ABSTRACT

Jung, Jinha Ph.D., Purdue University, May 2011. Integration of Full Waveform LiDAR and Hyperspectral Data for Characterization of Forest Structure. Major Professor: Melba M. Crawford.

Understanding characteristics of forest structure at large scales is of increasing importance because of its critical role in research and applications, including global carbon cycle studies, forest management planning, and wildlife habitat analysis. Structural characteristics provide valuable insights related to forest ecosystem function at individual tree, stand, forest, and global scales. Although field sampling based approaches are the most reliable and accurate means to characterize forest structure, they are not only destructive, time consuming, expensive, and limited to local scale studies, but also may be biased due to human interpretation and field conditions throughout a campaign period. To overcome the limitations of the field sampling based approaches, remote sensing technologies have been widely utilized to characterize forest structure at both local and global scales.

Recently, the use of multi-sensor data for characterizing forest structure has gained great attention of scientists, because data fusion approaches integrate information from multiple sensors that provide complementary information on targets, overcoming some of the problems of single source remote sensing data. Although several combinations of multi-sensor data and various fusion approaches have been investigated for mapping forest structure, little research has focused on the integration of full waveform LIDAR and hyperspectral data. While synergism clearly exists between hyperspectral and full waveform LIDAR data, two issues must be addressed in utilizing full waveform LiDAR and hyperspectral data for characterizing forest structure: 1) data integration strategies, and 2) feature extraction methodologies to discover a lower dimensional feature space from the high dimensional data.

The two issues motivate development of a new framework for characterizing forest structure using full waveform LiDAR and hyperspectral data. In the proposed framework, pixel-level data fusion is accomplished by reconstructing representative waveforms over a common grid structure, and low dimensional features are extracted using unsupervised machine learning algorithms. The extracted features are then used to develop predictive regression models in combination with field measurement data. For the framework development, this dissertation focuses on three primary research areas.

First, a sequential LiDAR waveform decomposition algorithm is developed since the data fusion step involves decomposing a waveform into a mixture of Gaussians. The proposed algorithm utilizes a sequential approach, where the goal is to reduce computation, but provide a good approximation to the waveform. It uses a region growing algorithm to estimate initial parameters of the newly added components, and utilizes different kinds of parameter optimization techniques depending on complexity of the waveforms. Experimental results demonstrate that the proposed algorithm utilizes a smaller number of components to decompose the waveforms, while it provides better approximation than traditional waveform decomposition algorithms.

Second, utilization of unsupervised machine learning algorithms is investigated to derive a lower dimensional feature space from the LiDAR waveform data. Waveform reconstruction scheme is proposed so that the unsupervised machine learning algorithms can be directly applied to the waveform data. Linear (PCA) and nonlinear (Isomap) feature extraction methods are applied to derive lower dimensional features from the waveforms, and they are compared to the traditional feature extraction methods. Experimental results indicate that the Isomap transformation successfully discovers a set of low dimensional features that can be used to characterize forest structure without manual interpretation.

Finally, the proposed framework to integrate full waveform LiDAR and hyperspectral data for forest structure characterization is applied to La Selva Biological Station which is one of the most actively studied tropical rain forests in the world. LAI (Leaf

Area Index) measurements acquired directly from a modular tower are used as ground reference data. Experimental results indicate that the best estimation results can be achieved when features extracted from the integrated data are utilized to build prediction models, while contribution from the hyperspectral data is greater for the high canopy sites. The main contribution of the proposed framework can be summarized as two factors; 1) the ability to integrate full waveform LiDAR and hyperspectral data at the pixel level, and 2) automatic discovery of meaningful lower dimensional features from the integrated data using unsupervised machine learning algorithms.