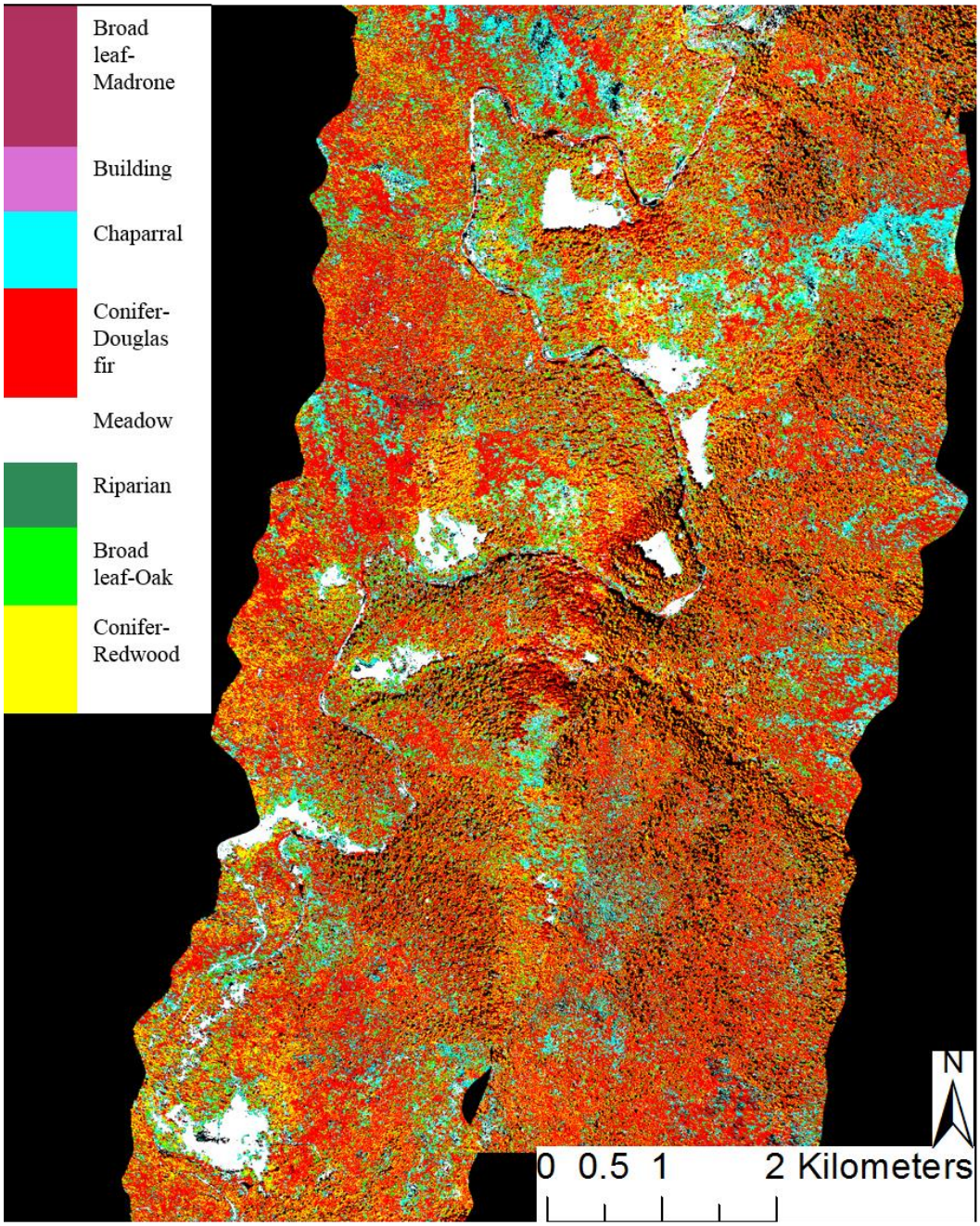


Figure 7-12 Rivendell data set classification map

Chapter 8 Conclusions and Future Work

8.1 Summary and Conclusions

In this dissertation, we investigated methods for the voxelization of FWL, fusion of FWL and HI, and then tested classification of a fused feature set on two representative data sets. The proposed methods were also compared to existing fusion methods found in the current literature. The proposed voxelization and fusion method synthesizes vertical waveforms, which are co-registered with the HI in a regular 3D grid, from the intersection of the original, slanting waveforms with the voxel columns using a maximum amplitude method. Novel VEDC and standard features were then extracted from the SWFs and combined with the HI for land cover classification and tree species identification tasks employing MLC and SVM classifiers using a the pairwise classification framework.

A novel waveform feature descriptor, VEDC, was proposed to extract information from the synthesized waveforms (i.e., SWFs) for use in subsequent classification algorithms. The VEDC feature provides information about the distribution of backscattered energy in the vertical direction within each column of voxels. Since VEDC is a data-dependent feature, parameters defining the dimension and span of VEDC are required for each data set. Both supervised and unsupervised VEDC parameter approaches were proposed and evaluated, with the unsupervised VEDC parameters producing slightly better classification performance when used in combination with a SVM classifier. This is a desirable results since the unsupervised method is generic and, by definition, requires much less user interaction. The discrimination performance of the

VEDC features was examined in the SFFS feature selection step, which is performed prior to the classification, and found to be effective. For most pairs of classes, VEDC was an effective feature for discrimination. The HI features were selected by keeping principal components from PCA that represents more than 99% of the HI imagery variance. The joint feature space was then used for classification.

The classification method used in this dissertation was adapted from a pairwise classifier framework, which enabled unique feature selection for each pair of classes, as well as the use of multiple classifiers. Two classifiers, MLC and SVM, were applied to simultaneously acquired FWL and HI data sets. The proposed voxelization, fusion and classification methods were first applied to the Moran data set with nine ground cover types containing both natural and man-made objects. The analysis led to the following conclusions:

1) Unsupervised VEDC shows similar or better performance compared with supervised VEDC when used in combination with different classifiers.

2) SWF features show promising potential for classification, with an overall classification accuracy of nearly 79.8% when using SWF features only in a pairwise MLC classifier. On the other hand, using only discrete LPC features was far less reliable for classification, with an accuracy of only 61.9%, for the same scheme.

3) Feature level fusion of SWF and HI improved the classification accuracy compared to using either feature set alone. Using fused SWF and HI with a

pairwise SVM classifier outperformed standalone HI - SVM classification by 9.4%.

4) SVM-based pairwise classifiers outperformed MLC-based pairwise classifiers in overall user accuracy and kappa coefficient values in seven of the nine individual classes. This is due to the nonlinear learning capacity of SVM using kernel functions. In addition, performance of SVM was less variable among classes, which was also an important advantage over MLC.

The Rivendell data set was introduced to test the performance of the proposed methods, particularly the value of the VEDC features, to separate different ground cover types having similar heights and spectral characteristics but differing vertical structure. The SVM classification of fused unsupervised VEDC and HI features shows significant improvement over standalone classification (7.4 percent better than HI and 12.0 percent better than SWF in overall accuracy). To further evaluate the power of fusing VEDC's structural information with HI spectral information, the general tree classes (broadleaf and conifer) were broken down into the individual species of madrone and oak from the broadleaf class and douglas-fir and redwood from the conifer class. Classification using SWFs only (92.5%) has a comparable performance with HI only (93.3%), but the fused data set is able to separate the tree species with an accuracy of 98.6%. These results indicate that the VEDC features are complementary to HI information for discriminating spectrally similar tree classes in complex forested areas by taking advantage of the ability of the VEDC features to encapsulate vertical structural differences that exist between the tree classes.

8.2 Future Work

This dissertation has proposed a complete framework for feature extraction from FWL and HI data sets, and the use of fused FWL and HI features for land cover classification. However, some subjects could benefit from further investigation. Even with the same data source, land cover classification applications may have differing focuses, e.g., man-made objects classification and extraction for urban planning, versus tree type characterization for forestry applications. To adapt the proposed methods for different applications, the process can be applied with different combinations of algorithms employed in each step, and even with different fusion levels of data sets.

8.2.1 Using Proposed Methods for Different Fusion Levels

In this dissertation, the fusion of FWL and HI is achieved on a data level since several similar object classes are broadly distributed in both data sets. Fusion of FWL and HI can also be applied at different levels, e.g., the object level or the decision level. Object level fusion has been discussed in the literature for discrete return LiDAR and HI data sets (Chen & Gao, 2014; Man et al., 2015), where LiDAR elevation and intensity information were used for image segmentation prior to fusion with HI for classification. The segmentation is usually achieved by use of elevation and intensity information of objects with relatively large contiguous area, e.g., building and parking lot (Man et al., 2015). With the inclusion of VEDC features derived from FWL data, the discrimination based on energy distribution at any layer of an object would be possible, thus achieving segmentation with less knowledge required from the user and potentially better accuracy.

8.2.2 Further Investigation into Implementation of the Pairwise Classifier

Three components determine the methodology for implementing a pairwise classifier: the individual classifier used, the feature selection tool, and the voting strategy. A broad variety of classification algorithms can be used under the framework of pairwise classification. For example, a boosted decision tree classifier (Nourzad & Pradhan, 2012) has the potential for use with fused FWL and HI features because the decision tree-based method does not impose strict requirements on the input features, and also works well with multiple-source input data.

The feature selection method used in this thesis is a wrapper based SFFS method, where the feature selection target was set to best accuracy for each individual classifier. The searching strategy of the wrapper method includes sequential, exhaustive, and random searching (Kohavi & John, 1997; Liu & Yu, 2005). A sequential method was used in this dissertation. An exhaustive method searches for all possible combinations, and while accurate is not practical for use on any large feature sets. Random searching starts from a randomly chose features set and randomly inserts of deletes features. Both exhaustive and random searching were not evaluated in this dissertation, but the use of these two search methods are worth investigating. For example exhaustive searching is the only method that looks for a global optimal feature subset, and therefore could be used for class pair that are difficult to separate. Within the framework of pairwise classification, different feature selection methods can also be combined.

The voting strategy is also crucial for successful pairwise classification because it directly affects the final output of the whole system. Based on different kinds of classifier outputs, the voting strategy can be applied to either class label, class ranking or soft

output, i.e., a value representing the likelihood that one sample falls in a particular class, for any classifier. This dissertation utilized as voting strategy for class label because different type of classifiers were used and the labeling didn't to be normalized between different classifiers. However, a voting strategy using soft output is still worth investigating because many classification algorithm1s yield s oft output including, e.g., MLC and SVM. The voting strategy for soft output aims to reduce the uncertainty in the final labeling. For example, Bayesian classifier fusion methods could be applied on classifiers whose output is posterior possibilities. Choosing an appropriate voting strategy gives the system extra flexibility to adapt different classifiers and application requirements.

8.2.3 Hyperspectral FWL

Several prototype hyperspectral FWL systems have been demonstrated in laboratory environments, where a laser with a broad spectrum range is employed as the LiDAR light source. This enables both target range and spectral response to be simultaneously measured, with the collected spectrum shape reported to be similar to passive HI (Hakala et al., 2012; Li et al., 2014). Assuming the future availability of airborne hyperspectral FWL data, the waveform voxelization and synthesizing approaches proposed in this dissertation would enable the ability to observe a vertical target spectrum, thereby supporting the extraction of target spectral features at different elevations within targets with complex vertical structures at the fine spatial resolution provided by digitizing FWL systems. Such systems hold great potential for use in land cover classification.

References

Abed, F.M., Mills, J.P. & Miller, P.E., 2014. Calibrated Full-Waveform Airborne Laser Scanning for 3D Object Segmentation. *Remote Sensing*, 6(5), pp.4109–4132.

Adams, T., Beets, P. & Parrish, C., 2012. Extracting More Data from LiDAR in Forested Areas by Analyzing Waveform Shape. *Remote Sensing*, 4(12), pp.682–702.

Aggarwal, C.C., 2015. *Data classification : algorithms and applications*,

Alonzo, M., Bookhagen, B. & Roberts, D.A., 2014. Urban Tree Species Mapping Using Hyperspectral and Lidar Data Fusion. *Remote Sensing of Environment*, 148, pp.70–83.

Anderson, J.E., Plourde, L.C., Martin, M.E., Braswell, B.H., Smith, M.L., Dubayah, R.O., Hofton, M. a. & Blair, J.B., 2008. Integrating Waveform Lidar with Hyperspectral Imagery for Inventory of a Northern Temperate Forest. *Remote Sensing of Environment*, 112(4), pp.1856–1870.

Anderson, K., Hancock, S., Disney, M. & Gaston, K.J., 2016. Is Waveform Worth It? A Comparison of LiDAR Approaches for Vegetation and Landscape Characterization D. Rocchini & D. Boyd, eds. *Remote Sensing in Ecology and Conservation*, 2(1), pp.5–15.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. & Field, C.B., 2007. Carnegie Airborne Observatory: In-Flight Fusion of Hyperspectral Imaging and Waveform Light Detection and Ranging for Three-Dimensional Studies of Ecosystems. *Journal of Applied Remote Sensing*, 1(1), p.013536.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. & Hughes, R.F., 2008. Invasive Species Detection in Hawaiian Rainforests Using Airborne Imaging Spectroscopy and LiDAR. *Remote Sensing of Environment*, 112(5), pp.1942–1955.

Baofeng, G., S.R.Gunn, Damper, R.I. & Nelson, J.D.B., 2008. Customizing Kernel Functions for SVM Based Hyperspectral Image Classification. *IEEE Trans. Im. Proc*, 17(4), pp.622–629.

Baraldi, a. & Parmiggiani, F., 1995. A Neural Network for Unsupervised Categorization of Multivalued\ninput Patterns: An Application to Satellite Image Clustering. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2).

Bateson, A. & Curtiss, B., 1996. A Method for Manual Endmember Selection and Spectral Unmixing. *Remote Sensing of Environment*, 55(3), pp.229–243.

Berrill, J.P., Beal, C.B., LaFever, D.H. & Dagley, C.M., 2013. Modeling Young Stand Development towards the Old-Growth Reference Condition in Evergreen Mixed-Conifer Stands at Headwaters Forest Reserve, California. *Forests*, 4(2), pp.455–470.

Bioucas-Dias, J.M. & Nascimento, J.M.P., 2008. Hyperspectral Subspace Identification. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8), pp.2435–2445.

Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P. & Chanussot, J., 2012. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2), pp.354–379.

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T. & Solberg, S., 2010. Prediction of

Species Specific Forest Inventory Attributes Using a Nonparametric Semi-Individual Tree Crown Approach Based on Fused Airborne Laser Scanning and Multispectral Data. *Remote Sensing of Environment*, 114(4), pp.911–924.

Buddenbaum, H., Seeling, S. & Hill, J., 2013. Fusion of Full-Waveform Lidar and Imaging Spectroscopy Remote Sensing Data for the Characterization of Forest Stands. *International Journal of Remote Sensing*, 34(13), pp.4511–4524.

Burges, C.J.C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2, pp.121–167.

Camps-Valls, G., Bandos Marsheva, T. V. & Zhou, D., 2007. Semi-Supervised Graph-Based Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), pp.3044–3054.

Camps-Valls, G. & Bruzzone, L., 2005. Kernel-Based Methods for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6), pp.1351–1362.

Canters, F., 1997. Evaluating the Uncertainty of Area Estimates Derived from Fuzzy Land-Cover Classification. *Photogrammetric Engineering & Remote Sensing*, 63(4), pp.403–414.

Chan, J.C.W. & Paelinckx, D., 2008. Evaluation of Random Forest and Adaboost Tree-Based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery. *Remote Sensing of Environment*, 112(6), pp.2999–3011.

Chang, Y., Hsieh, C.-J., Chang, K.-W., Ringgaard, M. & Lin, C., 2010. Training and

Testing Low-Degree Polynomial Data Mappings via Linear SVM. *The Journal of Machine …*, 11, pp.1471–1490.

Chen, Z. & Gao, B., 2014. An Object-Based Method for Urban Land Cover Classification Using Airborne Lidar Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10), pp.4243–4254.

Cibula, W.G. & Nyquist, M.O., 1987. Use of Topographic and Climatological Models in a Geographical Data Base to Improve Landsat Mss Classification for Olympic National Park. *Photogrammetric Engineering and Remote Sensing*, 53(1), pp.67–75.

Clint Slatton, K., Crawford, M.M. & Evans, B.L., 2001. Fusing Interferometric Radar and Laser Altimeter Data to Estimate Surface Topography and Vegetation Heights. *IEEE Transactions on Geoscience and Remote Sensing*, 39(11), pp.2470–2482.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), pp.37–46.

Congalton, R.G., 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment*, 37(1), pp.35–46.

Craig, D.M., Ulaby, F.T. & Pierce, L.E., 1995. Land-Cover Classification and Estimation of Terrain Attributes Using Synthetic Aperture Radar. *Remote Sensing of Environment*, 51(1), pp.199–214.

Crawford, M.M., 1999. Fusion of Airborne Polarimetric and Interferometric SAR for Classification of Coastal Environments. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3 I), pp.1306–1315.

Dalponte, M., Bruzzone, L. & Gianelle, D., 2012. Tree Species Classification in the Southern Alps Based on the Fusion of Very High Geometrical Resolution Multispectral/hyperspectral Images and LiDAR Data. *Remote Sensing of Environment*, 123, pp.258–270.

Dalponte, M., Member, S., Bruzzone, L., Member, S. & Gianelle, D., 2008. Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. , 46(5), pp.1416–1427.

Debes, C. et al., 2014. Hyperspectral and LiDAR Data Fusion: Outcome of the 2013 GRSS Data Fusion Contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), pp.2405–2418.

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M. a., Chazdon, R.L., Weishampel, J.F. & Prince, S., 2002. Estimation of Tropical Forest Structural Characteristics, Using Large-Footprint Lidar. *Remote Sensing of Environment*, 79(2-3), pp.305–319.

Du, Q. & Chang, C.-I., 2001. A Linear Constrained Distance-Based Discriminant Analysis for Hyperspectral Image Classification. *Pattern Recognition*, 34(2), pp.361–373.

Duda, T. & Canty, M., 2002. Unsupervised Classification of Satellite Imagery: Choosing a Good Algorithm. *International Journal of Remote Sensing*, 23(11), pp.2193–2212.

Erdody, T.L. & Moskal, L.M., 2010. Fusion of LiDAR and Imagery for Estimating Forest Canopy Fuels. *Remote Sensing of Environment*, 114(4), pp.725–737.

Foody, G.M., 2002. Status of Land Cover Classification Accuracy Assessment. *Remote*

Sensing of Environment, 80(1), pp.185–201.

Forzieri, G., Guarnieri, L., Vivoni, E.R., Castelli, F. & Preti, F., 2010. Spectral-ALS Data Fusion for Different Roughness Parameterizations of Forested Floodplains. *River Research and Applications*, 7(4), p.n/a–n/a.

Franklin, S.E. & Peddle, D.R., 1990. Classification of SPOT HRV Imagery and Texture Features. *International Journal of Remote Sensing*, 11(3), pp.551–556.

Freund, Y. & Schapire, R.E., 1996. Experiments with a New Boosting Algorithm. *Proc. ICML*, pp.148–156.

Friedl, M. a. M.A. & Brodley, C.E.C.E., 1997. Decision Tree Classification of Land Cover from Remotely Sensed Data. *Remote Sensing of Environment*, 61(3), pp.399–409.

Friedman, J., 1996. Another Approach To Polychotomous Classification. *Techinal report, Department of Statistics, Stanford University*.

Funk, C.C., Theiler, J., Roberts, D. a. & Borel, C.C., 2001. Clustering to Improve Matched Filter Detection of Weak Gas Plumes in Hyperspectral Thermal Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), pp.1410–1420.

Gao, B.C., Heidebrecht, K.B. & Goetz, A.F.H., 1993. Derivation of Scaled Surface Reflectances from AVIRIS Data. *Remote Sensing of Environment*, 44(2-3), pp.165–178.

Gat, N., 2000. Imaging Spectroscopy Using Tunable Filters : A Review. , 4056, pp.50–64.

Geerling, G.W., Labrador- Garcia, M., Clevers, J.G.P.W., Ragas, a. M.J. & Smits, a.

J.M., 2007. Classification of Floodplain Vegetation by Data Fusion of Spectral (CASI) and LiDAR Data. *International Journal of Remote Sensing*, 28(19), pp.4263–4284.

Gislason, P.O., Benediktsson, J.A. & Sveinsson, J.R., 2006. Random Forests for Land Cover Classification. *Pattern Recognition Letters*, 27(4), pp.294–300.

Glennie, C., 2007. Rigorous 3D Error Analysis of Kinematic Scanning LIDAR Systems. *Journal of Applied Geodesy*, 1(3), pp.147–157.

Glennie, C.L., Carter, W.E., Shrestha, R.L. & Dietrich, W.E., 2013. Geodetic Imaging with Airborne LiDAR: The Earth's Surface Revealed. *Reports on progress in physics. Physical Society (Great Britain)*, 76(8), p.086801.

Goel, P.K., Prasher, S.O., Patel, R.M., Landry, J.A., Bonnell, R.B. & Viau, A.A., 2003. Classification of Hyperspectral Data by Decision Trees and Artificial Neural Networks to Identify Weed Stress and Nitrogen Status of Corn. *Computers and Electronics in Agriculture*, 39(2), pp.67–93.

Goetz, A.F., Vane, G., Solomon, J.E. & Rock, B.N., 1985. Imaging Spectrometry for Earth Remote Sensing. *Science (New York, N.Y.)*, 228(4704), pp.1147–1153.

Gowda, K.C., 1984. A Feature Reduction and Unsupervised Classification Algorithm for Multispectral Data. *Pattern Recognition*, 17(6), pp.667–676.

Green, A. a., Berman, M., Switzer, P. & Craig, M.D., 1988. Transformation for Ordering Multispectral Data in Terms of Image Quality With Implications for Noise Removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1), pp.65–74.

Gualtieri, J. a & Cromp, R.F., 1998. Support Vector Machines for Hyperspectral Remote Sensing Classification. *Spie 3584*, 3584, pp.221–232.

Guanter, L., Estellés, V. & Moreno, J., 2007. Spectral Calibration and Atmospheric Correction of Ultra-Fine Spectral and Spatial Resolution Remote Sensing Data. Application to CASI-1500 Data. *Remote Sensing of Environment*, 109(1), pp.54–65.

Guo, L., Chehata, N., Mallet, C. & Boukir, S., 2011. Relevance of Airborne Lidar and Multispectral Image Data for Urban Scene Classification Using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(1), pp.56–66.

Hakala, T., Suomalainen, J. & Kaasalainen, S., 2012. Full Waveform Active Hyperspectral LIDAR. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B7(September), pp.459–462.

Ham, J., Chen, Y., Crawford, M.M. & Ghosh, J., 2005. Investigation of the Random Forest Framework for Classification of Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3), pp.492–501.

Hara, Y., Atkins, R.G., Yueh, S.H., Shin, R.T. & Kong, J.A., 1994. Application of Neural Networks to Radar Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 32(1), pp.100–109.

Harsanyi, J.C. & Chang, C.I., 1994. Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 32(4), pp.779–785.

Hastie, T. & Tibshirani, R., 1998. Classification by Pairwise Coupling. *Annals of*

Statistics, 26(2), pp.451–471.

Heinzel, J. & Koch, B., 2012. Investigating Multiple Data Sources for Tree Species Classification in Temperate Forest and Use for Single Tree Delineation. *International Journal of Applied Earth Observation and Geoinformation*, 18, pp.101–110.

Hermosilla, T., Coops, N.C., Ruiz, L. a. & Moskal, L.M., 2014. Deriving Pseudo-Vertical Waveforms from Small-Footprint Full-Waveform LiDAR Data. *Remote Sensing Letters*, 5(4), pp.332–341.

Herries, G., Selige, T. & Danaher, S., 1996. Singular Value Decomposition in Applied Remote Sensing. *Image Processing for Remote Sensing, IEE Colloquium on*, pp.5/1 – 5/6.

Hofton, M.A., Minster, J.B. & Blair, J.B., 2000. Decomposition of Laser Altimeter Waveforms. , 38(4), pp.1989–1996.

Hollaus, M. & Höfle, B., 2010. Terrain Roughness Parameters from Fullwaveform Airborne LiDAR Data. *... Archives of Photogrammetry, Remote Sensing and ...*, XXXVIII, pp.287–292.

Hovi, A., Korhonen, L., Vauhkonen, J. & Korpela, I., 2016. LiDAR Waveform Features for Tree Species Classification and Their Sensitivity to Tree- and Acquisition Related Parameters. *Remote Sensing of Environment*, 173, pp.224–237.

Huang, C., Davis, L.S. & Townshend, J.R.G., 2002. An Assessment of Support Vector Machines for Land Cover Classification. *International Journal of Remote Sensing*, 23(4), pp.725–749.

Huang, X. & Zhang, L., 2013. An SVM Ensemble Approach Combining Spectral, Structural, and Semantic Features for the Classification of High-Resolution Remotely Sensed Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1), pp.257–272.

Jensen, J.R., Hodgson, M.E., Mackey, H.E. & Krabill, W., 1987. Correlation between Aircraft MSS and LIDAR Remotely Sensed Data on a Forested Wetland. *Geocarto International*, 2(4), pp.39–54.

Jia, X. & Richards, J.A., 1994. Efficient Maximum Likelihood Classification for Imaging Spectrometer Data Sets. *IEEE Transactions on Geoscience and Remote Sensing*, 32(2), pp.274–281.

Jia, X. & Richards, J.A., 1999. Segmented Principal Components Transformation for Efficient Hyperspectral Remote-Sensing Image Display and Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1 II), pp.538–542.

Jung, J., 2011. *Utilization of Full Waveform Lidar and Hyperspectral Data for Forest Structure Characterization*. Purdue University.

Jung, J., Pasolli, E., Prasad, S., Tilton, J.C. & Crawford, M.M., 2014. A Framework for Land Cover Classification Using Discrete Return LiDAR Data: Adopting Pseudo-Waveform and Hierarchical Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(2), pp.491–502.

Jutzi, B. & Stilla, U., 2006. Range Determination with Waveform Recording Laser Systems Using a Wiener Filter. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(2), pp.95–107.

Kawaguchi, S. & Nishii, R., 2007. Hyperspectral Image Classification by Bootstrap AdaBoost with Random Decision Stumps. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11), pp.3845–3851.

Kempeneers, P., Deronde, B., Provoost, S. & Houthuys, R., 2009. Synergy of Airborne Digital Camera and Lidar Data to Map Coastal Dune Vegetation. *Journal of Coastal Research*, 10053(10053), pp.73–82.

Kerekes, J.P. & Baum, J.E., 2002. Spectral Imaging System Analytical Model for Subpixel Object Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 40(5), pp.1088–1101.

Keshava, N. & Mustard, J.F., 2002. Spectral Unmixing. *IEEE Signal Processing Magazine*, 19(1), pp.44–57.

Kirchhof, M., Jutzi, B. & Stilla, U., 2008. Iterative Processing of Laser Scanning Data by Full Waveform Analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1), pp.99–114.

Koetz, B., Morsdorf, F., van der Linden, S., Curt, T. & Allgöwer, B., 2008. Multi-Source Land Cover Classification for Forest Fire Management Based on Imaging Spectrometry and LiDAR Data. *Forest Ecology & Management*, 256(3), pp.263–271.

Kohavi, R. & John, G.H., 1997. Wrappers for Feature Subset Selection. *Artificial Intelligence*, 97(1-2), pp.273–324.

Korhonen, L., Korpela, I., Heiskanen, J. & Maltamo, M., 2011. Airborne Discrete-Return LIDAR Data in the Estimation of Vertical Canopy Cover, Angular Canopy Closure

and Leaf Area Index. *Remote Sensing of Environment*, 115(4), pp.1065–1080.

Kotsiantis, S., 2007. Supervised Machine Learning: A Review of Classification Techniques. *Informatica*, 31(3), pp.249–268.

Kulawardhana, R.W., Popescu, S.C. & Feagin, R. a., 2014. Fusion of Lidar and Multispectral Data to Quantify Salt Marsh Carbon Stocks. *Remote Sensing of Environment*, 154, pp.345–357.

Kumar, S., Ghosh, J. & Crawford, M.M., 2001. Best-Bases Feature Extraction Algorithms for Classification of Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), pp.1368–1379.

Kumar, S., Ghosh, J. & Crawford, M.M., 2002. Hierarchical Fusion of Multiple Classifiers for Hyperspectral Data Analysis. *Pattern Analysis & Applications*, 5(2), pp.210–220.

Kuplich, T.M., Freitas, C.C. & Soares, J. V, 2000. The Study of ERS-1 SAR and Landsat TM Synergism for Land Use Classification. *International Journal of Remote Sensing*, 21(10), pp.2101–2111.

Kwon, H. & Nasrabadi, N.M., 2005. Kernel Orthogonal Subspace Projection for Hyperspectral Signal Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(12), pp.2952–2962.

Lance, G.N. & Williams, W.T., 1966. Computer Programs for Hierarchical Polythetic Classification (“Similarity Analyses”). *The Computer Journal*, 9(1), pp.60–64.

Lawrence, K.C.K.C., Park, B., Windham, W.R. & Mao, C., 2003. Calibration of a

Pushbroom Hyperspectral Imaging System for Agricultural Inspection. *Transactions of the American Society of Agricultural Engineers*, 46(2), pp.513–521.

Lefsky, M. a., Keller, M., Pang, Y., De Camargo, P.B. & Hunter, M.O., 2007. Revised Method for Forest Canopy Height Estimation from Geoscience Laser Altimeter System Waveforms. *Journal of Applied Remote Sensing*, 1(1), p.013537.

Lennon, M., Mercier, G., Mouchot, M.C. & Hubert-Moy, L., 2001. Independent Component Analysis as a Tool for the Dimensionality Reduction and the Representation of Hyperspectral Images. In *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)*. IEEE, pp. 2893–2895.

Li, W., Sun, G., Niu, Z., Gao, S. & Qiao, H., 2014. Estimation of Leaf Biochemical Content Using a Novel Hyperspectral Full-Waveform LiDAR System. *Remote Sensing Letters*, 5(8), pp.693–702.

Liew, S.C.L.S.C., Chang, C.W.C.C.W. & Lim, K.H.L.K.H., 2002. Hyperspectral Land Cover Classification of EO-1 Hyperion Data by Principal Component Analysis and Pixel Unmixing. *IEEE International Geoscience and Remote Sensing Symposium*, 6(119260), pp.3111–3113.

Liu, C., Frazier, P. & Kumar, L., 2007. Comparative Assessment of the Measures of Thematic Classification Accuracy. *Remote Sensing of Environment*, 107(4), pp.606–616.

Liu, H. & Yu, L., 2005. Toward Integrating Feature Selection Algorithms for Classification and Clustering. *Ieee Transactions on Knowledge and Data*

Engineeering, 17, pp.491–502.

Lu, D. & Weng, Q., 2007. A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing*, 28(5), pp.823–870.

Lunga, D., Prasad, S., Crawford, M.M. & Ersoy, O., 2014. Manifold-Learning-Based Feature Extraction for Classification of Hyperspectral Data: A Review of Advances in Manifold Learning. *IEEE Signal Processing Magazine*, 31(1), pp.55–66.

Ma, L., Crawford, M.M. & Tian, J., 2010. Local Manifold Learning-Based -Nearest-Neighbor for Hyperspectral Image Classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(11), pp.4099–4109.

MacQueen, J.B., 1967. Kmeans Some Methods for Classification and Analysis of Multivariate Observations. *5th Berkeley Symposium on Mathematical Statistics and Probability 1967*, 1(233), pp.281–297.

Mallet, C. & Bretar, F., 2009. Full-Waveform Topographic Lidar: State-of-the-Art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), pp.1–16.

Mallet, C., Bretar, F., Roux, M., Soergel, U. & Heipke, C., 2011. Relevance Assessment of Full-Waveform Lidar Data for Urban Area Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6 SUPPL.).

Man, Q., Dong, P. & Guo, H., 2015. Pixel- and Feature-Level Fusion of Hyperspectral and Lidar Data for Urban Land-Use Classification. *International Journal of Remote Sensing*, 36(March), pp.1618–1644.

Manolakis, D., Marden, D. & Shaw, G. a, 2003. Hyperspectral Image Processing for Automatic Target Detection Applications. *Lincoln Laboratory Journal*, 14(1), pp.79–116.

Mason, D.C., Horritt, M.S., Hunter, N.M. & Bates, P.D., 2007. Use of Fused Airborne Scanning Laser Altimetry and Digital Map Data for Urban Flood Modelling. *Hydrological Processes*, 21(11), pp.1436–1447.

Matthew, M.W., Adler-Golden, S.M., Berk, A., Felde, G., Anderson, G.P., Gorodetzky, D., Paswaters, S. & Shippert, M., 2002. Atmospheric Correction of Spectral Imagery: Evaluation of the FLAASH Algorithm with AVIRIS Data. *Proceedings - Applied Imagery Pattern Recognition Workshop*, 2002-Janua, pp.157–163.

McCombs, J.W., Roberts, S.D. & Evans, D.L., 2003. Influence of Fusing Lidar and Multispectral Imagery on Remotely Sensed Estimates of Stand Density and Mean Tree Height in a Managed Loblolly Pine Plantation. *Forest Science*, 49(3), pp.457–466.

Melgani, F. & Bruzzone, L., 2004. Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), pp.1778–1790.

Miller, C.J., 2001. Fusion of High Resolution LIDAR Elevation Data with Hyperspectral Data to Characterize Tree Canopies. *Algorithms for Multispectral, Hyperspectral and Ultraspectral Imagery Vii*, 4381, pp.246–252.

Mouroulis, P. & Green, R.O., 2003. Optical Design for High-Fidelity Imaging Spectrometry. *Proceedings of SPIE - The International Society for Optical*

Engineering, 4829 II, p.1048.

Mundt, J.T., Streutker, D.R. & Glenn, N.F., 2006. Mapping Sagebrush Distribution Using Fusion of Hyperspectral and Lidar Classifications. *Photogrammetric Engineering Remote Sensing*, 72(1), pp.47–54.

Muss, J.D., Mladenoff, D.J. & Townsend, P. a., 2011. A Pseudo-Waveform Technique to Assess Forest Structure Using Discrete Lidar Data. *Remote Sensing of Environment*, 115(3), pp.824–835.

Nascimento, J.M.P. & Dias, J.M.B., 2005. Does Independent Component Analysis Play a Role in Unmixing Hyperspectral Data? *IEEE Transactions on Geoscience and Remote Sensing*, 43(1), pp.175–187.

Nelson, R.F. & Hoffer, R.M., 1981. *Landsat Data Procedure 1 and Forestland Classification Using Landsat Data*,

Neuenschwander, A.L., Magruder, L. a. & Tyler, M., 2009. Landcover Classification of Small-Footprint, Full-Waveform Lidar Data. *Journal of Applied Remote Sensing*, 3(1), p.033544.

Nourzad, S.H.H. & Pradhan, A., 2012. Ensemble Methods for Binary Classifications of Airborne LIDAR Data. *Journal of Computing in Civil Engineering*, 28(215), p.121130224813001.

Ormsby, J.P., 1982. Classification of Simulated And Actual NOAA-6 AVHRR Data for Hydrologic Land-Surface Feature Definition. *IEEE Transactions on Geoscience and Remote Sensing*, GE-20(3), pp.262–268.

Pal, M. & Mather, P.M., 2003. An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification. *Remote Sensing of Environment*, 86(4), pp.554–565.

Pan, Z., Glennie, C., Hartzell, P., Fernandez-Diaz, J., Legleiter, C. & Overstreet, B., 2015. Performance Assessment of High Resolution Airborne Full Waveform LiDAR for Shallow River Bathymetry. *Remote Sensing*, 7(5), pp.5133–5159.

Paris, C. & Bruzzone, L., 2015. A Three-Dimensional Model-Based Approach to the Estimation of the Tree Top Height by Fusing Low-Density LiDAR Data and Very High Resolution Optical Images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1), pp.467–480.

Parrish, C.E. & Nowak, R.D., 2009. Improved Approach to LIDAR Airport Obstruction Surveying Using Full-Waveform Data. *Journal of Surveying Engineering*, 135(2), pp.72–82.

Parrish, C.E., Rogers, J.N. & Calder, B.R., 2014. Assessment of Waveform Features for Lidar Uncertainty Modeling in a Coastal Salt Marsh Environment. *IEEE Geoscience and Remote Sensing Letters*, 11(2), pp.569–573.

Pedergnana, M., Marpu, P.R., Dalla Mura, M., Benediktsson, J.A. & Bruzzone, L., 2011. Fusion of Hyperspectral and Lidar Data Using Morphological Attribute Profiles. In L. Bruzzone, ed. *SPIE-Image and Signal Processing for Remote Sensing XVII*. p. 81801G–81801G–8.

Perry, E.M., Foote, H.P., Petrie, G.M., Steinmaus, K.L., Irwin, D.E. & Stephan, a. J., 1998. Exploitation of Hyperspectral Imagery and Lidar for Landuse Classification.

IGARSS '98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings. (Cat. No.98CH36174), (509), pp.1013–1015.

Pirotti, F., 2011. Analysis of Full-Waveform LiDAR Data for Forestry Applications: A Review of Investigations and Methods. *IForest*, 4(JUNE), pp.100–106.

Popescu, S.C., Wynne, R.H. & Scrivani, J. a., 2004. Fusion of Small-Footprint Lidar and Multispectral Data to Estimate Plot-Level Volume and Biomass in Deciduous and Pine Forests in Virginia, USA. *Forest Science*, 50(4), pp.551–565.

Puttonen, E., Suomalainen, J., Hakala, T., Räikkönen, E., Kaartinen, H., Kaasalainen, S. & Litkey, P., 2010. Tree Species Classification from Fused Active Hyperspectral Reflectance and LIDAR Measurements. *Forest Ecology and Management*, 260(10), pp.1843–1852.

Qin, Y., Vu, T.T. & Ban, Y., 2012. Toward an Optimal Algorithm for LiDAR Waveform Decomposition. *System*, 9(3), pp.482–486.

Qu, Z., Kindel, B.C. & Goetz, A.F.H., 2003. The High Accuracy Atmospheric Correction for Hyperspectral (Hatch) Data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6), pp.1223–1231.

Reitberger, J., Krzystek, P. & Stilla, U., 2008. Analysis of Full Waveform LIDAR Data for the Classification of Deciduous and Coniferous Trees. *International Journal of Remote Sensing*, 29(5), pp.1407–1431.

Richards, J. a & Jia, X., 2006. *Remote Sensing Digital Image Analysis: An Introduction*,

Roger, R.E., 1996. Sparse Inverse Covariance Matrices and Efficient Maximum Likelihood Classification of Hyperspectral Data. *International Journal of Remote Sensing*, 17(3), pp.589–613.

Roncat, A., Bergauer, G. & Pfeifer, N., 2011. B-Spline Deconvolution for Differential Target Cross-Section Determination in Full-Waveform Laser Scanning Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), pp.418–428.

Rottensteiner, F., Trinder, J., Clode, S. & Kubik, K., 2005. Using the Dempster-Shafer Method for the Fusion of LIDAR Data and Multi-Spectral Images for Building Detection. *Information Fusion*, 6(4), pp.283–300.

Sadro, S., Gastil-Buhl, M. & Melack, J., 2007. Characterizing Patterns of Plant Distribution in a Southern California Salt Marsh Using Remotely Sensed Topographic and Hyperspectral Data and Local Tidal Fluctuations. *Remote Sensing of Environment*, 110(2), pp.226–239.

Sarrazin, D., van Aardt, J., Asner, G.P., McGlinchy, J., Messinger, D.W. & Wu, J., 2010. Fusing Waveform Lidar and Hyperspectral Data for Species-Level Structural Assessment in Savanna Ecosystems. , 7684(585), p.76841H–76841H–10.

Sarrazin, M.J.D., van Aardt, J. a N., Asner, G.P., Mcglinchy, J., Messinger, D.W. & Wu, J., 2011. Fusing Small-Footprint Waveform LiDAR and Hyper Spectral Data for Canopy-Level Species Classification and Herbaceous Biomass Modeling in Savanna Ecosystems. *Canadian Journal of Remote Sensing*, 37(6), pp.653–665.

Sarrazin, M.J.D., Aardt, J.A.N. Van, Asner, G.P., Mcglinchy, J., Messinger, D.W. & Wu, J., 2012. Fusing Small-Footprint Waveform LiDAR and Hyperspectral Data for

Canopy-Level Species Classification and Herbaceous Biomass Modeling in Savanna Ecosystems. , 37(6).

Schaer, P., Skaloud, J., Landtwing, S. & Legat, K., 2007. Accuracy Estimation for Laser Point Cloud Including Scanning Geometry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, p.N/A.

Schwarz, K.P., Chapman, M.A., Cannon, M.E. & Gong, P., 1993. An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data. *Photogrammetric Engineering & Remote Sensing*, 59(11), pp.1667–1674.

Sellers, P.J. et al., 1995. Remote Sensing of the Land Surface for Studies of Global Change: Models — Algorithms — Experiments. *Remote Sensing of Environment*, 51(1), pp.3–26.

Serpico, S.B. & Bruzzone, L., 2001. A New Search Algorithm for Feature Selection in Hyperspectral Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), pp.1360–1367.

Shimoda, H., Fukue, K., Yamaguchi, R., Zhang, Z. jue & Sakata, T., 1988. Accuracy of Landcover Classification of TM and SPOT Data. In *International Geoscience and Remote Sensing Symposium (IGARSS)*. Publ by IEEE, pp. 529–535.

Shippert, P., 2003. Introduction to Hyperspectral Image Analysis. *Online Journal of Space Communication*.

Shlien, S. & Smith, A., 1975. A Rapid Method to Generate Spectral Theme Classification of LANDSAT Imagery. *Remote Sensing of Environment*, 4(C), pp.67–77.

Sinnock, S., Melhorn, W.N. & Montgomery, O.L., 1974. Machine-Aided Analysis of Land Use - Landform Relations From ERTS-1 MSS Imagery, Sand Hills Region, NEBRASKA. *Conf on Earth Resour Obs and Inf Anal Syst, Remote Sensing of Earth Resour*, 3, pp.503–526.

Skaloud, J. & Lichti, D., 2006. Rigorous Approach to Bore-Sight Self-Calibration in Airborne Laser Scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1), pp.47–59.

Smith, L.I., 2002. A Tutorial on Principal Components Analysis Introduction. *Statistics*, 51, p.52.

Smits, P.C., Dellepiane, S.G. & Schowengerdt, R. a., 1999. Quality Assessment of Image Classification Algorithms for Land-Cover Mapping: A Review and a Proposal for a Cost-Based Approach. *International Journal of Remote Sensing*, 20(8), pp.1461–1486.

Sohn, G. & Dowman, I., 2007. Data Fusion of High-Resolution Satellite Imagery and LiDAR Data for Automatic Building Extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1), pp.43–63.

Strahler, A.H., 1980. The Use of Prior Probabilities in Maximum Likelihood Classification of Remotely Sensed Data. *Remote Sensing of Environment*, 10(2), pp.135–163.

Suárez, J.C., Ontiveros, C., Smith, S. & Snape, S., 2005. Use of Airborne LiDAR and Aerial Photography in the Estimation of Individual Tree Heights in Forestry. *Computers and Geosciences*, 31(2), pp.253–262.

Sugumaran, R. & Voss, M., 2007. Object-Oriented Classification of LIDAR-Fused Hyperspectral Imagery for Tree Species Identificationin an Urban Environment.

Sun, G. & Ranson, K.J., 2000. Modeling Lidar Returns from Forest Canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6), pp.2617–2626.

Swatantran, A., Dubayah, R., Roberts, D., Hofton, M. & Blair, J.B., 2011. Mapping Biomass and Stress in the Sierra Nevada Using Lidar and Hyperspectral Data Fusion. *Remote Sensing of Environment*, 115(11), pp.2917–2930.

Thompson, D.R., Gao, B.C., Green, R.O., Roberts, D.A., Dennison, P.E. & Lundeen, S.R., 2015. Atmospheric Correction for Global Mapping Spectroscopy: ATREM Advances for the HyspIRI Preparatory Campaign. *Remote Sensing of Environment*, 167, pp.64–77.

Tolt, G. & Larsson, H., 2007. Waveform Analysis of Lidar Data for Targets in Cluttered Environments. *Proceedings of SPIE*, 6739, p.67390A.

Tran, T.N., Wehrens, R. & Buydens, L.M.C., 2005. Clustering Multispectral Images: A Tutorial. *Chemometrics and Intelligent Laboratory Systems*, 77(1-2 SPEC. ISS.), pp.3–17.

Ulaby, F. & McNaughton, J., 1975. Classification of Physiography from ERTS Imagery. *Photogrammetric Engineering and Remote Sensing*, 41(8), pp.1019–1027.

Ververidis, D. & Kotropoulos, C., 2008. Fast and Accurate Sequential Floating Forward Feature Selection with the Bayes Classifier Applied to Speech Emotion Recognition. *Signal Processing*, 88(12), pp.2956–2970.

Voss, M. & Sugumaran, R., 2008. The Seasonal Effect on Tree Species Classification in an Urban Environment Using Hyperspectral Data, LiDAR, and an Object-Oriented Approach. *Sensors*, 8(5), pp.3020–3036.

Wagner, W., Ullrich, a & Melzer, T., 2004. From Single-Pulse to Full-Waveform Airborne Laser Scanners: Potential and Practical Challenges. *International Archives of Photogramme-try, Remote Sensing and Spatial Information Sciences*, pp.201–206.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T. & Studnicka, N., 2006. Gaussian Decomposition and Calibration of a Novel Small-Footprint Full-Waveform Digitising Airborne Laser Scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2), pp.100–112.

Wang, H., Glennie, C. & Prasad, S., 2013. Voxelization of Full Waveform LiDAR Data for Fusion with Hyperspectral Imagery. *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp.3407–3410.

Weed, C., Crawford, M., Neuenschwander, A. & Gutierrez, R., 2002. Classification of LIDAR Data Using a Lower Envelope Follower and Gradient-Based Operator. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 00(C), pp.1384–1386.

Wehr, A. & Lohr, U., 1999. Airborne Laser Scanning—an Introduction and Overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3), pp.68–82.

Wozencraft, J.M., Macon, C.L. & Lillycrop, W.J., 2007. CHARTS-Enabled Data Fusion for Coastal Zone Characterization. In *Coastal Sediments '07*. Reston, VA: American Society of Civil Engineers, pp. 1827–1836.

Wu, H. & Prasad, S., 2013. Infinite Gaussian Mixture Models for Robust Decision Fusion of Hyperspectral Imagery and Full Waveform LiDAR Data. *2013 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013 - Proceedings*, pp.1025–1028.

Wu, J., Van Aardt, J. a N., McGlinchy, J. & Asner, G.P., 2012. A Robust Signal Preprocessing Chain for Small-Footprint Waveform LiDAR. *IEEE Transactions on Geoscience and Remote Sensing*, 50(8), pp.3242–3255.

Yamazaki, T. & Gingras, D., 1999. Unsupervised Multispectral Image Classification Using MRF Models and VQ Method. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2 II), pp.1173–1176.

Zhang, Q., Tian, Y., Yang, Y. & Pan, C., 2015. Automatic Spatial – Spectral Feature Selection for Hyperspectral Image via Discriminative Sparse Multimodal Learning. , 53(1).

Zhang, Y., Member, S., Yang, H.L., Member, S., Prasad, S., Member, S., Pasolli, E., Jung, J. & Crawford, M., 2015. For Classification of Multisource Remote Sensing Data. *IEEE journal of selected topics of Earth Observation*, 8(2), pp.845–858.

Zhang, Y. & Shen, X., 2013. Direct Georeferencing of Airborne LiDAR Data in National Coordinates. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84, pp.43–51.