APPROVAL SHEET

Title of Dissertation: Remote sensing of vegetation structure using computer vision

Name of Candidate: Jonathan P. Dandois
Ph.D. in Geography and Environmental Systems, 2014

Dissertation and Abstract Approved:

Dr. Erle C. Ellis
Associate Professor
Dept. of Geography and Environmental Systems

Date Approved: August 3, 2014

Curriculum Vitae

Name: Jonathan P. Dandois

Degree and date to be conferred: Ph.D., 2014

Secondary education: Bethesda Chevy Chase High School, Bethesda, Maryland, 1999

Collegiate institutions attended:

University of Maryland, Baltimore County, Ph.D., 2008-2014
University of Maryland, Baltimore County, B.S., 1999-2003

Major: Geography and Environmental Systems

Professional Publications:

Dandois, J.P., & Ellis, E.C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.

Dandois, J.P., & Ellis, E.C. (2010). Remote sensing of vegetation structure using computer vision. Remote Sensing, 2, 1157-1176.

Professional positions held:

Post-Doctoral Research Associate, Smithsonian Tropical Research Institute, Panama, 2014-Present.

Graduate Research Assistant, UMBC, Baltimore, MD. 2008-2014.

Teaching Assistant, UMBC, Baltimore, MD. 2010-2012.

GIS Manager, BSA Philmont Scout Ranch, Cimarron, NM. 2007-2008.

CAD Drafter / ADA Inspector, Alan Hamm Architects, Kensington, MD. 2004-2007.

ABSTRACT

Title of Document: REMOTE SENSING OF VEGETATION STRUCTURE USING COMPUTER VISION

Jonathan P. Dandois, Ph.D., 2014

Directed By: Dr. Erle C. Ellis, Department of Geography and Environmental Systems

High-spatial resolution measurements of vegetation structure are needed for improving understanding of ecosystem carbon, water and nutrient dynamics, the response of ecosystems to a changing climate, and for biodiversity mapping and conservation, among many research areas. Our ability to make such measurements has been greatly enhanced by continuing developments in remote sensing technology – allowing researchers the ability to measure numerous forest traits at varying spatial and temporal scales and over large spatial extents with minimal to no field work, which is costly for large spatial areas or logistically difficult in some locations. Despite these advances, there remain several research challenges related to the methods by which three-dimensional (3D) and spectral datasets are joined (remote sensing fusion) and the availability and portability of systems for frequent data collections at small scale sampling locations. Recent advances in the areas of computer vision structure from motion (SFM) and consumer unmanned aerial systems (UAS) offer the potential to address these challenges by enabling repeatable measurements of vegetation structural and spectral traits at the scale of individual trees. However, the potential advances offered by computer vision remote sensing also present unique challenges and questions that need to be addressed before this approach can be

used to improve understanding of forest ecosystems. For computer vision remote sensing to be a valuable tool for studying forests, bounding information about the characteristics of the data produced by the system will help researchers understand and interpret results in the context of the forest being studied and of other remote sensing techniques. This research advances understanding of how forest canopy and tree 3D structure and color are accurately measured by a relatively low-cost and portable computer vision personal remote sensing system: 'Ecosynth'. Recommendations are made for optimal conditions under which forest structure measurements should be obtained with UAS-SFM remote sensing. Ultimately remote sensing of vegetation by computer vision offers the potential to provide an 'ecologist's eye view', capturing not only canopy 3D and spectral properties, but also seeing the trees in the forest and the leaves on the trees.

REMOTE SENSING OF VEGETATION STRUCTURE USING COMPUTER VISION

By

Jonathan P. Dandois

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, Baltimore County, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

UMI Number: 3637314

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3637314

Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

© Copyright by
Jonathan P. Dandois
2014

Dedication

To Justin and Julia, team Keswick.

Acknowledgements

The author was supported as an NSF IGERT Trainee (Grant #0549469). Research was also supported by USDA Forest Service Joint Venture Agreement 06-JV-11242300-135 and NSF DBI Grant #1147089.

I wish to thank my advisor, Dr. Erle Ellis, as well as my committee members, Dr. Matthew Baker, Dr. Marc Olano, Dr. Geoffrey Parker, and Dr. Junmei Tang for all of their support and feedback along the way.

Thanks to the Ecosynth team for support in collecting forestry field data as well as UAS engineering; field team: Dana Boswell, Erik Anderson, Andrew Bofto, Dana Boyd, Natalie Cheetoo, Maeve Tilly, Darryl Wise, Svetlana Peshkoff, and Zachary Bergeron; algorithm coders: William Bierbower, Ecaterina Coman, Conor Fitzpatrick, Terrence Seneschal, Ashley Ryan, and Yu Wang; UAS engineers: Garrett Krol, Nisarg Patel, Lindsey Digman, and Stephen Gienow; and technician Andrew Jablonski.

Finally, thanks to my parents John and Diane for their love and support, and to my best friends, my brother Justin and wife Julia, for always trying to keep me smiling.

Table of Contents

Dedication ix
Acknowledgements x
Table of Contents xi
List of Tables xiii
List of Figures xvi
Chapter 1: Introduction and Research Background 1
1.0 Research Objectives 4
1.1 Background 8
1.1.1 Current progress and challenges in remote sensing of forests 8
1.1.2 UAS for remote sensing 13
1.1.3 Computer vision structure from motion 14
1.1.4 UAS-SFM remote sensing 16
1.1.5 Background on computer vision image features 20
1.1.6 Review of remote sensing data quality research 25
1.1.7 Sources of error in remote sensing canopy structure measures 30
Chapter 2: High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision 47
1.0 Introduction 48
1.1 Computer vision for remote sensing 53
1.2 UAS for remote sensing 54
2.0 Materials and methods 55
2.1 Study areas 55
2.2 Forestry field methods 56
2.3 Aerial LIDAR 57
2.4 Ecosynth—computer vision remote sensing 58
3.0 Results 69
3.1 Image acquisition and point cloud generation 69
3.2 Characteristics of Ecosynth point clouds 70
3.3 Digital terrain models 75
3.4 Canopy height, biomass and carbon estimates 77
3.5 Vertical canopy profiles 82
4.0 Discussion 84
4.1 Ecosynth canopy height models (CHMs) 84
4.2 Predictions of aboveground biomass and carbon 88
4.3 Observing canopy spectral dynamics in 3D 89
4.4 General characteristics of Ecosynth 3D point clouds 92
4.5 Practical challenges in producing Ecosynth point cloud measurements 93
5.0 Conclusions 98
Chapter 3: What is a point? Analysis of computer vision image features of vegetation 108
1.0 Introduction 109
1.1 SFM for remote sensing 3D-spectral fusion 111
1.2 Bundler 'Photo-Tourism' SFM data model 112
1.3 SIFT Scale Invariant Feature Transform feature detector 114

1.4 Research objectives 116
2.0 Materials and methods 117
2.1 Data collection 117
2.2 Data processing 118
2.3 Data analysis 121
3.0 Results 122
3.1 Evaluation of SFM point cloud fusion quality 122
3.2 Evaluation of point cloud image features 131
4.0 Discussion 135
4.1 What is a point? Whatever the feature detector wants it to be 136
4.2 Point as an image sample 137
4.3 Point as a feature descriptor 139
4.4 Point as a 3D-RGB point 140
4.5 SFM and TLS see vegetation differently 142
4.6 The fusion of computer vision and ecological remote sensing 143
5.0 Conclusions 146
Chapter 4: Optimal conditions for aerial measurements of forest canopy structure using computer vision 153
1.0 Introduction 154
1.1 UAS-SFM remote sensing 155
1.3 Research objectives 157
2.0 Materials and methods 158
2.1 Data collection 158
2.2 Data processing 162
2.3 Data analysis 164
3.0 Results 168
3.1 Ecosynth SFM estimates of canopy structure 168
3.2 Influence of scanning conditions on point cloud positioning accuracy 174
3.3 Influence of scanning conditions on measures of canopy sampling 175
3.4 Influence of wind of point cloud quality 176
3.5 Radiometric quality of Ecosynth point clouds 177
3.6 Influence of computation resources on Ecosynth point cloud quality 179
4.0 Discussion 183
4.1 Ecosynth remote sensing of forest canopies under varying conditions 183
4.2 Recommendations for optimal data collection 185
4.3 Future research on Ecosynth UAS-SFM point cloud data quality 189
5.0 Conclusions 192
Chapter 5: Conclusions 198
1.0 Major findings and contributions of research chapters 199
2.0 The rise of computer vision ecological image features 204
3.0 Concluding remarks 208
Appendices 211
Appendix 1: Supplemental material for Chapter 2 211
Appendix 2: Supplemental material for Chapter 3 217
Appendix 3: Supplemental material for Chapter 4 241
Glossary 244

List of Tables

Chapter 1

Table 1: Summary of recent SFM remote sensing research, sorted by publication year, including information about the image collection platform and data collection configuration, SFM processing algorithm, major research application and findings. AGL = Above Ground Level. ……..……...……18

Chapter 2

Table 1: Ecosynth and LIDAR point cloud density for different land cover types. Blanks indicate land cover type not present within study site. ……...……72

Table 2: Point cloud georeferencing error and accuracy across study sites for the GCP method……74

Table 3: Understory digital terrain model (DTM) errors (meters) across entire study sites compared to LIDAR Bare Earth DTM.……....……75

Table 4: Understory DTM error compared to LIDAR bare earth DTM across different land cover types for terrain slopes ≤ 10°. Reported as Mean error ± SD (RMSE) in meters……..…77

Table 5: Best linear regression predictors (canopy height metric with the highest R^2) of field measured mean heights of the 5 tallest trees per subplot (AvgTop5) across forested areas of the Knoll, Herbert Run, and SERC sites based on Ecosynth methods with different DTMs and LIDAR alone. RMSE is deviation in meters between field measured AvgTop5 and the specified subplot canopy height metric.…….……...79

Table 6: Predicted aboveground biomass (AGB) from Ecosynth and LIDAR canopy height metrics across forested 25 m subplots at Knoll, Herbert Run, and SERC sites. Simple linear models predicting AGB are presented for the canopy height metric producing the highest R^2. Estimated AGB values are from allometric models...……..81

Table 7: Key factors influencing the quality of data obtained by computer vision remote sensing. ……..….98

Chapter 3

Table 1: Target classification accuracy as error matrices between leaf-on and leaf-off TLS and SFM point clouds as the average observation rates per target across repetitions. A target was classified as red if more than 50 % of the points found within the target 3D search area had a red hue (330° – 20°). Breakdown of classification of points per target and per replicate are provided in Appendix A2.7. …….....……....123

Table 2: SFM point color precision as measured by the difference between point color and average color of pixels at the corresponding point location within original images for random samples of points from aerial image sets and points located within target areas for ground sets……..……..127

Table 3: What is a point? Three primary traits of a SIFT-based SFM point cloud point…….136

Chapter 4

Table 1: Experimental design of Ecosynth flights (a). Boxes with "X" indicate treatments that were successfully flown, gray shaded boxes indicate treatments that were not covered by data collection. (b) shows configurations of individual flights……..……..161

Table 2: Summary of Ecosynth TCH estimates relative to field and LIDAR measurements averaged across all replicates at Herbert Run, Knoll and SERC sites. ...169

Table 3: Summary of the response of Ecosynth quality traits and metrics from all replicates (n = 82) across major scanning condition treatments of lighting, altitude above the canopy, and image side and forward overlap. Values are mean and standard deviation for all replicates within each level. Significant differences / trends are indicated with the p-value for lighting conditions, or correlation (R^2) for altitude and overlap. 'NS' indicates no significant trend at the $\alpha = 0.05$ level.. ……...……...171

Table 4: Variation in Ecosynth point cloud quality traits and metrics as a function of average wind speed during each flight based on the Beaufort wind force scale…...…..177

Table 5: Mean and standard deviation of the percent channel variation for red, green, and blue point colors as well as the point color gray scale intensity and mean rugosity by landcover at Herbert Run…….……179

Table 6: Computation time, quality traits and metrics for the same 'standard' dataset run across 6 computer configurations, running version 0.91 of Photoscan.……..……180

Table 7: Comparison of Ecosynth point cloud quality traits and metrics for a single replicate processed under different SFM algorithms..……..……...……182

Table 8: Average Ecosynth point cloud quality traits and metrics for the 'optimal' Ecosynth data collection: clear skies, 80 m altitude, 75 % side overlap processed with Photoscan v0.91, n=7……186

Appendix 1

A1.2: Image acquisition and computer vision data processing characteristics across study sites (6.25 ha) for Knoll and Herbert Run sites……...212

A1.3: Summary of field measured tree heights of the average of the 5 tallest trees per subplot (AvgTop5) at Knoll, Herbert Run, and SERC……....214

A1.4: Point cloud georeferencing error and accuracies across sites based on the spline method.……..…...214

Appendix 2

A2.2: Table describing UAS aerial image datasets.…….……..…....217

A2.6: Set of common tags spanning five descriptive categories (color, shape, surface, vegetation objects, and other objects provided to users for tagging image feature tiles in custom graphical user interface (Appendix A 2.5)..……...…....223

A2.7: Table showing breakdown of classification of points per target and per replicated for SFM and TLS single tree scans under leaf-on and leaf-off conditions. Values show proportion of points located within the target area for each replicate that fell within the red hue cutoff region (330 – 20). Blank cells ('—') indicate that no points were observed at the target area. Cells with value of 0.00 had > 1 point within the target area, but no points were within the red cutoff region. Averages, for which classification accuracy is calculated on, consider blank values ('—') as 0. Values of 'Average Red' indicate the average proportion of points located at a target that were red. Values greater than 0.5 are highlighted in bold face and indicate that a target was 'accurately classified'...…….....224

Appendix 3

A3.1c Raw calibration values……..…243

List of Figures

Chapter 2

Figure 1: Workflow for Ecosynth remote sensing…….……59

Figure 2: Ecosynth RGB point clouds across 250 m x 250 m sites (Knoll, Herbert Run and SERC; purple outline) under leaf-on (a, b) and leaf-off (c, d) conditions and viewed from overhead (a, c) and obliquely with same heading (b, d). Point clouds have black background; brightness and contrast enhanced using autocorrect settings in Microsoft Visio software……71

Figure 3: Point cloud density maps from Ecosynth under (a) leaf-on and (b) leaf-off conditions compared with (c) LIDAR across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). LIDAR densities combine first and last returns. Commercial LIDAR at Knoll and Herbert Run sites have lower density map legend as indicated….....…73

Figure 4: Digital terrain maps (DTM) from (a) LIDAR and (b) leaf-off Ecosynth and differences (Δ) between LIDAR and (c) leaf-off and (d) leaf-on Ecosynth DTMs across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). Negative differences highlight areas where Ecosynth DTM is lower than LIDAR. DTM legends for (a) and (b) differ from (c), as indicated.. ……...76

Figure 5: Overhead maps of (a) Ecosynth and (b) LIDAR canopy height models (CHM) across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). Ecosynth CHMs produced using LIDAR DTMs..…….……...…78

Figure 6: Ecosynth estimated best linear regression predictors (Table 5) of field measured average maximum height per subplot (AvgTop5) across forested areas of the Knoll (a), Herbert Run (b), and SERC (c) sites. Ecosynth canopy heights estimated using LIDAR DTMs. Linear regression lines (dashed), R^2, linear models, and RMSE (m) are presented for each comparison, with solid gray reference lines along the one-to-one ratio. Circled data points are outliers based on Grubb's test (> 3 SD from the mean) and are not included in regression analysis…..……...……...…….…80

Figure 7: Time series of Ecosynth point clouds (overhead view), canopy height profiles (CHPs; (a), (c), (e), (g), (i), (k)) and mean RGB channel brightness ((b), (d), (f), (h), (j), (l)) of a 50 m x 50 m sample area of SERC site (red box in Figure 2a, 2c). Black dashed line is LIDAR CHP (all returns; 2011-10-05). Horizontal lines are estimates of mean canopy height, with standard error in (i): solid line - field height 38.4 m (2.1 m), dotted line - Ecosynth Q-90 35.2 m (1.2 m), dashed line - LIDAR Q-75 (34.3 m (0.8 m); Table 5). Overhead views (on black background) and RGB channel brightness are shown without color enhancement. CHPs show mean RGB color per 1 m bin, uniformly scaled to enhance brightness and contrast for viewing using min-max linear correction……83

Figure 8: Time series of Ecosynth relative green brightness (green line) and MODIS NDVI (black line) across the SERC site. Relative green brightness is for each Ecosynth point cloud (Figure 7) at the 38 m – 39 m height bin; error bars are standard deviation. Linear regression line (dashed), R^2, and linear models are presented in subset, with solid gray reference line along the one-to-one ratio……...……..…84

Chapter 3

Figure 1: Diagram of Bundler data model: (a) a 3D-RGB point cloud, (b) a point 'view list', (c) a set of point 'views', (d) 'image feature tiles' associated with each view, (e) a single 'image feature tile' with a circle indicating the scale of the original feature, and (f) the 128-D SIFT image feature descriptor as a histogram and numerical vector…….....113

Figure 2: Mean histograms of point hue for points that fell 'inside' the target zone (white bars) and points that fell 'outside' the target zone (gray bars) for SFM (top row) and TLS (bottom row) point clouds under leaf-on and leaf-off conditions. Error bars are standard error (n=10). Black vertical lines are red hue cutoff region (330° – 20°). Color bar source: http://commons.wikimedia.org/wiki/File:HueScale.svg, accessed 2014-03-26. …...….124

Figure 3: Example of the role of the SIFT detector in the placement of point cloud points. (a) original image overlaid with all SIFT points (green triangles) and point cloud points (pink circles), (b) – (d) insets show targets that were not assigned a point cloud point for this view with the associated gray-scale version of each inset at right. …….126

Figure 4: Tomographic slices of a 0.1 m section of leaf-on (a) and leaf-off (b) SFM and TLS point clouds. Pixels represent a 0.1 m cube within which at least 1 point was located. All slices are co-registered to the same coordinate system and viewpoint……..128

Figure 5: Mean vertical foliage profiles (VFPs) for SFM and TLS-subsampled point clouds under leaf-on and leaf-off conditions, and with leaf-on and leaf-off point clouds combined. Differences between SFM and TLS VFPs were not statistically significant based on Kolmogorov-Smirnov (K-S) test of distributions ($p < 0.0001$). ……...…...129

Figure 6: Example images used in single tree SFM reconstruction under leaf on (top) and leaf off (bottom) conditions and image view maps showing the density of where point cloud points were viewed from the perspective of the images. This corresponds to the density of points across all images for a single SFM replicate as diagrammed by pink circles in Figure 3…….....130

Figure 7: Clustering results (k =2) on SIFT 128D descriptor for 2000 points from aerial SFM point clouds. Left panels show 12 image feature tiles closest to cluster centroid. Center panels show the frequency distribution of the scale (meters) of cluster features. Right panels show the 'mean gray-scale intensity' of all cluster image feature tiles resized to the same size (200 pixels square).…….....131

Figure 8: Scatter plot of the gray-scale intensity of points versus the mean gray-scale intensity of the image feature tile around each point, symbolized by the same SIFT clusters as in Figure 7. Points located below the one-to-one line have a point that is brighter than its surroundings, points above the line are darker than the surroundings, as shown in the example images..…….....132

Figure 9: Sorted frequency distributions of manually identified 'tags' per image feature cluster for leaf-on aerial point cloud sample datasets. Clustering was performed separately on bright and dark points (Figure 8). Cluster frequency plots are truncated to the top 85 % to aid readability. Dotted horizontal lines indicate the 50th percentile. Starred tags ('*') are unique non-color tags in the top 50 % within the clustered dataset……..134

Chapter 4

Figure 1: Regression between Ecosynth average TCH per plot across all replicates (n = 82) to (a) field and (b) LIDAR TCH, (c) LIDAR TCH and field measurements, (d) the standard deviation in Ecosynth TCH by height, and (e) Ecosynth TCH RMSE error relative to vertical displacement of Ecosynth from LIDAR DTM (e). Solid line (a - c) is one to one line, error bars in are standard deviation, dotted lines are linear model (a, b, c, and e) and the horizontal dashed line in (d) is the mean…….169

Figure 2: Scatter plots of the relationship between mean point cloud density and mean Ecosynth TCH error relative to field measurements and LIDAR TCH estimates for a single set of 5 replicates sub-sampled from every image to every 10th image to simulate decreasing forward photographic overlap. Top axis indicates corresponding forward overlap: 60%, 64%, 68%, 72%, 76%, 80%, 84%, 88%, 92%, and 96%.....173

Figure 3: Average Ecosynth (a) and LIDAR (b) TCH estimated AGB (Mg ha-1) relative to field estimated AGB. Solid line is one to one line, error bars are standard deviation of Ecosynth AGB per plot……...174

Figure 4: Radiometric precision of Ecosynth point clouds per channel averaged within different landcovers across all main replicates at the Herbert Run site (n = 82). Average channel percent deviation measures variation (standard deviation) in point color within 1 m x 1 m bins, interpreted as a percentage of the maximum potential channel brightness (255).……...178

Figure 5: Average variation or contrast of point cloud point color values per channel within forest areas only under either clear or cloudy (diffuse) lighting conditions. All per channel differences in average variation under different lighting were significantly different based on analysis of variance ($p < 0.0001$)..…….179

Figure 6: Computation time required for SFM processing in Photoscan v0.91 based on the number of photos……...181

Chapter 5

Figure 1: Conceptual diagram highlighting a single flower cluster from SFM point cloud image features. *Dipteryx panamaenis* flower cluster viewed on ground (≈ 1m, a), and by a UAS (≈ 150 m, b, c). Flower 'feature' was observed and matched across multiple images (d), and as a single point cloud point (e). Image (a) is for reference and is not the same as (b). UAS imagery and 3D point clouds from the 50 ha permanent plot on Barro Colorado Island (BCI), Panama, 2013-06-11, produced by J. P. Dandois, unpublished research, (a) courtesy Stephanie Bohlman, BCI, 2013-06-11…….206

Appendix 1

A1.1: Workflow for computer vision (CV) remote sensing (Ecosynth). ……...…211

A1.5: LIDAR-estimated best linear regression predictor (Table 6) of field measured average maximum height per subplot (AvgTop5; a) and linear regression model of Ecosynth and LIDAR-estimated mean canopy height per plot (b) for the SERC site. Ecosynth and LIDAR canopy heights estimated using LIDAR DTM. Linear regression lines (dashed), R^2, linear models, and RMSE (m) are presented for each comparison, with solid gray reference lines along the one-to-one ratio. Circled data point in (a) is an outlier based on Grubb's test (> ±3 SD from the mean) and is not included in regression analysis…….215

A1.6: Map showing close correspondence of Ecosynth time-series data with stem map locations of the dominant Liriodendron tulipifera canopy trees at a 50 m x 50 m sample area of the SERC site. Point cloud points from Ecosynth Spring 1 subset at SERC (Figure 7) with heights ≥ 37 m are shown in green colored by the point relative green brightness or 'strength of green' (i.e., Sgreen = G/(R+G+B)). Grey and white circles correspond to the tree stem locations (stems ≥ 10 cm DBH), with white circles indicating Liriodendron tulipifera, and grey circles representing all other tree species. Crown class is labeled for all dominant and co-dominant canopy trees. Map is shown on a grey background for visualization only. Blank areas of the map mean that there were no Ecosynth points in that area with heights ≥ 37 m in the Spring 1 dataset and no stems with DBH ≥ 10 cm within the stem map. Tree stems without a 'D' or 'CD' crown class label are of either an intermediate or suppressed class….…….216

Appendix 2

A2.1: Diagram of computer vision structure from motion (SFM) and terrestrial LIDAR scanner (TLS) single tree scan configuration. Scanning was carried out around circle of 7 m radius from tree trunk center. TLS was positioned as indicated by gray boxes at north, east, south, west orthogonal positions around tree. SFM photos were taken 2.5° / 0.3 m intervals around circumference of scan area with camera centered on tree trunk……...217

A2.5: Example screenshot of custom graphical user interface developed to help users provide semantic tags (Appendix A 2.6) to SFM SIFT image feature tiles……223

A2.8: Tomographic slices of a 0.1 m (10 cm) section of combined leaf-on and leaf-off SFM and TLS point clouds centered at roughly the center of the tree stem. Pixels represent a 0.1 m x 0.1 m x 0.1 m cube (0.001 m^3; 1,000 cm^3) within which at least 1 point cloud point was located. All slices are co-registered to the same coordinate system (scale, rotation and translation) such that pixels are the same in each model…….226

A2.9: Samples of SFM image feature tiles from leaf-on and leaf-off single tree point cloud scans highlighting differences in the background around each feature that may contribute to differences in point vs. image feature tile RGB channel correlation and ratio. All SFM image feature tiles associated located within target areas for a single leaf-on replicate shown at actual relative scale, n = 55.....227

A2.10: Clustering results on SIFT 128D numeric descriptors for all datasets (Appendix A 2.2) and on a sample of points from leaf-on and leaf-off single tree datasets.....229

A2.11: Clustering results on all SIFT 128D descriptors for all datasets at k = 10 clusters.....236

A2.12: Clustering results on all SIFT 128D numeric descriptors broken up by bright or dark spot and by leaf-on or leaf-off aerial datasets.....237

Appendix 3

A3.1a: Boxplots of the relationship between planned altitude (meters above mean sea level) and mean values of 'absolute' (gray) and 'relative' (black) altitude from Arducopter telemetry files from n =72 flight replicates. Mean altitude values are shown as points with horizontal 'jitter' to highlight variability in actual flight altitude values for a given planned altitude.....241

A3.1b: Calibration model used to correct relative altitude values to laser rangefinder measured altitude for georeferencing of point clouds.....242

Chapter 1: Introduction and Research Background

Forests cover roughly 30% of global land area and hold 70-90% of terrestrial aboveground and belowground biomass, a key sink of global carbon (FAO and JRC 2012). Accurate understanding of the spatial extent, condition, quality, and dynamics of forests is vital for understanding their role in the biosphere (Houghton et al. 2009). Obtaining such information through field work alone would be extremely labor intensive and has been made possible within the last four decades by remote sensing technologies that map the extent and dynamics of structural, spectral, and even taxonomic traits of forests at spatial scales ranging from the single leaf to the entire planet (Defries et al. 2000; Lefsky et al. 2002; Zhang et al. 2003; Parker and Russ 2004; Hansen et al. 2008; Asner and Martin 2009; Richardson et al. 2009). Even so, no one remote sensing instrument can simultaneously capture structural and spectral traits and dynamics of forests at high spatial resolution (Lefsky and Cohen 2003).

Such technical limitations limit the potential research that can be carried out for improving understanding of forest ecosystems using remote sensing. In other words, ecologists are often unable to answer important questions about forest ecology due to the lack of tools needed to collect the necessary measurements. Several such questions are posed here along with the potential benefits and challenges of using remote sensing to answer them.

What are the spatial and temporal patterns of forest canopy phenology at the scale of individual crowns / tree species and how does this vary within and across sites? Remote sensing has made great advances to improving understanding of canopy dynamics at regional and global extents, but due to limitations of the remote sensing

instruments these observations are often only made at coarse resolutions (> 250 m pixel size), with the result that information about the phenology of individual trees is lacking. Such fine scale observations are expected to provide important insights into the relationship between the climate and the forest canopy, particularly as it pertains to the response of individual species to climate change (Richardson et al. 2007; Richardson et al. 2009).

What are the dynamics of canopy biomass density at seasonal and longer time intervals across different forest types? Fine scale observations of canopy three-dimensional (3D) structure are vital for understanding the density and dynamics of forest biomass as they provide information on canopy height as well as canopy, but these observations are difficult to obtain from satellite remote sensing systems due to a lack of spatial resolution and 3D information gaps (Houghton et al. 2009). While observations obtained from aircraft-based remote sensing (e.g., LIDAR, light detection and ranging) do provide the necessary measurements of 3D structure, these are too costly to collect at the necessary time scales.

How does the amount of tree flower and fruit production vary through time and as a function of climate, species, and local edaphic conditions? One of the main benefits of remote sensing to ecological research is that it enables a 'bird's-eye' view of the forest canopy, providing observations from a point of view that is difficult or impossible for a person to achieve otherwise, for example the timing and amount of canopy fruits and flowers (Wright et al. 1999). Even so, due to technical limitations and trade-offs of different remote sensing instruments (Lefsky and Cohen 2003) such observations are often unavailable at either the spatial or temporal resolution needed, or may be

unavailable due to frequent cloud cover, for example in the humid tropics (Hansen et al. 2008; Beuchle et al. 2011).

Recent advances in consumer, automated unmanned aerial systems (UAS) and automated computer vision structure from motion algorithms (SFM) offer the potential to address these challenges by enabling remote sensing of forest structural and spectral traits at high spatial resolution and frequencies, albeit at relatively small spatial extents (< 100 ha). UAS equipped with consumer digital cameras enables automated and repeatable aerial image acquisitions and SFM enables the creation of 3D 'point cloud' models that are similar to those produced by LIDAR but with multi-spectral (red-green-blue, RGB) color spectral information associated with each point. Combined, UAS-SFM remote sensing has already enabled mapping of canopy structure and spectral dynamics at the scale of individual trees ('Ecosynth', Dandois and Ellis 2010, 2013), yet despite this progress important information about the nature and quality of Ecosynth point clouds remains unstudied.

Ecosynth 3D-spectral point clouds are produced from images alone based on algorithms that were originally intended for visualization (Snavely et al. 2010). The image feature algorithms used for automatically matching images, a pre-requisite for the photogrammetric 3D reconstruction algorithms used in SFM, are often developed to work on hard objects with discrete boundaries, like the edges and corners of buildings, books, walls, or solid objects (Gil et al. 2010; Wu 2010; Gauglitz et al. 2011) and vegetation is often considered noise in such research (Omerčević et al. 2008). This means that unlike with other remote sensing, a definition of what is being measured, what an individual datum (a point cloud point) represents, is lacking. This makes more in-depth

interpretability or comparison of the data to other types of measurements difficult. And while measures of data quality that are used for LIDAR remote sensing of forests can be applied to SFM systems (e.g., evaluating point clouds based on metric accuracy to a standard like tree height for LIDAR), because the mechanisms of recording a measurement are different, it is unknown how changes in the conditions under which an observation is made will influence the quality and subsequent applications of the product. This dissertation research aims to address these uncertainties and in doing so introduce a new form of 'personal remote sensing' for observation of forest structural and spectral traits.

1.0 Research Objectives

The main objectives of this research are to (1) define a new form of 'personal remote sensing' of forest canopies using UAS and SFM (Ecosynth), (2) to improve understanding of what is being measured by this system ('What is a point cloud point?'), and (3) to evaluate the quality of Ecosynth point clouds of forest canopies and how changing the way the data is collected will influence data quality. To address these objectives, this dissertation is divided into three main research chapters, summarized below. In addition to these chapters, section 1.1 below summarizes the relevant literature on current progress in remote sensing of forests, the increasing role of UAS in remote sensing, as well as background in computer vision SFM research, and the ways in which the quality of remote sensing datasets is evaluated.

- **Chapter 2** presents the proof-of-concept of Ecosynth UAS-SFM remote sensing with the objective of characterizing the 3D structure and color-spectral dynamics of Temperate Deciduous forest canopies. This chapter presents research that has already

been published by the lead author (Dandois and Ellis 2013) and helps sets the stage for more in depth exploration of SFM remote sensing.

- **Chapter 3** addresses the question of what is being measured by SFM remote sensing, with particular attention on the individual point cloud point and 'image features', 'What is a point?' A novel graphical user interface (GUI) is developed to facilitate manual classification and interpretation of computer vision points as 'image features'.
- **Chapter 4** defines relevant measures of the quality of Ecosynth point clouds of forest canopies, including measures of canopy height, geometric accuracy, point cloud density, and canopy penetration. A rigorous remote sensing experiment is carried out to evaluate how the quality of Ecosynth point clouds varies as a function of flight configuration, weather conditions, and image processing. Recommendations are made for optimal flight configuration strategies for remote sensing of forest canopies.
- The dissertation concludes with a review of the major findings of these research chapters and also suggestions for future research using Ecosynth UAS-SFM remote sensing of vegetation, in particular as it pertains to the exploitation of computer vision image features as a means for classifying point cloud points into ecologically distinct categories like leaves and branches or potentially even fruits and flowers.

Research on the use of Ecosynth SFM remote sensing began in 2010 when the author used SFM and images collected with a kite to make accurate estimates of canopy height compared to field based measurements (Dandois and Ellis 2010). This dissertation builds upon that research and is organized to reflect the path of scientific discovery that took place as research questions were posed that aimed to better understand how traits of the

canopy could be measured by Ecosynth and to improve understanding of the measurement system itself. Chapter 2 of this dissertation reflects a major advance over earlier Ecosynth research as the use of automated UAS overcomes many of the challenges with consistently and reliably collecting images over a fixed area that occurred when a kite was used (Dandois and Ellis 2010, 2013). Research in Chapter 2 also explores the potential of using the inherent fusion of 3D-color spectral information from Ecosynth to characterize the dynamics of canopy structural and spectral traits at the scale of individual trees, a feat not practical or possible with other forms of remote sensing to date. However, the research of Chapter 2 also raises its own set of questions: what are the SFM point cloud points and how are they related to canopy objects? How does changing the conditions under which data are collected (lighting, altitude, overlap) influence point cloud data quality and what might be the optimal data collection conditions? Chapters 3 and 4 of this dissertation address those questions and in doing so place a bounding understanding around Ecosynth SFM 3D-spectral remote sensing of forest canopies as a valuable tool for forest ecology research.

The organization of this dissertation as well as its context and contribution to remote sensing of forest canopies in general, builds upon research into the use of LIDAR remote sensing for observing forests. The goal of early LIDAR research aimed to evaluate how well the new technology could be used to measure canopy height, and was often limited to a single line scan along the flight path and required significant manual processing and analysis (Nelson et al. 1988). As the technology advanced, so too did the applications and ways in which the canopy could be described, for example through the use of more sophisticated sensors that could map canopy structure over larger areas and

provide a more detailed depiction of canopy vertical structure (Lefsky et al. 2002). At the same time, research was carried out to improve understanding of the properties of the LIDAR measurement itself, including sources of measurement error as well the relationship between the LIDAR measurements and the properties of the objects being observed (Baltsavias 1999a, Glennie 2007). More recently, with increased access to multiple LIDAR datasets over the same locations, researchers also examined the relationship between the data collection configuration (altitude, sensor settings) and observations of the canopy (Næsset 2009a, 2009b). By developing a bounding understanding of the way in which LIDAR observes the forest canopy, researchers are able to evaluate change in forest biomass density based on data collected from different sensors with different configurations (Hudak et al. 2012). Similarly, this understanding makes it possible to develop advanced canopy height metrics that are invariant to changes in both canopy structure and sensor configurations (e.g., top-of-canopy height TCH; Asner and Mascaro 2014). Ultimately, it is the goal of this dissertation research to take a similar approach for evaluating how Ecosynth UAS-SFM remote sensing can be used for observing forest canopies and to place a bounding understanding on the properties of those observations. In doing so, this research will open the door for new research applications that are not possible or practical with existing forms of remote sensing, including characterizing canopy structural and spectral dynamics at the scale of individual trees and potentially even observing the flowering of individual crowns.

1.1 Background

1.1.1 Current progress and challenges in remote sensing of forests

Remote sensing is recognized as a fundamental part of earth system monitoring and is a core component of the National Ecological Observation Network (NEON; Schimel et al. 2011), the monitoring requirements of the REDD+ program (Reducing Emissions from Deforestation and forest Degradation; Goetz and Dubayah 2011; Saatchi et al. 2011), and several national scale forest monitoring programs (Shimabukuro et al. 2014). Global estimates of forest area extent and change are made possible by optical satellite remote sensing systems that produce images of the reflectance of electromagnetic energy from the land surface at frequencies ranging from days to weeks and spatial resolutions ranging from 10's - 100's of meters (Achard and Hansen 2014). In particular, the MODIS (250 – 1000 m resolution, Moderate Resolution Imaging Spectrometer; Zhang et al. 2003) and Landsat (30 m resolution; Hansen et al. 2008; Roy et al. 2014) satellite systems are increasingly preferred for this purpose over the use of national or regional reporting due to inconsistent methods, updating, and levels of quality (FAO and JRC 2012). Remote sensing systems with high temporal resolutions like MODIS (1 to 2 day repeat time) are driving similar advances in understanding ecosystem dynamics (e.g., phenology, Zhang and Goldberg 2011), including the response of terrestrial ecosystems to changes in climate and land use (Morisette et al. 2008; Frolking et al. 2009). Even so, there are inevitable tradeoffs between the spatial resolution, spatial extent, and temporal frequency possible with any one particular sensor, with finer spatial resolution coming at the cost of reduced extent and frequency (Lefsky and Cohen 2003). In addition, many areas of the planet, for example the humid tropics, are frequently under-sampled by

satellite imagery due to occlusion of the land surface by cloud cover and haze (Hansen et al. 2008; Beuchle et al. 2011).

Mapping forest cover alone is also insufficient for accurately characterizing changes in biomass and ecosystem functions as knowledge of the quality of forests is needed, including estimates of forest structure and density as it relates to forest type, climate, topography, and other environmental variables (DeFries et al. 2007; Dolan et al. 2009; Houghton et al. 2009; Zolkos et al. 2013; Magdon et al. 2014). Measurement of forest canopy structure is best suited to active systems like LIDAR (Light Detection and Ranging) and SAR (Synthetic Aperture Radar) which simultaneously capture information about canopy structure and the ground surface below the canopy to provide estimates of canopy height, from which it is possible to estimate canopy biomass density when combined with measurements obtained in the field (Dubayah and Drake 2000; Treuhaft et al. 2004). One short-lived satellite LIDAR sensor (GLAS, Geoscience Laser Altimeter System) demonstrated the value of such measurements over large spatial extents (U.S. states) for capturing forest structure to assess biomass change when combined with maps of forest area change from Landsat (Dolan et al. 2009). However no satellite LIDAR systems currently exist for capturing forest structure across large scales (Dubayah et al. 2010; Goetz and Dubayah 2011) and satellite-based SAR often lacks the resolution to accurately discriminate between the ground and the top of canopy or struggles with other issues related to data co-registration (Ni et al. 2014). Estimation of canopy height is possible by applying stereo photogrammetric techniques to pairs of satellite-based optical images, but this too requires an accurate mapping of the understory terrain under most forest conditions (St. Onge et al. 2004; St. Onge et al. 2008).

Active remote sensing by airplane, in particular by LIDAR, often remains the preferred or only choice for characterizing canopy structure. LIDAR systems estimate the height and structure of the land surface based on the laser energy returned to the sensor after reflecting off of physical surfaces and is often depicted in the form of three-dimensional (3D) point clouds (Dubayah and Drake 2002; Lefsky et al. 2002). LIDAR remote sensing of vegetation 3D structure is a vital tool for carbon accounting (Lefsky et al. 2002; Lim et al. 2003; Frolking et al. 2009; Houghton et al. 2009; Goetz and Dubayah 2011), commercial and scientific forestry (Næsset and Gobakken 2008; Næsset 2011), ecosystem modeling (Thomas et al. 2008; Zhao and Popescu 2009; Antonarakis et al. 2011), and quantitative assessments of habitat suitability and biodiversity (Goetz et al. 2007; Vierling et al. 2008; Jung et al. 2012).

Recently, an advanced airborne remote sensing system was introduced that enables high-accuracy and high-resolution observations of vegetation structural and spectral traits through an integrated 'fusion' of LIDAR and hyperspectral remote sensing, yielding unprecedented capabilities for observing biodiversity and ecosystem functioning beyond the capabilities of either sensor alone (CAO, Carnegie Airborne Observatory; Asner et al. 2007; Asner and Martin 2009; Baldeck and Asner 2013, 2014). CAO achieves high-accuracy fusion based on an integrated sensor system that records LIDAR and hyperspectral imaging data at the same time from the same aircraft, combined with high-precision and accuracy GPS (Global Positioning System) and IMU (Inertial Monitoring Unit) equipment that enables precise location of both datasets relative to each other and the real world (≈ 1/3 pixel positioning, 0.45 – 1.45 m spatial resolution, Asner et al. 2007). This system overcomes many of the challenges faced when fusion is

attempted in a post-hoc fashion with LIDAR and imaging data collected from different sensors on separate missions, which includes issues with accurate data co-registration (Hudak et al. 2002; Mundt et al. 2006; Geerling et al. 2007) and the mis-match between 2D image pixels and 3D structure data points that occurs when a forest canopy is observed from different viewpoints (Popescu and Wynne 2004; Packalén et al. 2009). However, this integrated fusion technology remains prohibitively expensive for repeat monitoring of the same area at anything more frequent than annual intervals (Kampe et al. 2010; Schimel et al. 2011). Airborne LIDAR flights in general often cost a minimum of USD$20,000 per flight regardless of study area size (Erdody and Moskal 2010), representing a significant barrier to widespread application, especially for local environmental management and in ecological field studies based on annual or more frequent observations at numerous small sites or sampling plots (e.g., Holl et al. 2011).

The challenges presented by relying on airborne and satellite remote sensing alone have given rise to a variety of 'personal remote sensing systems' for measuring vegetation structure and spectral traits at high spatial resolution and frequency. Terrestrial LIDAR scanning (TLS) equipment that was primarily developed and used for engineering and industrial 3D scanning are now being deployed for modeling individual tree structure (Lefsky and McHale 2008; Hosoi and Omasa 2009; McHale et al. 2009), for modeling the distribution of photosynthetic and non-photosynthetic vegetation in the canopy (Seielstad et al. 2011), and for studying forest understory and canopy structure (Dassot et al. 2011). LIDAR scanning equipment commonly used for downward looking aerial scanning of forest structure has also been used from the ground looking up into the canopy for creating detailed profiles of canopy structure along sampling transects to

study the differences in forest vertical structure across different stand types (Parker et al. 2004; Hardiman et al. 2011). Seidel et al. (2011) provide a detailed review of the literature related to ground based measurements of canopy structure using manual and close-range sensor methods, including hemispherical photography, radiation sensors, and ground-based LIDAR.

Similar advances are being made in personal ground based remote sensing of vegetation optical properties. Annual time-series of close-range red-green-blue (RGB) color images of a fixed location above a forest collected daily from ordinary digital cameras can be used to track forest canopy phenology (Richardson et al. 2007). This has proven useful for understanding ecosystem carbon dynamics and when studied across multiple seasons may be useful for studying ecosystem response to climate change (Richardson et al. 2009; Mizunuma et al. 2013). Ryu et al. (2014) measured the phenology of multiple height strata within a forest canopy based on measurements from light emitting diodes (LEDs) of narrow wavelengths operating in 'reverse' mode (Ryu et al. 2010) for recording the amount of incoming light instead emitting light themselves. These technologies demonstrate that remote sensing does not have to be carried out from airplanes and satellites and that operating within a sampling framework can enable more in-depth insight into canopy structure and optical properties, often with relatively inexpensive equipment.

Two recent technological developments offer similar potential for overcoming the challenges of capturing forest structural and spectral traits at high spatial resolution and frequency: automated consumer-grade, hobbyist unmanned aerial systems (UAS) and relatively easy-to-use automated computer vision structure from motion algorithms

(SFM). Consumer grade UAS have reached a level of technical maturity to enable on-demand observation of forests at high spatial resolution and frequency using off-the-shelf digital cameras, ushering in an era of personal aerial remote sensing of forest canopies (Anderson and Gaston 2013). Easy-to-use SFM algorithms provide an automated image processing pipeline for producing inherently fused 3D-spectral point clouds of forest canopies from images alone (Snavely et al. 2006), providing LIDAR fusion-like datasets without the need for advanced integrated sensor technology.

1.1.2 UAS for remote sensing

The use of UAS has become increasingly common for remote sensing applications for studying vegetation, addressing some of the issues of portability and repeat collection of airborne and satellite remote sensing data. High spatial resolution photo mosaics were generated from photos taken with regular digital cameras mounted on remote controlled aircraft for monitoring crop health and nutrient status (Hunt et al. 2005; Hunt et al. 2009; Hunt et al. 2010) and rangelands (Rango et al. 2009). Recently, plantation tree height was estimated with high precision (≈ 0.5 m error) from a miniature LIDAR on a consumer UAS (Wallace et al. 2012; Wallace et al. 2014). Several recent studies provide a relatively comprehensive review of UAS for a variety of remote sensing applications and include information about different platforms, payloads, flight controllers, and potential applications (Eisenbeiss 2009; Anderson and Gaston 2013; Colomina and Molina 2014).

Compared to prior work that involved manually controlled UAS flight, Eisenbeiss (2009) demonstrated that UAS had reached a level of technological maturity to facilitate the systematic collection of photogrammetric image 'blocks' over small areas (0.4 ha – 10

ha), similar to what could be achieved with manned aircraft but at significantly reduced cost. They used existing photogrammetric software packages to manually align image sequences of between 40 - 160 photos and produce 3D digital surface models (DSMs) and found sub-meter accuracy in elevations compared to LIDAR and other reference data. They suggested that potential applications of UAS remote sensing could include quantifying sediment volume change, landscape slope and topography, and estimated the outer surface height of corn crop canopy to study cross-pollination dynamics (Eisenbeiss 2007; Vogler et al. 2009). But they also note major logistical challenges to repeatedly using this approach including the need to manually register image sets, which was time consuming and limited the number of photos, and therefore area, that could be covered.

1.1.3 Computer vision structure from motion

At around the same time that UAS was increasing in popularity for remote sensing, Snavely et al. (2006) demonstrated highly automated 3D reconstruction of scenes from unordered sets of digital photos collected from the Internet. Based on prior research from both photogrammetry and computer vision, they demonstrated that with images alone, the 3D structure of a scene could be modeled automatically with no initial knowledge of camera location based on the concepts of multi-view geometry using a technique referred to as structure from motion (SFM; Hartley and Zisserman 2000). The free and open-source software Bundler produced from that research represented a breakthrough in 3D modeling of scenes as it did not require the manual input and photogrammetric constraint of precise camera location information that limited the research of Eisenbeiss (2009). Instead of requiring time-consuming manual registration of images or information on camera location, the Bundler SFM algorithm uses automatically extracted 'image features'

to automatically match and align corresponding points across images. Given images alone, Bundler automatically produces a LIDAR-like 3D point cloud model of the scene with red-green-blue (RGB) color information from images inherently 'fused' with each point.

SFM algorithms follow three main processing steps to produce 3D point clouds from images alone: feature identification, feature matching, and bundle adjustment based on feature matches. In the feature identification stage, an 'image feature detector' algorithm is applied to all of the images to determine potential 'features' within each image that may be suitable for being tracked or matched in other images based on feature locations and 'feature descriptors'. While the terms are often used interchangeably, in this work the term 'detector' is used to refer to the feature detection algorithm and 'descriptor' is used to refer to that which is produced by the algorithm. There are many kinds of feature detectors, but most generally employ a two-step process of first 'locating' potential features based on image morphology to find specific patterns and then producing a 'feature descriptor' at or around the feature location that will be used in the feature matching stage (Szeliski 2011). Image features are reviewed in greater detail in Section 1.1.5. The result is a collection of features for each image that contain information about where the feature was located within in an image (XY coordinates of a point) and a numeric descriptor that acts like a feature signature or fingerprint.

In the feature matching stage, features are matched across images based on comparison of the numeric descriptor associated with each image, for example by determining which feature numeric descriptors are mathematically closest to each other by a measure of distance within the descriptor space (Snavely et al. 2006). Information

on feature matches is then used to parameterize the bundle adjustment process which seeks to optimize for the location of features as points in 3D space and the location and calibration of the cameras that observed those points within the same 3D space by minimizing the reprojection error of matching points across images (Triggs et al. 1999). The result is a 3D point cloud model of the observed scene where each point corresponds to a location that was observed by the feature detector, matched to other features, and used as part of the optimization process. SFM was used to produce 3D models of bare substrates from images collected with a camera positioned over a cliff (de Matías et al. 2009). Dandois and Ellis (2010) demonstrated the use of Bundler for characterizing forest canopy structure from images collected with a kite and ordinary digital camera. After georeferencing to the local geographic coordinate system, they showed that this relatively simple remote sensing system ('Ecosynth') produced estimates of canopy height comparable to those made in the field and LIDAR. Bundler is used in Chapter 3 where its data model is described in detail.

1.1.4 UAS-SFM remote sensing

Other similar SFM software solutions quickly appeared after Bundler, including the commercial, closed-source Agisoft Photoscan (http://www.agisoft.ru/). Photoscan is used in Chapter 2 and Chapter 4 as its computation optimization enables processing of large image sets in less time compared to Bundler (Dandois and Ellis 2013). Inspired by the ease with which UAS could be programmed for autonomous image collection and SFM could be used for automated 3D reconstruction, these technologies were quickly adopted for environmental research and a new form of remote sensing emerged. Tao et al. (2011) suggested the potential of using autonomous UAS and SFM for characterizing canopy

structure, but provided no quantitative assessments. Researchers began using SFM and images collected from kites, on the ground, or from consumer grade autonomous UAS to characterize bare ground, stream channel topography, and archaeological sites in 3D (Lucieer et al. 2011; Verhoeven 2011; Castillo et al. 2012; d'Oleire-Oltmanns et al. 2012; Dey et al. 2012; Harwin and Lucieer 2012; James et al. 2012; James and Robson 2012; Rosnell and Honkavaara 2012; Turner et al. 2012; Westoby et al. 2012; Fonstad et al. 2013). Recent UAS-SFM remote sensing research is summarized in Table 1.

Table 1: Summary of recent SFM remote sensing research, sorted by publication year, including information about the image collection platform and data collection configuration, SFM processing algorithm, major research application and findings. AGL = Above Ground Level.

Study	Platform	Data collection configuration	Algorithm	Application	Major findings related to data quality
de Matías et al. 2009	Handheld	Camera held over outcropping above rock surface	Unspecified, custom SFM	Mapping of bare ground topography	< 25 cm horizontal and vertical error compared to a total station
Dandois and Ellis 2010	Kite	Unordered, wind dependent 20 – 110 m AGL	Bundler	Estimating canopy height, biomass	R^2 0.53 – 0.80, RMSE < 4.3 m to field heights, > 40 % error in biomass estimates
Lucieer et al. 2011	Manual RC Helicopter	50 m AGL	Bundler and PMVS	Imaging of moss beds	3 cm spatial alignment, visually satisfactory
Tao et al. 2011	Unspecified UAS	Unspecified, > 60 % forward overlap & > 40 % side overlap	Unspecified, Bundler-like SFM	Mapping canopies	No comparison to field heights, merely suggests potential
Verhoeven 2011	Kite, manned aircraft	Unspecified	Photoscan	Mapping archaeological site topography	Suggests potential for mapping archaeological sites
Castillo et al. 2012	Handheld	Ground – along stream channel	Bundler and PMVS	Stream channel modeling and gully erosion and change	< 5% error in stream cross-sectional error compared to LIDAR
Dey et al. 2012	Handheld	Ground, walking along vineyard rows	Bundler and PMVS	Classification of grapes on a vine	Accuracy > 95 % for accurately identifying grapes on a vine
Harwin et al. 2012	Oktokopter multirotor UAS	30 – 50 m AGL	Bundler and PMVS	Mapping bare ground topography	< 10 cm horizontal and vertical accuracy compared to high quality GCPs
James et al. 2012	Handheld and manned ultralight	Ground, 800 m AGL	Bundler and PMVS	Mapping bare ground topography	Comparable to LIDAR for mapping volume change with ~80 % reduction in workflow time
James and Robson 2012	Handheld, fixed point	Ground – around lava flow areas	Bundler and PMVS	Mapping lava flow area morphology and channels	Suggest sub-meter accuracies of topography
Rosnell and Honkavaara 2012	Quadcopter, multirotor UAS	70 m AGL, 55 – 70 % side, 75 – 90 % forward overlap	Photosynth SFM, SOCET SET photogrammetry	Producing digital surface models over field, grass, forest, asphalt, and gravel	High image overlap is desirable – SFM produced more surface noise than photogrammetric software
Turner et al. 2012	Oktokopter multirotor UAS	50 m AGL	Bundler and PMVS	Mapping bare ground topography	Sub-meter accuracy < 15 cm
Westoby et al. 2012	Handheld	Ground – along terrain profile	Bundler and PMVS	Mapping bare ground topography	Sub-meter precision of terrain models compared to TLS
Dandois and Ellis 2013	Hexacopter Multirotor UAS	~ 80 – 90 m AGL, 50 % sidelap, > 90 % forward overlap	Photoscan	Estimate canopy height, biomass, canopy spectral phenology	R^2 > 0.83, RMSE < 4.4 m to field heights, strong correlation of canopy greenness to MODIS NDVI
Fonstad et al. 2013	Kite	Unordered, 100 m AGL	Photosynth	Mapping bare topography	R^2 = 0.97, RMSE 1.0 m compared to LIDAR, sub-meter positioning precision to high-accuracy GPS

Table 1: Continued

Study	Platform	Data collection configuration	Algorithm	Application	Major findings related to data quality
Liesin et al. 2013	Gatewing, fixed wing UAS	225 m AGL, 75 % forward and side overlap	MicMac SFM	Estimate canopy height	$R^2 > 0.96$ to LIDAR, >0.82 to field measurements with < 1.7 m RMSE error
Matthews and Jensen 2013	Hawkeye kitewing UAS	125 m AGL	Photoscan	Predict vineyard LAI	R^2 0.56 compared to field measured LAI
Javernick et al. 2014	Manned helicopter	600 – 800 m AGL	Photoscan	Mapping stream channel morphology	< 1m RMSE error in mapping channel bathymetry
Morgenroth and Gomez 2014	Handheld	Ground – around tree	Photoscan	Mapping single tree structure	Close agreement (< 5 % error) between field and laser measured tree measurements

This rapid growth culminated recently in the mapping of forest canopy structure and spectral dynamics using UAS and SFM at the scale of individual trees and at the extent of field plots (< 10 ha), offering the potential to address some of the challenges to remote sensing of forests (Dandois and Ellis 2013). Applications of UAS-SFM remote sensing continue, including for mapping forest height (Lisein et al. 2013), vineyard structure (Mathews and Jensen 2013), stream channels (Javernick et al. 2014), and single tree structure (Morgenroth and Gomez 2014). Leberl et al. (2010) argue that the rapid rise in computer vision SFM 3D reconstructions from images ('3D-vision') is saving photogrammetry due to otherwise complicated workflows made obsolete by LIDAR remote sensing, which provides users with 'instant gratification.' Colomina and Molina (2014), in their review of UAS for remote sensing, mention Bundler and Photoscan as photogrammetric software tools for producing digital surface models (DSMs), but do not mention the potential for modeling of forest canopy structure from SFM and only briefly reference applications relative to forests as it pertains to the ability of UASs to carry spectral and thermal imaging, LIDAR or SAR payloads.

1.1.5 Background on computer vision image features

Aside from the rising popularity in the use of SFM for remote sensing, the use of computer vision in remote sensing is not new and there are many examples of the use of computer vision concepts to improve, speed-up, automate, and otherwise enhance remote sensing research tasks and applications. A growing body of recent research also highlights the role that computer vision plays in advancing ecological research directly and not through remote sensing per se, often serving a similar role as it does for remote sensing by enhancing or automating existing ecological research tasks. The use of computer vision in remote sensing and ecological research span many applications, including image matching, tracking, and classification. In all cases, the benefit provided by computer vision often rests in the concepts and algorithms for establishing 'image features.' The reader is referred to the Glossary at the end of the dissertation for reference to common terms relevant for computer vision ecological remote sensing.

Image features are a fundamental component of computer vision research as they represent the quantification of particular aspects of an image that can be described by a numerical vector and used by a computer algorithm to attempt to recreate human vision tasks (Szeliski 2011). Image features take many forms and may be as simple as a set of three numbers describing the mean red, green, and blue color value of all pixels within an image or as complex as a large collection of 100's – 1000's of numerical vectors that describe patterns of image intensity, texture, and shapes. The creation of image features involves two main steps: feature detection and feature description (Li and Allinson 2008; Szeliski 2011). In the feature detection stage, image features are located within an image based on the identification of a particular signal or pattern in the image which is often

divided into geometric categories of points or corners, regions, patches or blobs, and lines or edges. Several studies review common feature detectors within these major categories (Zhang and Lu 2004; Mikolajczyk and Schmid 2005; Li and Allinson 2008; Szeliski 2011). In the feature description stage, a numerical vector is computed around each detected feature location that describes aspects of the pixel values of the image within a region around the feature based on patterns or filters, distributions of pixel intensity values, texture patterns, shapes, derivatives, or the image color moments (mean, variance, skew, kurtosis). Combined, an image feature refers to a specific location (XY coordinate) within an image and a numerical vector feature descriptor at that location. Image features are often classified based on their invariance to particular types of image distortion that would occur when the same point is observed from different locations, including invariance to changes in scale, rotation, and illumination (Lowe 2004; Mikolajczyk and Schmid 2004).

Applications of computer vision to ecological research often involve leveraging what computer vision does well, recreating and automating human vision tasks (Szeliski 2011), to tasks that would otherwise be difficult or impossible with manual labor alone. One popular example is the free software ImageJ (Abràmoff et al. 2004) which enables a variety of image processing and analysis tasks including image thresholding and edge detection as well as the extraction of image statistics in an easy-to-use and user-friendly graphical user interface. Other common applications of computer vision to ecological research involve the automatic characterization of animal behavior through the use of computer vision motion tracking techniques based on image features. Several have used image morphology techniques (e.g., differencing, Gaussian smoothing; Szeliski 2011) to

track the motion of mosquitos (Spitzen et al. 2013) and pigs (Ahrendt et al. 2011) within fixed areas. Ohayon et al. (2006) used computer vision motion tracking to improve understanding of the relationship between barn owl head movement leading up to attack movement. Computer vision image mosaicking and segmentation algorithms have been used to map and delineate coral reef communities (Lirman et al. 2007; Beijbom et al. 2012).

Remote sensing has also benefitted from algorithms and concepts of computer vision that provide semi-automated or automated solutions to remote sensing problems, often opening the door for new forms of image analysis, and often also based on image features. One of the most common ways that computer vision overlaps with remote sensing is in the use of image features for automated alignment or registration of multiple images (de Matías et al. 2009; Huang et al. 2009; Lingua et al. 2009; Zhang et al. 2009; Schwind et al. 2010; Oh et al. 2011). In these cases, image features are detected in each input image and corresponding features are matched to each other based on a measure of distance or difference based on the numeric feature descriptor. Images can then be aligned to each other based on knowledge of these correspondences and the known location of features. Image features serve this role as a core component of the algorithms behind the recently popular use of computer vision structure from motion (SFM) for remote sensing (Section 1.1 Background). Image features are used by SFM as a substitute for the manual identification of tiepoints that represents one of the major limitations and logistical challenges of traditional photogrammetric approaches (Eisenbeiss 2007; Eisenbeiss 2009). Wallace et al. (2012) used image features and SFM on frames extracted from video to estimate the trajectory of a UAS equipped with a

miniature LIDAR and found that the use of SFM significantly improved the geometric accuracy of the LIDAR product compared to using UAS GPS and IMU information alone.

Along with their use for image registration and image tracking, image features are also used for automated image classification. In image classification, a library of images that have already been classified is used to generate representative feature descriptors, which might be considered like image signatures, which are then used to attempt to classify new images or sections of images based on the knowledge obtained from the reference images (Szeliski 2011). This form of image classification is often referred to as using 'local features' or a 'bag of features' approach (BOF; Lazebnik et al. 2006). The theory behind classification with BOF is that the sematic content of an image or parts of an image can be identified based on the collection of image features that make up the image and relative to an existing library of already classified images and image features (Oliva and Torralba 2001; Csurka et al. 2004; Vogel and Schiele 2004; Fei-Fei and Perona 2005). In a BOF approach, multiple image features are produced and aggregated into a 'bag' for a single image or segmented regions of an image. The collection of features within a given bag is then interpreted to represent a feature signature for that image or image segment and is compared to an existing library of image classification bags to attempt to automatically describe what is in the un-classified image. This is conceptually similar to the use of spectral signatures in remote sensing for the characterization of land cover types (Rees 2001).

Beijbom et al. (2012) used image feature descriptors from a reference library of images of corals to automatically classify coral communities from new image surveys.

Image features, along with image segmentation, are used for automatic detection and classification of leaves and flowers (Nilsback 2009; Kumar et al. 2012; Kendal et al. 2013). Recently, the use of image features combined with citizen science has gained increasing popularity for producing classification databases from camera trap images and videos of animals, including for estimating regional animal abundance and tracking the behavior of insects, bears, and apes, among many organisms (as reviewed by Desell et al. 2013). Recent research has also extended the use of image features and classification for 'geographic image retrieval' of high resolution remote sensing images (Newsam and Yang 2007; Yang and Newsam 2008, 2013). In these studies, the Scale Invariant Feature Transform (SIFT; Lowe 2004), the same algorithm that is used for image matching by Bundler and in other research, was used to classify segments of high resolution aerial imagery into different categories of land cover, including forest, suburbs, marinas, urban areas, and parks.

The use of computer vision image features spans a range of applications including recognition of image and scene content and semantic classification. By limiting the interpretation of computer vision SFM points clouds to only considering 3D structure or the fusion of 3D structure and color, a large area of potential new research applications is discarded, including the automated mapping in 3D of individual canopy objects like leaves, branches, fruits, and flowers. Detailed mapping of such canopy objects could prove valuable for understanding the spatial heterogeneity of canopy structure, for example by providing estimates of leaf clumping and canopy gaps (Hilker et al. 2008). The use of image features may also be useful for automatically quantifying the amount of canopy fruits and flowers, which is a significant factor for understanding the relationship

between forest plant and animal communities (Wright et al. 1999), but which is challenging to measure in practice by manual methods alone due to the difficulties with accessing the forest canopy. The role, quality, and nature image features in SFM remote sensing of vegetation is explored in greater detail in Chapter 3.

1.1.6 Review of remote sensing data quality research

The evaluation of remote sensing data quality can be divided into three categories of theoretical, empirical, and applied methodological approaches, depending upon the goal of the research and the resources available. In a theoretical data quality approach, mathematical or physical-based understanding of the process of data collection by the remote sensing sensor is used to simulate observations and the error in observations by manipulation of system variables, which may include sensor configurations, scene weather conditions, and the properties of the land surface being observed (Börner et al. 2001). In optical imaging remote sensing the theoretical approach to data validation involves first developing a detailed understanding of the spectral response of the imaging sensor and then employing a model of surface light reflectance (i.e., radiative transfer model) to predict the characteristics of light that will be reflected to the sensor (Koetz et al. 2007). Radiative transfer models are also used to simulate how canopy structure is characterized by LIDAR to understand the relationship between such measurements and canopy traits like height, leaf clumping, leaf area (leaf area index, LAI), or leaf angle (Hilker et al. 2008). Theoretical understanding of LIDAR sensors facilitates analysis of the sources of error in range estimates given different conditions of local topographic variation and sensor configuration (Baltsavias 1999a; Habib and Cheng 2006; Glennie 2007; Habib et al. 2009).

A theoretical model for simulating SFM reconstructions of forest canopies from UAS imagery would require a system for producing synthetic aerial images of a scene with the ability to control for differences in altitude, overlap, and lighting, among many variables, while at the same time taking into account observable changes in the scene due to changing field of view and parallax from the perspective of a moving camera. These synthetic image sets could then be supplied to SFM to produce synthetic 3D point cloud models, from which it could be possible to characterize changes in point cloud quality as a function of different tunable parameters within the model. Existing tools do not capture the necessary combination of attributes to simulate photo-like images of a canopy. Some image simulation algorithms exist for producing single synthetic satellite images, however these are produced at coarse resolutions (> 30 m x 30 m) and spatial extents (> 100 km x 100 km) and are based on other existing images of the same place as reference (Schott et al. 1999; Ientilucci and Brown 2003; Schott et al. 2012). Radiative transfer models that are used to simulate optical image remote sensing do not reproduce a true 'image' of the scene, even though they are able to capture the expected spectral reflectance of that area (Börner et al. 2001; Guanter et al. 2009). And while substantial progress has been made in the development of radiative transfer models for simulating and evaluating remote sensing observations of vegetated surfaces, there are still challenges with accurately representing the optical properties of canopies in the form of model parameters, including measures of the spatial distribution, size and clumping of canopy foliage (Kuusk et al. 2014), factors which may strongly influence the way in which a canopy is observed in a photograph and 'seen' by SFM feature descriptors (explored in Chapter 3). 3D radiative transfer and ray-tracing models are used for

simulating the 3D canopy structure observed by LIDAR (Lovell et al. 2005; Morsdorf et al. 2009; Antonarakis et al. 2011) but such systems are not designed to produce photo-like images and would likely require the development of new feature descriptors in order to carry out the necessary feature matching steps in SFM. As such, a theoretical approach to evaluating the data quality of Ecosynth UAS-SFM remote sensing data products is outside the scope of the current work.

While theoretical data models exist for optical imaging and LIDAR, it is more common that research takes on an empirical or applied methodological approach to evaluate the quality of remote sensing products. Empirical data quality measures can be used to characterize the precision of a dataset as it exists on its own, without necessarily being applied to any particular research task. These empirical quality measures can then be considered quality 'traits' of the dataset that may be useful for understanding how good the dataset will be for a given application, compared to other similar datasets with different trait values. Empirical data quality traits also have the advantage of being relatively easy to compute without the need for dataset specific field work or manual analysis. Empirical measures of data quality often seek to address two issues: are the data in the correct place and how accurately is something observed. The geometric quality of optical images is often evaluated by comparison of the horizontal and vertical error of manually identified ground control points (Poli et al. 2014; Wolf et al. 2014). The radiometric quality of optical images is evaluated by quantifying the noise in spectral bands based on statistical measures of variability (standard deviation, roughness) within areas that should otherwise be relatively spectrally homogenous like large bare areas or water (Baltsavias et al. 2001; Poli et al. 2014). Image radiometric quality may also be

evaluated by comparing the spectral information at a homogenous location across multiple image scenes for image calibration and comparing spectral information of image features with spectral reflectance recorded with field spectrometer equipment at approximately the time of image collection (Yang and Lo 2000; Hunt et al. 2010). Empirical quality measures are also used as a form of quality control. For example, dedicated spectral bands are used to automatically estimate potential sources of error in satellite remote sensing, including the amount of cloud cover within a satellite image at the time of collection (Roy et al. 2014).

The primary measure of data quality for LIDAR is an estimate of the precision of 3D structure or vertical height measurements. Typically, the quality of LIDAR is reported as a measure of the horizontal and vertical error of LIDAR points relative to ground control points or relatively flat and homogenous surfaces with known elevations (Habib and Cheng 2006; Gatziolis and Andersen 2008; Habib et al. 2009). Habib et al. (2009) proposed a measure of LIDAR quality based on the amount of horizontal and vertical displacement between overlapping LIDAR swaths over the same areas. They found that by applying an automated fitting algorithm (Iterative Closest Point, ICP; Besl and McKay 1992) they could measure how much one swath had to be moved in order to be optimally aligned with another swath over the same area and quantified horizontal and vertical error of the LIDAR point cloud based on the magnitude of this displacement.

Finally, it is common to evaluate the quality of a dataset through application to a particular research task. In an applied approach, the dataset is treated as the final product and is used to estimate some other characteristic of the scene, evaluated based on a better measurement of that characteristic. This applies to when LIDAR, SFM, and stereo or

multiple-stereo photogrammetry 3D data are used to estimate tree height and are compared to field-based estimates of height using regression analysis (Lefsky et al. 2002; Lim et al. 2003; St. Onge et al. 2004; Ofner et al. 2006; Hirschmugl et al. 2007; St. Onge et al. 2008; Dandois and Ellis 2010, 2013). The applied accuracy approach is also employed when optical or optical-3D fusion datasets are used to classify different classes of vegetation and other land covers, and are commonly evaluated using the kappa statistic (Key et al. 2001; Dalponte et al. 2008; Ke et al. 2010). In the case of both LIDAR and SFM, data quality is commonly evaluated by considering how well a 3D surface is characterized. Points in a LIDAR or computer vision 3D model are compared to either the nearest other point on a surface or to a surface interpolated from a reference 3D dataset (Baltsavias 1999b; Seitz et al. 2006; Habib et al. 2009; Strecha et al. 2009).

Recently, several LIDAR studies have taken advantage of the coincidence of multiple datasets collected over the same area to provide insights into how different data collection strategies contribute to differences in datasets and the forest metrics derived from those datasets. Previous studies showed that while changes in flight altitude and pulse rate will change the sampling of the forest canopy, as measured by point cloud density (points m^{-2}), forest metrics derived from point clouds (e.g., canopy height and related measures of canopy volume and biomass density) are relatively robust and unaffected by different scan configurations (Næsset 2009b; Næsset 2009a). A study that examined the use LIDAR of collected over the course of several years found that even though the data were collected with different parameters, estimates of canopy height as well as derived estimates of change in canopy height and biomass, were closely matched to field estimates and were again relatively invariant to differences in scan configuration

(Hudak et al. 2012). However, others have shown that some canopy metrics are strongly dependent upon scanning configurations, particularly the resolution with which the canopy is sampled. For example, the scale of canopy gaps detected by LIDAR depends strongly upon the footprint size of the laser beam (Cook et al. 2009) and it is often accepted that LIDAR will under-predict the height of the peak of coniferous trees as the laser beam will be more likely to hit the crown 'shoulders' (Magnussen et al. 2010).

1.1.7 Sources of error in remote sensing canopy structure measures

Using applied techniques, remote sensing has enabled the observation of many aspects of forest canopy structural, spectral, and recently even taxonomic traits. There are numerous reviews available that characterize the state-of-the-science for remote sensing of vegetation traits, including structure and diversity (Nagendra 2001, Lefsky et al. 2002, Lim et al. 2003, Turner et al. 2003, Bergen et al. 2009). These reviews often characterize the types of measurements that can be obtained from different remote sensing sensors along with the different resolutions that are possible with each (spectral, spatial, and temporal). However, studies often do not describe sources of error and potential challenges that arise when actually trying to apply these remote sensing techniques to characterizing vegetation traits.

Estimates of vegetation structure and canopy height in particular from remote sensing, while becoming increasingly routine, are subject to several significant sources of error. One of the greatest sources of error is in the quality of the digital terrain model (DTM) used for interpreting heights from surface elevations. This includes errors that arise in attempting to filter a dataset to find the ground under the forest, or when canopy surface and ground measurements are obtained from different sensors where mis-

registration between the two datasets can lead to significant errors in height estimates (Sithole and Vosselman 2004; Næsset 2009b; Ni et al. 2014). Another major source of error in remote sensing estimates of canopy structure involves the relationship between the measurement of height in the field and the way the canopy is observed by the sensor. Field estimates of canopy height are often based on field sampling plots which may contain significant errors in height measurements (Bragg 2008; Goetz and Dubayah 2011; Larjavaara and Muller-Landau 2013). Errors in geolocation between field and remote sensing data and the fact that remote sensing observations of plots may observe canopy structure for crowns that fall within the plot boundary but which were not included in field estimates because the stem was outside the plot, can also lead to significant errors in height estimates (Frazer et al. 2011). Canopy height is often estimated from remote sensing 3D structure measurements from distribution-based metrics (e.g., mean, median or the N^{th} quantile of 3D observations within a plot) which may produce a biased estimate of plot canopy height due to the fact that the remote sensing instrument may observe more of the upper parts of the canopy (Magnussen and Boudewyn 1998; Parker and Russ; 2004; Hopkinson et al. 2006; Frazer et al. 2011). Application of a remote sensing instrument to measuring canopy properties therefore requires an understanding of the way in which the instrument observes the canopy as well as the quality of those observations as a function of the sensor itself and the thing being observed.

References

Abràmoff, M.D., Magalhães, P.J., & Ram, S.J. (2004). Image processing with ImageJ. Biophotonics international, 11, 36-43.

Achard, F., & Hansen, M.C. (2014). Global Forest Monitoring from Earth Observation. CRC Press, Boca Raton, FL.

Ahrendt, P., Gregersen, T., & Karstoft, H. (2011). Development of a real-time computer vision system for tracking loose-housed pigs. Computers and Electronics in Agriculture, 76, 169-174.

Anderson, K., & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. Frontiers in Ecology and the Environment, 11, 138-146.

Antonarakis, A.S., Saatchi, S.S., Chazdon, R.L., & Moorcroft, P.R. (2011). Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. Ecological Applications, 21, 1120-1137.

Asner, G.P., & Martin, R.E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Frontiers in Ecology and the Environment, 7, 269-276.

Asner, G.P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. Remote Sensing of Environment, 140, 614-624.

Baldeck, C., & Asner, G. (2013). Estimating Vegetation Beta Diversity from Airborne Imaging Spectroscopy and Unsupervised Clustering. Remote Sensing, 5, 2057-2071.

Baldeck, C., & Asner, G. (2014). Improving Remote Species Identification through Efficient Training Data Collection. Remote Sensing, 6, 2682-2698.

Baltsavias, E., Pateraki, M., & Zhang, L. (2001). Radiometric and geometric evaluation of Ikonos GEO images and their use for 3D building modelling. In (pp. 19-21): Citeseer.

Baltsavias, E.P. (1999a). Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing, 54, 199.

Baltsavias, E.P. (1999b). A comparison between photogrammetry and laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing, 54, 83-94.

Beijbom, O., Edmunds, P.J., Kline, D.I., Mitchell, B.G., & Kriegman, D. (2012). Automated annotation of coral reef survey images. In, Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 1170-1177).

Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M., Hunsaker, C.T., Imhoff, M.L., Nelson, R.F., Parker, G.G., & Radeloff, V.C. (2009). Remote sensing of vegetation

3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. Journal of Geophysical Research, 114.

Besl, P.J., & McKay, H.D. (1992). A method for registration of 3-D shapes. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 14, 239-256.

Beuchle, R., Eva, H.D., Stibig, H.-J., Bodart, C., Brink, A., Mayaux, P., Johansson, D., Achard, F., & Belward, A. (2011). A satellite data set for tropical forest area change assessment. International Journal of Remote Sensing, 32, 7009-7031.

Börner, A., Wiest, L., Keller, P., Reulke, R., Richter, R., Schaepman, M., & Schläpfer, D. (2001). SENSOR: a tool for the simulation of hyperspectral remote sensing systems. ISPRS Journal of Photogrammetry and Remote Sensing, 55, 299-312.

Bragg, D. (2008). An improved tree height measurement technique tested on mature southern pines. Southern Journal of Applied Forestry, 32, 38-43.

Castillo, C., Pérez, R., James, M.R., Quinton, J.N., Taguas, E.V., & Gómez, J.A. (2012). Comparing the Accuracy of Several Field Methods for Measuring Gully Erosion. Soil Science Society of America Journal, 76, 1319-1332.

Colomina, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 79-97.

Cook, B.I., Miller, R.L., & Seager, R. (2009). Amplification of the North American "Dust Bowl" drought through human-induced land degradation. Proceedings of the National Academy of Sciences, 106, 4997-5001.

Csurka, G., Dance, C.R., Fan, L., Willamowski, J., & Bray, C. (2004). Visual categorization with bags of keypoints. In, ECCV Statistical Learning in Computer Vision.

d'Oleire-Oltmanns, S., Marzolff, I., Peter, K., & Ries, J. (2012). Unmanned Aerial Vehicle (UAV) for Monitoring Soil Erosion in Morocco. Remote Sensing, 4, 3390-3416.

Dalponte, M., Bruzzone, L., & Gianelle, D. (2008). Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. Geoscience and Remote Sensing, IEEE Transactions on, 46, 1416-1427.

Dandois, J.P., & Ellis, E.C. (2010). Remote sensing of vegetation structure using computer vision. Remote Sensing, 2, 1157-1176.

Dandois, J.P., & Ellis, E.C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.

Dassot, M., Constant, T., & Fournier, M. (2011). The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Annals of Forest Science, 68, 959-974.

de Matías, J., Sanjosé, J.J.d., López-Nicolás, G., Sagüés, C., & Guerrero, J.J. (2009). Photogrammetric Methodology for the Production of Geomorphologic Maps: Application to the Veleta Rock Glacier (Sierra Nevada, Granada, Spain). Remote Sensing, 1, 829-841.

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B., & de Souza Jr, C. (2007). Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. Environmental Science and Policy, 10, 385-394.

Defries, R.S., Hansen, M.C., Townshend, J.R.G., Janetos, A.C., & Loveland, T.R. (2000). A new global 1-km dataset of percentage tree cover derived from remote sensing. Global Change Biology, 6, 247-254.

Desell, T., Bergman, R., Goehner, K., Marsh, R., VanderClute, R., & Ellis-Felege, S. (2013). Wildlife@Home: Combining Crowd Sourcing and Volunteer Computing to Analyze Avian Nesting Video. In, eScience (eScience), 2013 IEEE 9th International Conference on (pp. 107-115).

Dey, D., Mummert, L., & Sukthankar, R. (2012). Classification of plant structures from uncalibrated image sequences. In, Applications of Computer Vision (WACV), 2012 IEEE Workshop on (pp. 329-336).

Dolan, K., Masek, J.G., Huang, C., & Sun, G. (2009). Regional forest growth rates measured by combining ICESat GLAS and Landsat data. J. Geophys. Res., 114.

Dubayah, R.O., & Drake, J.B. (2000). Lidar Remote Sensing for Forestry. Journal of Forestry, 98, 44-46.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., & Chazdon, R.L. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. Journal of Geophysical Research, 115, G00E09.

Eisenbeiss, H. (2007). Applications of photogrammetric processing using an autonomous model helicopter. In, Revue Francaise de Photogrammetrie et de Teledetection. Symposium ISPRS Commission Technique I “Des capteurs a l’Imagerie”. Saint-Mande Cedex, France.

Eisenbeiss, H. (2009). UAV Photogrammetry. In. Zurich: University of Technology Dresden.

Erdody, T.L., & Moskal, L.M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sensing of Environment, 114, 725-737.

FAO, & JRC (2012). Global forest land-use change 1990–2005, by E.J. Lindquist, R. D'Annunzio, A. Gerrand, K. MacDicken, F. Achard, R. Beuchle, A. Brink, H.D. Eva, P. Mayaux, J. San-Miguel-Ayanz & H-J. Stibig. FAO Forestry Paper No. 169. Food and Agriculture Organization of the United Nations and European Commission Joint Research Centre. Rome, FAO. In.

Fei-Fei, L., & Perona, P. (2005). A Bayesian hierarchical model for learning natural scene categories. In, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (pp. 524-531 vol. 522).

Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L., & Carbonneau, P.E. (2013). Topographic structure from motion: a new development in photogrammetric measurement. Earth Surface Processes and Landforms, 38, 421-430.

Frazer, G.W., Magnussen, S., Wulder, M.A., & Niemann, K.O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR derived estimates of forest stand biomass. Remote Sensing of Environment, 115, 636-649.

Frolking, S., Palace, M.W., Clark, D.B., Chambers, J.Q., Shugart, H.H., & Hurtt, G.C. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. Journal of Geophysical Research, 114.

Gatziolis, D., & Andersen, H.E. (2008). A Guide to LiDAR Data Acquisition and Processing for Forests of the Pacific Northwest - General Technical Report PNW-GTR-768. In (p. 32 p). Portland, OR: USDA, Forest Service, Pacific Northwest Research Station.

Gauglitz, S., Höllerer, T., & Turk, M. (2011). Evaluation of Interest Point Detectors and Feature Descriptors for Visual Tracking. International Journal of Computer Vision, 94, 335-360.

Geerling, G., Labrador-Garcia, M., Clevers, J., Ragas, A., & Smits, A. (2007). Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. International Journal of Remote Sensing, 28, 4263-4284.

Gil, A., Mozos, O., Ballesta, M., & Reinoso, O. (2010). A comparative evaluation of interest point detectors and local descriptors for visual SLAM. Machine Vision and Applications, 21, 905-920.

Glennie, C. (2007). Rigorous 3D error analysis of kinematic scanning LIDAR systems. Journal of Applied Geodesy, 1, 147-157.

Goetz, S., & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management, 2, 231-244.

Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sensing of Environment, 108, 254-263.

Guanter, L., Segl, K., & Kaufmann, H. (2009). Simulation of Optical Remote-Sensing Scenes With Application to the EnMAP Hyperspectral Mission. Geoscience and Remote Sensing, IEEE Transactions on, 47, 2340-2351.

Habib, A., Bang, K.I., Kersting, A.P., & Lee, D.-C. (2009). Error budget of LiDAR systems and quality control of the derived data. Photogrammetric Engineering and Remote Sensing, 75, 1093-1108.

Habib, A., & Cheng, R. (2006). Surface Matching Strategy for Quality Control of LIDAR Data. Innovations in 3D Geo Information Systems, 67-83.

Hansen, M.C., Roy, D.P., Lindquist, E., Adusei, B., Justice, C.O., & Altstatt, A. (2008). A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. Remote Sensing of Environment, 112, 2495-2513.

Hardiman, B.S., Bohrer, G., Gough, C.M., Vogel, C.S., & Curtis, P.S. (2011). The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. Ecology, 92, 1818-1827.

Hartley, R., & Zisserman, A. (2000). Multiple View Geometry in Computer Vision. Cambridge: Cambridge University Press.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sensing, 4, 1573-1599.

Hilker, T., Coops, N.C., Hall, F.G., Black, T.A., Chen, B., Krishnan, P., Wulder, M.A., Sellers, P.J., Middleton, E.M., & Huemmrich, K.F. (2008). A modeling approach for upscaling gross ecosystem production to the landscape scale using remote sensing data. J. Geophys. Res., 113.

Hirschmugl, M., Ofner, M., Raggam, J., & Schardt, M. (2007). Single tree detection in very high resolution remote sensing data. Remote Sensing of Environment, 110, 533-544.

Holl, K.D., Zahawi, R.A., Cole, R.J., Ostertag, R., & Cordell, S. (2011). Planting Seedlings in Tree Islands Versus Plantations as a Large-Scale Tropical Forest Restoration Strategy. Restoration Ecology, 19, 470-479.

Hosoi, F., & Omasa, K. (2009). Detecting seasonal change of broad-leaved woody canopy leaf area density profile using 3D portable LIDAR imaging. Functional Plant Biology, 36, 998-1005.

Hopkinson, C., Chasmer, L., Lim, K., Treitz, P., & Creed, I. (2006). Towards a universal LIDAR canopy height indicator. Canadian Journal of Remote Sensing, 32, 139-152.

Houghton, R.A., Hall, F., & Goetz, S.J. (2009). Importance of biomass in the global carbon cycle. Journal of Geophysical Research, 114.

Huang, H., Gong, P., Cheng, X., Clinton, N., & Li, Z. (2009). Improving Measurement of Forest Structural Parameters by Co-Registering of High Resolution Aerial Imagery and Low Density LiDAR Data. Sensors, 9, 1541-1558.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sensing of Environment, 82, 397-416.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S.n., & Falkowski, M.J. (2012). Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. Remote Sensing of Environment, 123, 25-40.

Hunt, E., Hively, W., Daughtry, C., McCarty, G., Fujikawa, S., Ng, T., Tranchitella, M., Linden, D., & Yeol, D. (2009). Remote Sensing Crop Leaf Area Index Using Unmanned Airborne Vehicles. In, The Future of Land Imaging...Going Operational. Denver, Colorado.

Hunt, E.R., Cavigelli, M., Daughtry, C.S.T., McMurtrey, J.E., & Walthall, C.L. (2005). Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. Precision Agriculture, 6, 359-378.

Hunt, J., E. Raymond, Hively, W.D., Fujikawa, S., Linden, D., Daughtry, C.S., & McCarty, G. (2010). Acquisition of NIR-Green-Blue Digital Photographs from Unmanned Aircraft for Crop Monitoring. Remote Sensing, 2, 290-305.

Ientilucci, E.J., & Brown, S.D. (2003). Advances in wide-area hyperspectral image simulation. In (pp. 110-121).

James, M., Applegarth, L.J., & Pinkerton, H. (2012). Lava channel roofing, overflows, breaches and switching: insights from the 2008–2009 eruption of Mt. Etna. Bulletin of Volcanology, 74, 107-117.

James, M., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of Geophysical Research: Earth Surface (2003–2012), 117.

Javernick, L., Brasington, J., & Caruso, B. (2014). Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. Geomorphology, 213, 166-182.

Jung, K., Kaiser, S., Böhm, S., Nieschulze, J., & Kalko, E.K.V. (2012). Moving in three dimensions: effects of structural complexity on occurrence and activity of insectivorous bats in managed forest stands. Journal of Applied Ecology, 49, 523-531.

Kampe, T.U., Johnson, B.R., Kuester, M., & Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. Journal of Applied Remote Sensing, 4, 043510.

Ke, Y., Quackenbush, L.J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. Remote Sensing of Environment, 114, 1141-1154.

Kendal, D., Hauser, C.E., Garrard, G.E., Jellinek, S., Giljohann, K.M., & Moore, J.L. (2013). Quantifying Plant Colour and Colour Difference as Perceived by Humans Using Digital Images. PLoS ONE, 8, e72296.

Key, T., Warner, T.A., McGraw, J.B., & Fajvan, M.A. (2001). A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. Remote Sensing of Environment, 75, 100-112.

Koetz, B., Sun, G., Morsdorf, F., Ranson, K., Kneubuhler, M., Itten, K., & Allgower, B. (2007). Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization. Remote Sensing of Environment, 106, 449-459.

Kumar, N., Belhumeur, P., Biswas, A., Jacobs, D., Kress, W.J., Lopez, I., & Soares, J.B. (2012). Leafsnap: A Computer Vision System for Automatic Plant Species Identification. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), Computer Vision – ECCV 2012 (pp. 502-516): Springer Berlin Heidelberg.

Kuusk, A., Kuusk, J., & Lang, M. (2014). Modeling directional forest reflectance with the hybrid type forest reflectance model FRT. Remote Sensing of Environment, 149, 196-204.

Larjavaara, M., & Muller-Landau, H.C. (2013). Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. Methods in Ecology and Evolution, 4, 793-801.

Lazebnik, S., Schmid, C., & Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In, CVPR (pp. 2169-2178): IEEE.

Leberl, F., Irschara, A., Pock, T., Gruber, M., Scholz, S., & Wiechart, A. (2010). Point clouds: Lidar versus 3D vision. Photogrammetric Engineering and Remote Sensing, 76, 1123-1134.

Lefsky, M., & Cohen, W. (2003). Selection of Remotely Sensed Data. In M. Wulder, & S. Franklin (Eds.), Remote Sensing of Forest Environments (pp. 13-46): Springer US.

Lefsky, M., & McHale, M.R. (2008). Volume estimates of trees with complex architecture from terrestrial laser scanning. Journal of Applied Remote Sensing, 2, 023521.

Lefsky, M.A., Cohen, W.B., Parker, G.G., & Harding, D.J. (2002). Lidar remote sensing for ecosystem studies. Bioscience, 52, 20-30.

Li, J., & Allinson, N.M. (2008). A comprehensive review of current local features for computer vision. Neurocomputing, 71, 1771-1787.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LIDAR remote sensing of forest structure. Progress in Physical Geography, 27, 88-106.

Lingua, A., Marenchino, D., & Nex, F. (2009). Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. Sensors, 9, 3745-3766.

Lirman, D., Gracias, N., Gintert, B., Gleason, A., Reid, R., Negahdaripour, S., & Kramer, P. (2007). Development and application of a video-mosaic survey technology to document the status of coral reef communities. Environmental Monitoring and Assessment, 125, 59-73.

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., & Lejeune, P. (2013). A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. Forests, 4, 922-944.

Lovell, J., Jupp, D., Newnham, G., Coops, N., & Culvenor, D. (2005). Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. Forest Ecology and Management, 214, 398-412.

Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60, 91-110.

Lucieer, A., Robinson, S., & Turner, D. (2011). Unmanned aerial vehicle (UAV) remote sensing for hyperspatial terrain mapping of Antarctic moss beds based on structure from motion (SFM) point clouds. In, Proceedings of the 34th International Symposium on Remote Sensing of Environment (ISRSE34) (p. x).

Magdon, P., Fischer, C., Fuchs, H., & Kleinn, C. (2014). Translating criteria of international forest definitions into remote sensing image analysis. Remote Sensing of Environment, 149, 252-262.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 28, 1016-1031.

Magnussen, S., Naesset, E., & Gobakken, T. (2010). Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. Remote Sensing of Environment, 114, 700-712.

Mathews, A., & Jensen, J. (2013). Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. Remote Sensing, 5, 2164-2183.

McHale, M.R., Burke, I.C., Lefsky, M.A., Peper, P.J., & McPherson, E.G. (2009). Urban forest biomass estimates: is it important to use allometric relationships developed specifically for urban trees? Urban Ecosystems, 12, 95-113.

Mikolajczyk, K., & Schmid, C. (2004). Scale & affine invariant interest point detectors. International Journal of Computer Vision, 60, 63-86.

Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27, 1615-1630.

Mizunuma, T., Wilkinson, M., Eaton, E.L., Mencuccini, M., Morison, J.I.L., & Grace, J. (2013). The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern England. Functional Ecology, 27, 196-207.

Morgenroth, J., & Gomez, C. (2014). Assessment of tree structure using a 3D image analysis technique—A proof of concept. Urban Forestry & Urban Greening, 13, 198-203.

Morisette, J., Richardson, A., Knapp, A., Fisher, J., Graham, E., Abatzoglou, J., Wilson, B., Breshears, D., Henebry, G., & Hanes, J. (2008). Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. Frontiers in Ecology and the Environment, 7, 253-260.

Morsdorf, F., Nichol, C., Malthus, T., & Woodhouse, I.H. (2009). Assessing forest structural and physiological information content of multi-spectral LiDAR waveforms by radiative transfer modelling. Remote Sensing of Environment, 113, 2152-2163.

Mundt, J.T., Streutker, D.R., & Glenn, N.F. (2006). Mapping sagebrush distribution using fusion of hyperspectral and lidar classifications. Photogrammetric Engineering and Remote Sensing, 72, 47.

Næsset, E. (2009a). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment, 113, 148-159.

Næsset, E. (2009b). Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. Remote Sensing of Environment, 113, 2210-2223.

Næsset, E. (2011). Estimating above-ground biomass in young forests with airborne laser scanning. International Journal of Remote Sensing, 32, 473 - 501.

Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. Remote Sensing of Environment, 112, 3079-3090.

Nagendra, H. (2001). Using remote sensing to assess biodiversity. International Journal of Remote Sensing, 22, 2377-2400.

Newsam, S., & Yang, Y. (2007). Comparing global and interest point descriptors for similarity retrieval in remote sensed imagery. In, Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems. Seattle, Washington: ACM.

Ni, W., Ranson, K.J., Zhang, Z., & Sun, G. (2014). Features of point clouds synthesized from multi-view ALOS/PRISM data and comparisons with LiDAR data in forested areas. Remote Sensing of Environment, 149, 47-57.

Nilsback, M.-E. (2009). An automatic visual Flora - segmentation and classication of flower images. In, Robotics Research Group, Department of Engineering Science: University of Oxford.

Ofner, M., Hirschmugl, M., Raggam, H., & Schardt, M. (2006). 3D Stereo Mapping by Means of UltracamD Data. In, Workshop on 3D Remote Sensing in Forestry. Vienna, Austria.

Oh, J., Toth, C.K., & Grejner-Brzezinska, D.A. (2011). Automatic Georeferencing of Aerial Images Using Stereo High-Resolution Satellite Images. Photogrammetric Engineering and Remote Sensing, 77, 1157-1168.

Ohayon, S., Willigen, R., Wagner, H., Katsman, I., & Rivlin, E. (2006). On the barn owl's visual pre-attack behavior: I. Structure of head movements and motion patterns. Journal of Comparative Physiology A, 192, 927-940.

Oliva, A., & Torralba, A. (2001). Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. International Journal of Computer Vision, 42, 145-175.

Omerčević, D., Perko, R., Targhi, A., Eklundh, J., & Leonardis, A. (2008). Vegetation segmentation for boosting performance of MSER feature detector. In J. Perš (Ed.), Computer Vision Winter Workshop (pp. 17-23). Moravske Toplice, Slovenia, February 4–6 2008.

Packalén, P., Suvanto, A., & Maltamo, M. (2009). A two stage method to estimate speciesspecific growing stock. Photogrammetric Engineering and Remote Sensing, 75, 1451-1460.

Parker, G., Harding, D., & Berger, M. (2004). A portable LIDAR system for rapid determination of forest canopy structure. Journal of Applied Ecology, 41, 755-767.

Parker, G.G., & Russ, M.E. (2004). The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. Forest Ecology and Management, 189, 307-315.

Poli, D., Remondino, F., Angiuli, E., & Agugiaro, G. (2014). Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction. ISPRS Journal of Photogrammetry and Remote Sensing.

Popescu, S., & Wynne, R. (2004). Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering and Remote Sensing, 70, 589-604.

Rango, A., Laliberte, A., Herrick, J.E., Winters, C., Havstad, K., Steele, C., & Browning, D. (2009). Unmanned aerial vehicle-based remote sensing for rangeland assessment, monitoring, and management. Journal of Applied Remote Sensing, 3, 15.

Rees, W.G. (2001). Physical Principles of Remote Sensing Second Edition. Cambridge: Cambridge University Press.

Richardson, A., Jenkins, J., Braswell, B., Hollinger, D., Ollinger, S., & Smith, M. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. Oecologia, 152, 323-334.

Richardson, A.D., Braswell, B., Hollinger, D.Y., Jenkins, J.C., & Ollinger, S.V. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications, 19, 1417-1428.

Rosnell, T., & Honkavaara, E. (2012). Point Cloud Generation from Aerial Image Data Acquired by a Quadrocopter Type Micro Unmanned Aerial Vehicle and a Digital Still Camera. Sensors, 12, 453-480.

Roy, D.P., Wulder, M.A., Loveland, T.R., C.E, W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment, 145, 154-172.

Ryu, Y., Baldocchi, D.D., Verfaillie, J., Ma, S., Falk, M., Ruiz-Mercado, I., Hehn, T., & Sonnentag, O. (2010). Testing the performance of a novel spectral reflectance sensor, built with light emitting diodes (LEDs), to monitor ecosystem metabolism, structure and function. Agricultural and Forest Meteorology, 150, 1597-1606.

Ryu, Y., Lee, G., Jeon, S., Song, Y., & Kimm, H. (2014). Monitoring multi-layer canopy spring phenology of temperate deciduous and evergreen forests using low-cost spectral sensors. Remote Sensing of Environment, 149, 227-238.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings of the National Academy of Sciences, 108, 9899-9904.

Schimel, D., Keller, M., Berukoff, S., Kao, B., Loescher, H., Powell, H., Kampe, T., Moore, D., & Gram, W. (2011). NEON Science Infrastructure. In, NEON 2011 Science Strategy (pp. 36-38).

Schott, J., Gerace, A., Brown, S., Gartley, M., Montanaro, M., & Reuter, D.C. (2012). Simulation of Image Performance Characteristics of the Landsat Data Continuity Mission (LDCM) Thermal Infrared Sensor (TIRS). Remote Sensing, 4, 2477-2491.

Schott, J.R., Brown, S.D., Raqueno, R.V., Gross, H.N., & Robinson, G. (1999). Advanced synthetic image generation models and their application to multi/hyperspectral algorithm development. In (pp. 211-220).

Schwind, P., Suri, S., Reinartz, P., & Siebert, A. (2010). Applicability of the SIFT operator to geometric SAR image registration. International Journal of Remote Sensing, 31, 1959 - 1980.

Seidel, D., Fleck, S., Leuschner, C., & Hammett, T. (2011). Review of ground-based methods to measure the distribution of biomass in forest canopies. Annals of Forest Science, 1-20.

Seielstad, C., Stonesifer, C., Rowell, E., & Queen, L. (2011). Deriving Fuel Mass by Size Class in Douglas-fir (Pseudotsuga menziesii) Using Terrestrial Laser Scanning. Remote Sensing, 3, 1691-1709.

Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. In, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR06) (pp. 519-526).

Shimabukuro, Y.E., Santos, J.R.d., Formaggio, A.R., Duarte, V., & Rudorff, B.F.T. (2014). The Brazilian Amazon Monitoring Program: PRODES and DETER Projects. In F. Achard, & M.C. Hansen (Eds.), Global Forest Monitoring from Earth Observation: CRC Press.

Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 59, 85-101.

Snavely, N., Seitz, S., & Szeliski, R. (2006). Photo tourism: exploring photo collections in 3D. In (pp. 835-846): ACM New York, NY, USA.

Snavely, N., Simon, I., Goesele, M., Szeliski, R., & Seitz, S.M. (2010). Scene Reconstruction and Visualization from Community Photo Collections. Proceedings of the IEEE, 98, 1370-1390.

Spitzen, J., Spoor, C.W., Grieco, F., ter Braak, C., Beeuwkes, J., van Brugge, S.P., Kranenbarg, S., Noldus, L.P.J.J., van Leeuwen, J.L., & Takken, W. (2013). A 3D Analysis of Flight Behavior of <italic>Anopheles gambiae sensu stricto</italic> Malaria Mosquitoes in Response to Human Odor and Heat. PLoS ONE, 8, e62995.

St. Onge, B., Hu, Y., & Vega, C. (2008). Mapping the height and above-ground biomass of a mixed forest using lidar and stereo Ikonos images. International Journal of Remote Sensing, 29, 1277-1294.

St. Onge, B., Jumelet, J., Cobello, M., & Vega, C. (2004). Measuring individual tree height using a combination of stereophotogrammetry and lidar. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 34, 2122-2130.

Strecha, C., Von Hansen, W., Van Gool, L., Fua, P., & Thoennessen, U. (2009). On benchmarking camera calibration and multi-view stereo for high resolution imagery. In, Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR09). Miami, June 2009..

Szeliski, R. (2011). Computer Vision. Springer Algorithms and Applications.

Tao, W., Lei, Y., & Mooney, P. (2011). Dense point cloud extraction from UAV captured images in forest area. In, 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services. Fuzhou: June 29 2011 - July 1 2011.

Thomas, R.Q., Hurtt, G.C., Dubayah, R., & Schilz, M.H. (2008). Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain. Canadian Journal of Remote Sensing, 34, S351-S363.

Treuhaft, R.N., Law, B.E., & Asner, G.P. (2004). Forest Attributes from Radar Interferometric Structure and Its Fusion with Optical Remote Sensing. Bioscience, 54, 561-571.

Triggs, B., McLauchlan, P., Hartley, R., & Fitzgibbon, A. (1999). Bundle adjustment- A modern synthesis. Lecture Notes in Computer Science, 298-372.

Turner, D., Lucieer, A., & Watson, C. (2012). An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SFM) Point Clouds. Remote Sensing, 4, 1392-1410.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. Trends in Ecology & Evolution, 18, 306-314.

Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeological Prospection, 18.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., & Clawges, R.M. (2008). Lidar: shedding new light on habitat characterization and modeling. Frontiers in Ecology and the Environment, 6, 90-98.

Vogel, J., & Schiele, B. (2004). A Semantic Typicality Measure for Natural Scene Categorization. In C. Rasmussen, H. Bülthoff, B. Schölkopf, & M. Giese (Eds.), Pattern Recognition (pp. 195-203): Springer Berlin Heidelberg.

Vogler, A., Eisenbeiss, H., Aulinger-Leipner, I., & Stamp, P. (2009). Impact of topography on cross-pollination in maize (Zea mays L.). European Journal of Agronomy, 31, 99-102.

Wallace, L., Lucieer, A., Watson, C., & Turner, D. (2012). Development of a UAV-LiDAR System with Application to Forest Inventory. Remote Sensing, 4, 1519-1543.

Wallace, L., Musk, R., & Lucieer, A. (2014). An Assessment of the Repeatability of Automatic Forest Inventory Metrics Derived From UAV-Borne Laser Scanning Data. Geoscience and Remote Sensing, IEEE Transactions on, 52, 7160-7169.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., & Reynolds, J.M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314.

Wolf, P.R., Dewitt, B.A., & Wilkinson, B.E. (2014). Elements of Photogrammetry with Applications in GIS. (4th ed.). New York: McGraw Hill Education.

Wright, S.J., Carrasco, C., Calderón, O., & Paton, S. (1999). The El Niño southern oscillation, variable fruit production, and famine in a tropical forest. Ecology, 80, 1632-1647.

Wu, C. (2010). SiftGPU: A GPU implementation of scale invariant feature transform (SIFT) http://www.cs.unc.edu/~ccwu/siftgpu/. In.

Yang, X., & Lo, C. (2000). Relative radiometric normalization performance for change detection from multi-date satellite images. Photogrammetric Engineering and Remote Sensing, 66, 967-980.

Yang, Y., & Newsam, S. (2008). Comparing SIFT descriptors and gabor texture features for classification of remote sensed imagery. In, Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on (pp. 1852-1855).

Yang, Y., & Newsam, S. (2013). Geographic Image Retrieval Using Local Invariant Features. Geoscience and Remote Sensing, IEEE Transactions on, 51, 818-832.

Zhang, D., & Lu, G. (2004). Review of shape representation and description techniques. Pattern Recognition, 37, 1-19.

Zhang, X., Friedl, M., Schaaf, C., Strahler, A., Hodges, J., Gao, F., Reed, B., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. Remote Sensing of Environment, 84, 471-475.

Zhang, Z., Zhang, Y., Ke, T., & Guo, D. (2009). Photogrammetry for First Response in Wenchuan Earthquake. Photogrammetric Engineering and Remote Sensing, 75, 510-513.

Zhao, K., & Popescu, S. (2009). Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. Remote Sensing of Environment, 113, 1628-1645.

Zolkos, S.G., Goetz, S.J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sensing of Environment, 128, 289-298.

Chapter 2: High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision

Abstract

High spatial resolution three-dimensional (3D) measurements of vegetation by remote sensing are advancing ecological research and environmental management. However, substantial economic and logistical costs limit this application, especially for observing phenological dynamics in ecosystem structure and spectral traits. Here we demonstrate a new aerial remote sensing system enabling routine and inexpensive aerial 3D measurements of canopy structure and spectral attributes, with properties similar to those of LIDAR, but with RGB (red-green-blue) spectral attributes for each point, enabling high frequency observations within a single growing season. This 'Ecosynth' methodology applies photogrammetric 'structure from motion' computer vision algorithms to large sets of highly overlapping low altitude (< 120 m) aerial photographs acquired using off-the-shelf digital cameras mounted on an inexpensive (< USD$4000), lightweight (< 2 kg), hobbyist-grade unmanned aerial system (UAS). Ecosynth 3D point clouds with densities of 30 – 67 points m^{-2} were produced using commercial computer vision software from digital photographs acquired repeatedly by UAS over three 6.25 ha (250 m x 250 m) Temperate Deciduous forest sites in Maryland USA. Ecosynth point clouds were georeferenced with a precision of 1.2 – 4.1 m horizontal radial root mean square error (RMSE) and 0.4 – 1.2 m vertical RMSE. Understory digital terrain models (DTMs) and canopy height models (CHMs) were generated from leaf-on and leaf-off point clouds using procedures commonly applied to LIDAR point clouds. At two sites, Ecosynth CHMs were strong predictors of field-measured tree heights (R^2 = 0.63 to 0.84) and were highly correlated with a LIDAR CHM (R = 0.87) acquired 4 days earlier,

though Ecosynth-based estimates of aboveground biomass and carbon densities included significant errors (31 – 36% of field-based estimates). Repeated scanning of a 50 m x 50 m forested area at six different times across a 16 month period revealed ecologically significant dynamics in canopy color at different heights and a structural shift upward in canopy density, as demonstrated by changes in vertical height profiles of point density and relative RGB brightness. Changes in canopy relative greenness were highly correlated ($R^2 = 0.88$) with MODIS NDVI time series for the same area and vertical differences in canopy color revealed the early green up of the dominant canopy species, *Liriodendron tulipifera*, strong evidence that Ecosynth time series measurements can capture vegetation structural and spectral phenological dynamics at the spatial scale of individual trees. The ability to observe individual canopy phenology in 3D at high temporal resolutions represents a breakthrough in forest ecology. Inexpensive user-deployed technologies for multispectral 3D scanning of vegetation at landscape scales (< 1 km^2) heralds a new era of participatory remote sensing by field ecologists, community foresters and the interested public.

1.0 Introduction

High spatial resolution remote sensing of vegetation structure in three-dimensions (3D) has become an important tool for a broad range of scientific and environmental management applications, including national and local carbon accounting (Frolking et al. 2009; Houghton et al. 2009; Goetz and Dubayah 2011), fire spread and risk modeling (Andersen et al. 2005; Skowronski et al. 2011), commercial and scientific forestry (Næsset and Gobakken. 2008), ecosystem modeling (Thomas et al. 2008; Zhao and Popescu 2009; Antonarakis et al. 2011), quantitative assessments of habitat suitability

and biodiversity (Vierling et al. 2008; Jung et al. 2012) and serves as a core data product of the National Ecological Observation Network (NEON; Schimel et al. 2011). Recent advances in 3D remote sensing have combined 3D measurements with rich spectral information, yielding unprecedented capabilities for observing biodiversity and ecosystem functioning (Asner and Martin 2009). Remote sensing systems with high temporal resolutions are driving similar advances in understanding ecosystem dynamics in forests (Richardson et al. 2009) and globally (Zhang and Goldberg 2011), including the response of terrestrial ecosystems to changes in climate and land use (Morisette et al. 2008; Frolking et al. 2009; Richardson et al. 2009), yet no single instrument is technically or logistically capable of combining structural and spectral observations at high temporal and spatial resolutions. Here we demonstrate an inexpensive user-deployed aerial remote sensing system that enables high spatial resolution 3D multispectral observations of vegetation at high temporal resolutions, and discuss its prospects for advancing the remote sensing of forest structure, function and dynamics.

Tree heights, generally in the form of canopy height models (CHM), are the most common remotely sensed 3D vegetation measurements. CHMs can be produced using stereo-pair and multiple-stereo photogrammetry applied to images acquired from aircraft and satellites (Hirschmugl et al. 2007; St. Onge et al. 2008) and active synthetic aperture radar (SAR) sensors (Treuhaft et al. 2004), but are now most commonly produced using active LIDAR remote sensing (Light Detection and Ranging). LIDAR CHMs with precisions of 0.2 – 2 m can be produced across forest types and acquisition settings (i.e., altitude, point density, etc.; Andersen et al. 2006; Wang and Glenn 2008) based on the return times of laser pulses reflected from canopy surfaces and the ground, by generating

models of understory terrain elevations (digital terrain models; DTM) and top canopy surface heights, which are then subtracted (Dubayah and Drake 2000; Popescu et al. 2003). Canopy heights and other metrics of vertical structure are useful for estimating aboveground biomass and carbon density (Lefsky et al. 2002; Goetz and Dubayah 2011), biomass change (from multiple LIDAR missions; Hudak et al. 2012), fire risk (Andersen et al. 2005; Skowronski et al. 2011), and for individual tree extraction by species (Falkwoski et al. 2008; Vauhkonen et al. 2008) among many other scientific and management applications.

While conventional airborne LIDAR acquisitions have become less expensive over time, they remain very costly for researchers and other end-users, especially if required at high spatial resolution over a few small areas or at high temporal frequencies (Gonzalez et al. 2010; Schimel et al. 2011). When applied over large spatial extents (e.g., > hundreds of square kilometers) LIDAR can be used to map aboveground biomass at a cost of USD$0.05 - USD$0.20 per hectare (Asner 2009). However, typical commercial aerial LIDAR acquisitions often cost a minimum of USD$20,000 per flight regardless of study area size (Erdody and Moskal 2010), representing a significant barrier to widespread application, especially for local environmental management and in ecological field studies based on annual or more frequent observations at numerous small sites or sampling plots (e.g., Holl et al. 2011). Even LIDAR satellite missions require local calibration data from multiple small sampling locations dispersed across spatial scales (Defries et al. 2007; Frolking et al. 2009; Dubayah et al. 2010).

The fusion of active-3D and optical-image remote sensing datasets has become increasingly common for the mapping of vegetation structural and spectral traits for

applications including the measurement of aboveground biomass and carbon, identifying individual species, and modeling the spatial heterogeneity of vegetation biochemistry (Turner et al. 2003; Anderson et al. 2008; Vitousek et al. 2009; Ke et al. 2010). However, the need to combine data from different sensors presents multiple challenges to both analysis and application, including areas of no data, spatial misalignment, and the need to reduce the quality of one dataset to match the other, such as coarsening LIDAR structural observations to match optical image observations; Hudak et al. 2002; Geerling et al. 2007; Mundt et al. 2006; Anderson et al. 2008). Recent advances in 3D remote sensing have combined active 3D and spectral measurements in a calibrated sensor package (Asner and Martin 2009). Yet despite their high utility, integrated fusion instruments remain too costly to be deployed at the frequent time intervals needed to capture vegetation temporal dynamics at the same location within a growing season (Kampe et al. 2010; Schimel et al. 2011).

To overcome the cost and logistical barriers to routine and frequent acquisition of high spatial resolution 3D datasets, three rapidly emerging technologies can be combined: low-cost, hobbyist-grade Unmanned Aircraft Systems (UAS), high speed consumer digital cameras (continuous frame rates >1 s^{-1}), and automated 3D reconstruction algorithms based on computer vision. Recent advances in hobbyist grade UAS capable of autonomous flight make it possible for an individual to obtain over the Internet a small (< 1 m diameter), light-weight (< 2 kg), and relatively low-cost (< USD$4000) aerial image acquisition platform that can be programmed to fly a specified route over an area at a fixed altitude (e.g., 100 m above the ground). Dandois and Ellis (2010) demonstrated that high spatial resolution 3D 'point cloud' models of vegetation structure and color

(RGB; red-green-blue) can be produced by applying Structure from Motion computer vision algorithms (SFM; Snavely et al. 2010) to sets of regular digital photographs acquired with an off-the-shelf digital camera deployed on a kite, without any information about sensor position and orientation in space (Dandois and Ellis 2010; Snavely et al. 2010). While this early 'Ecosynth' system proved capable of yielding useful data, kite platforms proved incapable of supporting the consistent repeated acquisitions of overlapping high quality images needed to observe dynamics in vegetation structure and color at high spatial resolutions in 3D over larger areas.

This study will demonstrate that by enhancing Ecosynth methods using automated UAS image acquisition techniques, high spatial resolution multispectral 3D datasets can be repeatably and consistently produced, thereby enabling the structural and spectral dynamics of forest canopies to be observed in 3D; a major advance in the remote sensing of forest ecosystems. Ecosynth methods encompass the full process and suite of hardware and software used to observe vegetation structural and spectral traits from ordinary digital cameras using computer vision. Ecosynth methods are not presented as a replacement for remote sensing systems designed to map large extents, but rather as an inexpensive user-deployed system for detailed observations across local sites and landscapes at scales generally less than 1 km^2, much like ground-based Portable Canopy LIDAR (PCL; Parker et al. 2004; Hardiman et al. 2011), or web-cam phenology imaging systems deployed at carbon flux towers (PhenoCam; Richardson et al. 2009; Mizunuma et al. 2013). Nevertheless, the general utility and maturity of Ecosynth methods for routine and inexpensive forest measurements on demand will be demonstrated by comparing these with estimates of understory terrain, canopy height, and forest

aboveground biomass density produced by field and LIDAR methods across three > 6 ha forest study sites. The unprecedented ability of Ecosynth methods to simultaneously observe vegetation structural and spectral dynamics at high spatial resolutions is then demonstrated by comparing vertical profiles of vegetation structure (Parker and Russ 2004) and RGB relative brightness (Richardson et al. 2009; Mizunuma et al. 2013) acquired at six times across the Northern Temperate growing season to data from vegetation stem maps, discrete return LIDAR, and a MODIS NDVI time series.

1.1 Computer vision for remote sensing

Automated photogrammetric systems based on computer vision SFM algorithms (Snavely et al. 2008) enable the production of geometrically precise 3D point cloud datasets based entirely on large sets of overlapping digital photographs taken from different locations (Dandois and Ellis 2010; Dey et al. 2012; Rosnell and Honkavaara 2012). SFM relies on photogrammetric methods that have already been used for estimating tree height from overlapping images acquired using large-format, photogrammetric-grade cameras coupled with flight time GPS and IMU data, including automated feature extraction, matching and bundle adjustment (Ofner et al. 2006; Hirschmugl et al. 2007), and these methods have been discussed as a viable alternative to LIDAR for 3D forestry applications (Leberl et al. 2010). However, SFM differs from prior photogrammetric applications in that camera position and orientation data that are conventionally acquired using GPS and IMU instruments carried by the aircraft are removed from the 3D modeling equation, and instead the 3D reconstruction of surface feature points is determined automatically based on the inherent 'motion' of numerous overlapping images acquired from different locations (Snavely et al. 2008). The result is

an extremely simple remote sensing instrument: an ordinary digital camera taking highly overlapping images while moving around or along objects.

SFM techniques have already proved successful for accurate 3D modeling of built structures, bare geological substrates, and fine-spatial scale individual plant structure (de Matías et al. 2009; Snavely et al. 2010; Dey et al. 2012; Harwin and Lucieer 2012). SFM has been applied to generate 3D surface models of open fields, forests and trees from aerial images acquired from a remote-controlled multi-rotor aircraft (Tao et al. 2011; Rosnell and Honkavaara 2012). Recently, Wallace et al. (2012) used SFM algorithms to improve the calculation of sensor position and orientation on a lightweight UAS (≈ 5 kg with payload) carrying a mini-LIDAR sensor with lightweight GPS and new micro-electromechanical system (MEMS) based IMU equipment (2.4 kg), finding sub-meter horizontal and vertical spatial accuracies of ground targets (0.26 m and 0.15 m, respectively). That study found low variance (0.05 m - 0.25 m) of manually extracted individual tree height measurements from the LIDAR point cloud but did not compare these with field measured tree heights.

1.2 UAS for remote sensing

UAS are increasingly being deployed for low-cost, on-demand aerial photography and photogrammetry applications (Rango et al. 2009; Hunt et al. 2010; Harwin and Lucieer 2012). Rosnell and Honkavaara (2012) used an autonomous multirotor aircraft to take aerial photos in a grid pattern to generate orthomosaics and land surface elevation models using photogrammetry and computer vision software. Lin et al. (2011) recently explored the deployment of LIDAR sensors on relatively small UAS (11.5 kg with platform, battery and payload) suggesting a technology useful for measuring forest structure, but

without demonstrating the production of canopy height or other forestry measures. As both conventional LIDAR and photogrammetric techniques require precise measurements of sensor position and orientation during flight, these techniques require high-accuracy global positioning systems (GPS) and inertial monitoring units (IMU), both of which are relatively expensive and heavy instruments (> 10 kg) that tend to limit applications to the use of relatively large UASs (> 10 kg) and higher altitudes (> 120 m), invoking logistical and regulatory requirements similar to those of conventional manned aircraft.

2.0 Materials and methods

2.1 Study areas

Research was carried out across three 6.25 ha (250 m x 250 m) forest research study sites in Maryland USA; two areas on the campus of the University of Maryland Baltimore County (UMBC; 39°15'18"N 76°42'32"W) and one at the Smithsonian Environmental Research Center in Edgewater Maryland (SERC; 38°53'10"N 76°33'51"W). UMBC sites are centered on and expanded from the smaller study sites described by Dandois and Ellis (2010).

The first UMBC study site, 'Knoll' centers on a forested hill surrounded by turfgrass and paved surfaces, peaking at about ≈ 60 m ASL (above mean sea level) and gradually descending by 5 to 20 m. Forest is composed of a mixed-age canopy (mean canopy height 25 m, max. 42 m) dominated by American beech (*Fagus grandifolia*), oak (*Quercus spp.*), and hickory (*Carya spp.*) but also including several large mature white ash (*Fraxinus americana*) and tulip-poplar (*Liriodendron tulipifera*). The second UMBC study site, 'Herbert Run', consists of a remnant forest patch similar in size and composition (mean canopy height 20 m, max. 34 m) to the Knoll (elevation 55 m ASL)

but steeply sloping (up to 50% grade) down to a riparian forest along a small stream (Herbert Run; 40 m ASL) and back up to a road running parallel to the stream. The riparian forest canopy consists mostly of an even-aged stand of black locust (*Robinia pseudoacacia*) overstory with black cherry (*Prunus serotina*) understory along the steep stream banks, with honey locust (*Gleditsia triacanthos*) and green ash (*Fraxinus pennsylvanica*) becoming dominant in closest proximity to the stream.

The 'SERC' study site is located approximately at the center of the 'Big Plot' at the Smithsonian Environmental Research Center in Edgewater, Maryland that has been the long-term focus of a variety of forest ecology and remote sensing studies (Parker and Russ 2004; McMahon et al. 2010). The site is comprised of floodplain with a gradual slope (8 % mean grade) from a small hill (≈ 19 m ASL) at the north to a riparian area (≈ 0 m ASL) to the east and south. The canopy is dominated by tulip-poplar, American beech, and several oak (*Quercus spp.*) species in the overstory (mean canopy height 37 m, max. 50 m).

2.2 Forestry field methods

At UMBC sites, a 25 m x 25 m subplot grid was staked out within forested areas using a Sokkia Set 5A Total Station and Trimble TSC2 Data Logger based off of the local geodectic survey network (0.25 m horizontal radial RMSE, 0.07 m vertical RMSE; WGS84 UTM Zone 18N datum). Tree location, species, and diameter at breast height (DBH) of trees greater than 1 cm DBH were hand mapped within the subplot grid between June 2012 and March 2013. Tree heights were measured by laser hypsometer during leaf-off conditions over the same period for the five largest trees per subplot, based on DBH, as the average of three height measurements taken at approximately 120°

intervals around each tree at an altitude angle of < 45°. Subplot canopy height was then estimated as the mean height of the 5 tallest trees, i.e., average maximum height.

Field data for SERC were collected as part of a long-term forest inventory and monitoring program as described by McMahon et al. (2010). In that project, individual trees greater than 1 cm DBH were mapped to a surveyed 10 m x 10 m subplot grid using a meter tape placed on the ground and were identified to species. For the current study, a sample of field measured tree heights were obtained by overlaying a 25 m x 25 m subplot grid across the existing stem map in GIS and selecting the five largest trees per subplot based on DBH. During winter 2013, tree heights were measured as described above in 30 of the 100 25 m x 25 m subplots: 26 in randomly selected subplots and 4 in a group of subplots that comprise a 50 m x 50 m subset area.

2.3 Aerial LIDAR

LIDAR data covering UMBC sites were acquired in 2005 by a local contractor for the Baltimore County Office of Information Technology with the goal of mapping terrain at high spatial resolution across Baltimore County MD, USA. The collection used an Optech ALTM 2050 LIDAR with Airborne GPS and IMU under leaf-off conditions in the spring of 2005 (2005-03-18 – 2005-04-15; ≈ 800 – 1,200 m above ground surface; ≈ 140 knots airspeed; 36 Hz scan frequency; 20° scan width half angle; 50 kHz pulse rate; ≈ 150 m swath overlap; mean point density 1.5 points m^{-2}; NAD83 Harn Feet horizontal datum; NAVD88 Feet vertical datum). More recent LIDAR data for UMBC sites was not available (Baltimore County has a 10 year LIDAR collection plan), so the 2005 LIDAR dataset represents the only existing 3D forest canopy dataset at these sites. Airborne LIDAR data for SERC were collected 2011-10-05 by the NASA GSFC G-LiHT

(Goddard LIDAR-Hyperspectral-Thermal; Cook et al. 2012) remote sensing fusion platform (350 m above ground surface; 110 knots airspeed; 300 kHz pulse repetition frequency; 150 kHz effective measurement rate; 30° scan width half angle; 387 m swath width at 350 m altitude; mean point density 78 points m^{-2}; WGS84 UTM Zone 18N horizontal coordinate system; GEOID09 vertical datum; data obtained and used with permission from Bruce Cook, NASA GSFC on 2012-02-22).

2.4 Ecosynth—computer vision remote sensing

The term 'Ecosynth' is used here and in prior research (Dandois and Ellis 2010) to describe the entire processing pipeline and suite of hardware involved in generating ecological data products (e.g., canopy height models (CHMs), aboveground biomass (AGB) estimates, and canopy structural and spectral vertical profiles) and is diagrammed in Figure 1. The Ecosynth method combines advances and techniques from many areas of research, including computer vision structure from motion, UAS, and LIDAR point cloud data processing.

2.4.1 Image acquisition using UAS

An autonomously flying, hobbyist-grade multi-rotor helicopter, 'Mikrokopter Hexakopter' (Figure 1a; HiSystems GmbH, Moormerland, Germany; http://www.mikrokopter.de) was purchased as a kit, constructed, calibrated and programmed for autonomous flight according to online instructions. The flying system included a manufacturer-provided wireless telemetry downlink to a field computer, enabling real-time ground monitoring of aircraft altitude, position, speed, and battery life.

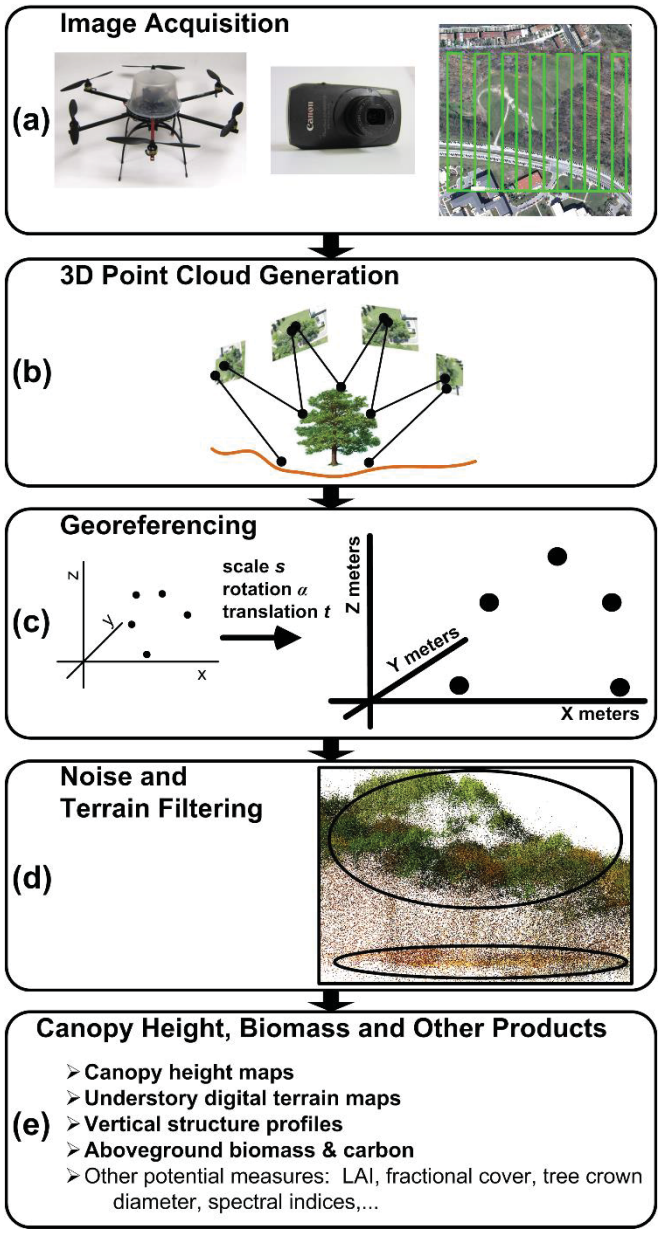


Figure 1: Workflow for Ecosynth remote sensing (details in Appendix A1.1).

Image acquisition flights were initiated at the geographic center of each study site because Hexakopter firmware restricted autonomous flight within a 250 m radius of takeoff, in compliance with German laws. This required manual piloting of the Hexakopter through a canopy gap at the Knoll and SERC sites; flights at Herbert Run were initiated from an open field near study site center. Flights were programmed to a predetermined square parallel flight plan designed to cover the study site plus a 50 m

buffer area added to avoid edge effects in image acquisitions, by flying at a fixed altitude approximately 40 m above the peak canopy height of each study site. Once the Hexakopter reached this required altitude, as determined by flight telemetry, automated flight was initiated by remote control. Flight paths were designed to produce a minimum photographic side overlap of 40 % across UMBC sites and 50 % at SERC owing to higher wind prevalence at that study site at the time of acquisition; forward overlap was > 90 % for all acquisitions.

A Canon SD4000 point-and-shoot camera was mounted under the Hexakopter to point at nadir and set to "Continuous Shooting mode" to collect 10 megapixel resolution photographs continuously at a rate of 2 frames s^{-1}. Camera focal length was set to "Infinity Focus" (≈ 4.90 mm) and exposure was calibrated to an 18 % grey camera target in full sun with a slowest shutter speed of 1/800 seconds. Images were acquired across each study site under both leaf-on and leaf-off conditions as described in Appendix 1.2. Two leaf-on acquisitions were produced at the Knoll study site to assess repeatability of height measurements and spectral changes caused by Fall leaf senescence (Leaf-on 2; Appendix 1.2). At SERC, four additional data sets were collected across a 16 month period to capture the structural and spectral attributes of the canopy at distinct points throughout the growing season (winter leaf-off, early spring, spring green-up, summer mature green, early fall leaf-on, senescing). Upon completion of its automated flight plan, the aircraft returned to the starting location and was manually flown vertically down to land.

2.4.2 3D point cloud generation using SFM

Multi-spectral (red-green-blue, RGB) three-dimensional (3D) point clouds were generated automatically from the sets of aerial photographs described in Appendix 1.2 using a purchased copy of Agisoft Photoscan, a commercial computer vision software package (http://www.agisoft.ru; v0.8.4 build 1289). Photoscan uses proprietary algorithms that are similar to, but not identical with, those of Bundler (Personal email communication with Dmitry Semyonov, Agisoft LLC, 2010-12-01) and was used for its greater computational efficiency over the open source Bundler software used previously for vegetation point cloud generation (estimated at least 10 times faster for photo sets >2000; Dandois and Ellis 2010). Photoscan has already been used for 3D modeling of archaeological sites from kite photos (Verhoeven 2011) and has been proposed for general image-based surface modeling applications (Remondino et al. 2011).

Prior to running Photoscan, image sets were manually trimmed to remove photos from the take-off and landing using the camera time stamp and the time stamp of GPS points recorded by the Mikrokopter. Photoscan provides a completely automated computer vision SFM pipeline, taking as input a set of images and automatically going through the steps of feature identification, matching and bundle adjustment. To generate each 3D RGB point, Photoscan performs several tasks as part of an automated computer vision SFM pipeline (Verhoeven 2011). This is accomplished by automatically extracting 'keypoints' from individual photos, identifying 'keypoint matches' among photos (e.g., Lowe 2004), and then using bundle adjustment algorithms to estimate and optimize the 3D location of feature correspondences together with the location and

orientation of cameras and camera internal parameters (Triggs et al. 1999; Snavely et al. 2008). SFM process is described in detail in Section 1.1.3 of Chapter 1.

Photoscan was run using the 'Align Photos' tool with settings: 'High Accuracy' and 'Generic Pair Pre-selection'. The 'Align Photos' tool automatically performs the computer vision structure from motion process as described above, but using proprietary algorithms. According the manufacturer's description, the 'High Accuracy' setting provides a better solution of camera position, but at the cost of greater computation time. Similarly, the 'Generic Pair Pre-selection' setting uses an initial low accuracy assessment to determine which photos are more likely to match, reducing computation time. After this, no other input is provided by the user until processing is complete, at which time the user exports the forest point cloud model into an ASCII XYZRGB file and the camera points into an ASCII XYZ file.

Photoscan was installed on a dual Intel Xeon X5670 workstation (12 compute cores) with 48GB of RAM, which required 2-5 days of continuous computation to complete the generation of a single point cloud across each study site, depending roughly on the size of the input photo set (Appendix 1.2). Point clouds thus produced consisted of a set of 3D points in an arbitrary but internally consistent geometry, with RGB color extracted for each point from input photos, together with the 3D location of the camera for each photo together with its camera model, both intrinsic (e.g., lens distortion, focal length, principle point) and extrinsic (e.g., XYZ location, rotational pose along all three axes), in the same coordinate system as the entire point cloud (Figure 1b).

2.4.3 SFM point cloud georeferencing and post-processing

Ground control point (GCP) markers (five-gallon orange buckets) were positioned across sites prior to image acquisition in configurations recommended by Wolf and Dewitt (2000). The XYZ locations of each GCP marker were measured using a Trimble GeoXT GPS with differential correction to within 1 m accuracy (UTM; Universal Transverse Mercator projection Zone 18N, WGS84 horizontal datum). The coordinates of each GCP marker in the point cloud coordinate system were determined by manually identifying orange marker points in the point cloud and measuring their XYZ coordinates using ScanView software (Menci Software; http://www.menci.com). Six GCPs were selected for use in georeferencing, the center-most and the five most widely distributed across the study site; remaining GCPs were reserved for georeferencing accuracy assessment.

A 7-parameter Helmert transformation was used to georeference SFM point clouds to GCPs by means of an optimal transformation model implemented in Python (v2.7.2; Scipy v0.10.1; Optimize module) obtained by minimizing the sum of squared residuals in X, Y, and Z between the SFM and GCP coordinate systems, based on a single factor of scale, three factors of translation along each axis, and three angles of rotation along each axis (Figure 1c; Wolf and Dewitt 2000). Georeferencing accuracy was assessed using National Standard for Spatial Data Accuracy (NSSDA) procedures (RMSE = Root Mean Square Error, RMSEr = Radial (XY) RMSE, RMSEz = Vertical (Z) RMSE, 95 % Radial Accuracy, and 95 % Vertical Accuracy; Flood 2004), by comparing the transformed coordinates of the GCP markers withheld from the transformation model with their coordinates measured by precision GPS in the field. This technique for georeferencing is referred to as the 'GCP method'.

GCP markers at SERC were obscured by forest canopy under leaf-on conditions, so georeferencing was only achievable using GPS track data downloaded from the Hexakopter. This method was also applied to the Knoll and Herbert Run datasets to evaluate its accuracy against the GCP method. Owing to hardware limitations of the Hexakopter GPS, positions could only be acquired every 5 seconds, a much lower frequency that was out of synch with photograph acquisitions (2 frames s^{-1}). To overcome this mismatch and the lower precision of the Hexakopter GPS, the entire aerial GPS track (UTM coordinates) and the entire set of camera positions along the flight path (SFM coordinate system) were fitted to independent spline curves, from which a series of 100 XYZ pseudo-pairs of GPS and SFM camera locations were obtained using an interpolation algorithm (Python v2.7.2; Scipy v0.10.1; Interpolate module) and then used as input for the georeferencing of point clouds using the same Helmert transformation algorithm used in the GCP method. This technique for georeferencing is referred to as the 'spline method'. SERC georeferencing accuracy with the spline method was then assessed during leaf-off conditions based on 12 GCP markers placed along a road bisecting the study site that were observable in the SFM point cloud, using the same methods as for UMBC sites (Appendix 1.3). However, the poor geometric distribution of these GCP markers across the SERC study site precluded their direct use for georeferencing.

2.4.4 Noise filtering of SFM point clouds

Georeferenced SFM point clouds for each study site included a small but significant number of points located far outside the possible spatial limits of the potential real-world features, most likely as the result of errors in feature matching (Triggs et al. 1999). As in

LIDAR postprocessing, these 'noise' points were removed from point clouds after georeferencing using statistical outlier filtering (Sithole and Vosselman 2004). First, georeferenced point clouds were clipped to a 350 m x 350 m extent: the 250 m x 250 m study site plus a 50 m buffer on all sides to avoid edge effects. A local filter was applied by overlaying a 10 m grid across the clipped point cloud, computing standardized Z-scores (Rousseeuw and Leroy 1987) within each grid cell, and removing all points with |Z-score| > 3; between 1 % and 2 % of input points were removed at this stage (Appendix 1.2). While filtering did remove some verifiable canopy points, filters were implemented instead of manual editing to facilitate automation. At this point, 'Ecosynth' point clouds were ready for vegetation structure measurements.

2.4.5 Terrain filtering and DTM creation

After georeferencing and noise-filtering of computer vision point clouds, a 1 m grid was imposed across the entire clipped point cloud of the study site and the median elevation point within each 1 m grid cell was retained; all other points were discarded. Understory digital terrain models (DTMs) were then generated from these median-filtered leaf-on and leaf-off point clouds using morphological filter software designed for discrete return LIDAR point clouds (Figure 1d; Zhang et al. 2003; Zhang and Cui 2007). This software distinguished terrain points based on elevation differences within varying window sizes around each point within a specified grid mesh. This algorithm enabled convenient batching of multiple filtering runs with different algorithm parameters, a form of optimization that is a common and recommended practice with other filtering algorithm packages (Zhang et al. 2003; Sithole and Vosselman 2004; Evans and Hudak 2007, Tinkham et al. 2012) and has previously been used across a range of different forest

types, including high biomass redwood forests of the Pacific northwest (Gonzalez et al. 2010), Florida mangroves (Simard et al. 2006; Zhang 2008) and in prior studies at similar sites (Dandois and Ellis 2010). Ordinary Kriging was then used to interpolate 1 m raster DTMs from terrain points using ArcGIS 10.0 (ESRI, Redlands, CA; Popescu et al. 2003).

Ecosynth DTM error was evaluated across 250 m x 250 m sites as a whole relative to slope and land cover classes (Clark et al. 2004) following NSSDA procedures (Flood 2004). Land cover across the Knoll and Herbert Run sites was manually interpreted and digitized in ArcGIS 10.0 from a 2008 leaf-off aerial orthophotograph (0.6 m horizontal accuracy, 0.3 m pixel resolution, collected 2008-03-01 – 2008-04-01) into seven categories: forest (woody vegetation > 2 m height), turfgrass, brush (woody vegetation < 2 m height), buildings, pavement, water, and other (i.e., rock rip-rap, unpaved trail). Land cover feature height (e.g., greater or less than 2 m) and aboveground feature outline (e.g., for buildings and forest canopy) was determined from the Ecosynth canopy height model for each study site. The SERC study site was classified as all forest.

LIDAR understory DTMs were generated at UMBC sites using a bare earth point cloud product provided by the LIDAR contractor and interpolated to a 1 m grid using Ordinary Kriging. Despite being collected 5 years prior to the current study, the 2005 LIDAR bare earth product still provides an accurate depiction of the relatively unchanged terrain at the UMBC study sites. A LIDAR understory DTM was generated for the SERC study site using the morphological terrain filter on the set of 'last return' points and interpolating to a 1 m grid using Ordinary Kriging.

2.4.6 CHM generation and canopy height metrics

Sets of aboveground point heights were produced from Ecosynth and LIDAR point clouds by subtracting DTM cell values from the elevation of each point above each DTM cell; points below the DTM were discarded (Popescu et al. 2003). To investigate the accuracy of Ecosynth methods, aboveground point heights for Ecosynth leaf-on point clouds were computed against three different DTMs; those from leaf-on Ecosynth, leaf-off Ecosynth, and LIDAR bare earth; LIDAR CHMs were only processed against LIDAR DTMs. All aboveground points ≥ 2 m in height were accepted as valid canopy points and used to prepare CHM point clouds. CHM point height summary statistics were calculated within 25 m x 25 m subplots across each study site, including median (Hmed), mean (Hmean), minimum (Hmin), maximum (Hmax), and quantiles (25th, 75th, 90th, 95th and 99th = Q-25, Q-75, Q-90, Q-95 and Q-99 respectively). At all sites, Ecosynth and LIDAR CHM metrics were compared with field measured heights of the five tallest trees within each subplot using simple linear regressions (Dandois and Ellis 2010), although for Knoll and Herbert Run, LIDAR comparisons at these sites must be considered illustrative only: the long time delay since LIDAR data acquisition biases these from any direct quantitative comparisons. At SERC, where Ecosynth and LIDAR were collected only a few days apart in 2010, Ecosynth canopy height statistics were also compared directly with LIDAR height statistics within 25 m x 25 m grid cells overlaid across the SERC study site and compared using simple linear regression. For each site, one outlier was identified (Grubbs 1969) and removed from analysis where Ecosynth overestimated field height by > 10 m due to: tree removal (Knoll), tall canopy spreading

into a plot with few small trees (Herbert Run), and a plot that had only one large tree and several smaller, suppressed understory trees (SERC).

2.4.7 Prediction of aboveground biomass and carbon from 3D point clouds

Ecosynth and LIDAR CHMs were used to predict forest canopy aboveground biomass density (AGB Mg ha^{-1}) at all study sites using linear regression to relate canopy height metrics to field based estimates of biomass within forested 25 m x 25 m subplots. Biomass density was estimated by first computing per tree biomass using standardized allometric equations for the 'hard maple/oak/hickory/beech' group (Jenkins et al. 2003; AGB = EXP(-2.0127 + 2.4342×LN(DBH))), summing total AGB per subplot and then standardizing to units of Mg ha^{-1} (Hudak et al. 2012). Linear regression was then used to predict subplot AGB from CHM height metrics, with prediction error computed as the RMSE error between observed and predicted AGB values (Lefsky et al. 1999; Drake et al. 2002). Aboveground forest carbon density was estimated by multiplying AGB by a factor of 0.5 (Hurtt et al. 2004). As with estimates of canopy height, AGB predictions obtained from LIDAR at Knoll and Herbert Run would be expected to show large errors due to the large time difference between LIDAR (2005) and field measurements (2011). Nevertheless, AGB predictions were made at all sites using both Ecosynth and LIDAR to demonstrate the general utility of Ecosynth for similar applications as LIDAR.

2.4.8 Repeated seasonal 3D RGB vertical profiles

Computer vision RGB point clouds were used to assess forest spectral dynamics in 3D by producing multiple point cloud datasets of the SERC site in leaf-off (Winter), early spring (Spring 1), spring green-up (Spring 2), mature green (Summer), early senescing leaf-on (Fall 1), and senescing (Fall 2) conditions between October 2010 – June 2012

(Appendix 1.2). A single 50 m x 50 m sample area was selected for its diverse fall colors, clipped from each point cloud and stratified into 1 m vertical height bins for analysis. Canopy height profiles (CHPs) were then generated for all points within the 50 m x 50 m sample area across the six point clouds, with each 1 m height bin colorized using the mean RGB channel value of all points within the bin. The relative RGB channel brightness (e.g., R/(R+G+B)) was computed based on the mean RGB point color within each 1 m bin (Richardson et al. 2009). A CHP of the sample area was also generated from the leaf-on G-LiHT point cloud for comparison, combining all returns. For each of the six seasonal point clouds, the relative green channel brightness (i.e., G/(R+G+B), Strength of green: Sgreen) was extracted for all points within the height bin corresponding to mean field measured canopy height within the 50 m x 50 m sample area (Richardson et al. 2009; Mizunuma et al. 2013). Ecoysnth Sgreen values were plotted based on day of year (DOY) against the MODIS NDVI time series for 2011 (MOD13Q1; 16-day composite; 2011-01-01 – 2011-12-19; one 250 m pixel centered on 38° 53' 23.2"N 76° 33' 35.8"W; ORNL DAAC 2012). Regression analysis was used to directly compare MODIS NDVI and Sgreen values based on the closest MODIS NDVI DOY value for each Ecosynth DOY Sgreen value, or the mean of two NDVI values when an Ecosynth observation fell between two MODIS observations.

3.0 Results

3.1 Image acquisition and point cloud generation

Image acquisition flight times ranged from 11 to 16 minutes, acquiring between 1600 and 2500 images per site, depending mostly on prevailing winds. As wind speeds approached 4.4 m s^{-1} (9.8 mph), flight times increased substantially, image acquisition trajectories

ranged further from plan, and photo counts increased. Wind speeds > 4.4 m s^{-1} generally resulted in incomplete image overlap and the failure of point cloud generation and were thus avoided. Point cloud generation using commercial SFM software required between 27 – 124 hours of continuous computation to complete image processing across 6.25 ha sites, depending in part on the number of photographs (Appendix 1.2).

3.2 Characteristics of Ecosynth point clouds

Ecosynth point clouds are illustrated in Figure 2 and described in Appendix 1.2. Point cloud density varied substantially with land cover and between leaf-on and leaf-off acquisitions (Table 1, Figure 3), with forested leaf-on point clouds generally having the highest densities (Table 1). Densities of leaf-off point clouds were similar across all three study sites (20 – 23 points m^{-2}), and leaf-on densities were similar across UMBC sites (27 – 37 points m^{-2}), but the leaf-on SERC cloud was twice as dense (67 points m^{-2}) as the leaf-on UMBC clouds. Point cloud density varied substantially with land cover type at the Knoll and Herbert Run, and was generally highest in types with the greatest structural and textural complexity such as forest, low brush and rock riprap (29 – 54 points m^{-2}; Table 1) and lowest in types that were structurally simple and had low variation in contrast like roads, sidewalks, and turfgrass (7 – 22 points m^{-2}). However, building roof tops had similar point densities to vegetated areas at Herbert Run (35 points m^{-2}), where shingled roofs were present, compared with simple asphalt roofs at Knoll.

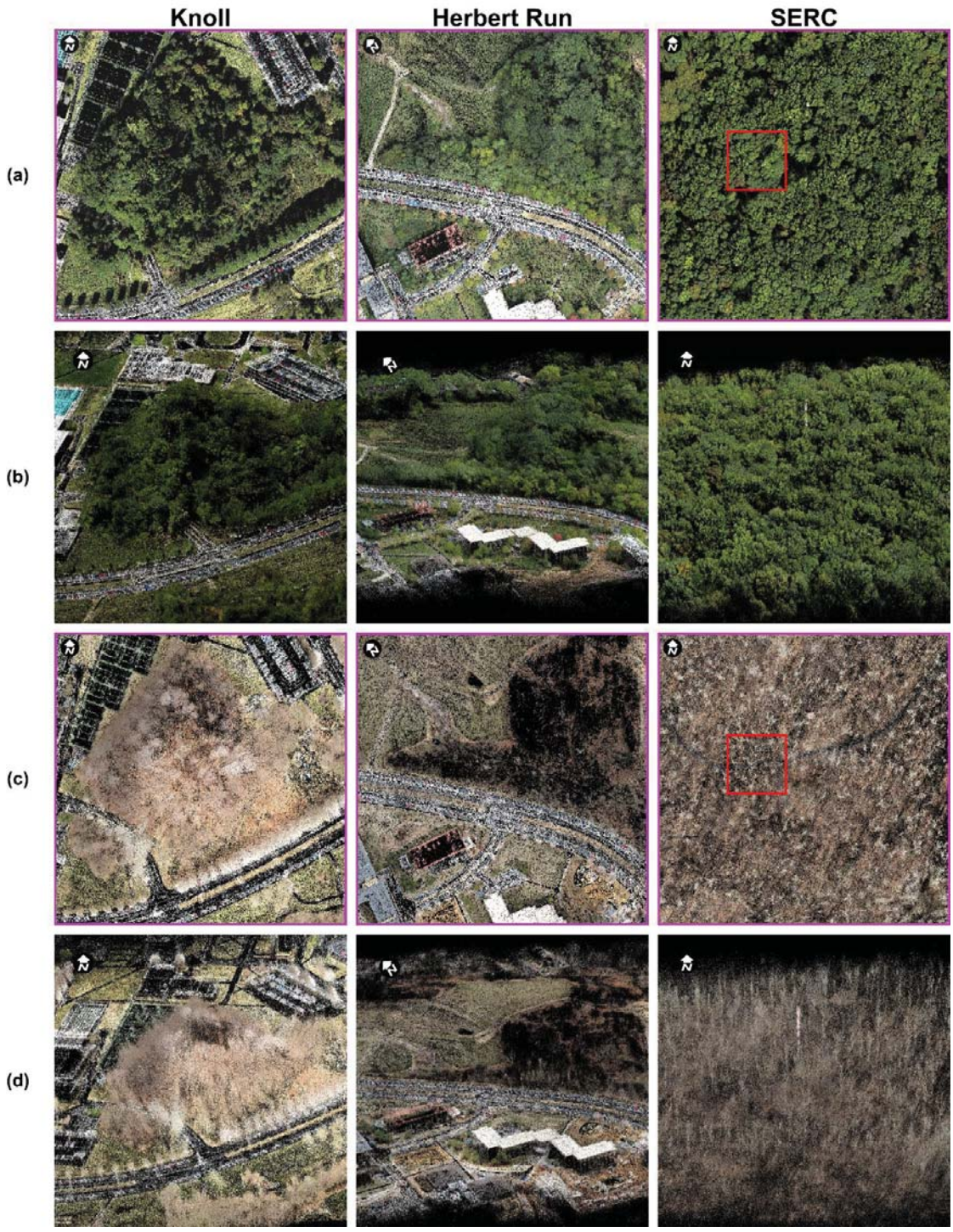


Figure 2: Ecosynth RGB point clouds across 250 m x 250 m sites (Knoll, Herbert Run and SERC; purple outline) under leaf-on (a, b) and leaf-off (c, d) conditions and viewed from overhead (a, c) and obliquely with same heading (b, d). Point clouds have black background; brightness and contrast enhanced using autocorrect settings in Microsoft Visio software.

Table 1: Ecosynth and LIDAR point cloud density for different land cover types. Blanks indicate land cover type not present within study site.

	Point cloud density by land cover class: mean(SD) [min,max] (points m^{-2})							
	Forest	Turfgrass	Brush	Buildings	Pavement	Water	Other	All
Knoll								
Leaf-on 1	54(72) [0,1030]	22(16) [0,160]	-	21(21) [0,168]	12(14) [0,147]	30(19) [1,92]	-	37(57) [0,1030]
Leaf-off	35(23) [0,207]	19(12) [0,153]	-	10(12) [0,156]	7(10) [0,183]	12(9) [0,68]	-	24(21) [0,207]
LIDAR[a]	1.7(0.9) [0,7]	1.8(1.0) [0,7]	-	1.7(1.1) [0,6]	1.7(1.0) [0,6]	0.3(0.7) [0,2]	-	1.7(1.0) [0,7]
Herbert Run								
Leaf-on	34(27) [0,249]	20(16) [0,249]	48(24) [1,144]	34(30) [0,251]	19(16) [0,170]	-	42(22) [3,140]	27(23) [0,251]
Leaf-off	12(15) [0,174]	26(14) [0,173]	39(19) [4,128]	36(28) [0,198]	17(13) [0,199]	-	29(18) [0,111]	20(17) [0,199]
LIDAR[a]	1.8(1.4) [0,12]	2.1(1.8) [0,18]	2.6(2.0) [0,8]	3.3(2.7) [0,20]	2.4(2.0) [0,12]	-	1.7(1.4) [0,8]	2.1(1.8) [0,20]
SERC								
Leaf-on	67(60) [0,829]	-	-	-	-	-	-	67(60) [0,829]
Leaf-off	23(19) [0,1194]	-	-	-	-	-	-	23(19) [0,1194]
LIDAR[b]	45(16) [0,173]	-	-	-	-	-	-	45(16) [0,173]

[a] Combined density of first return and bare earth points.

[b] Combined density of first and last return.

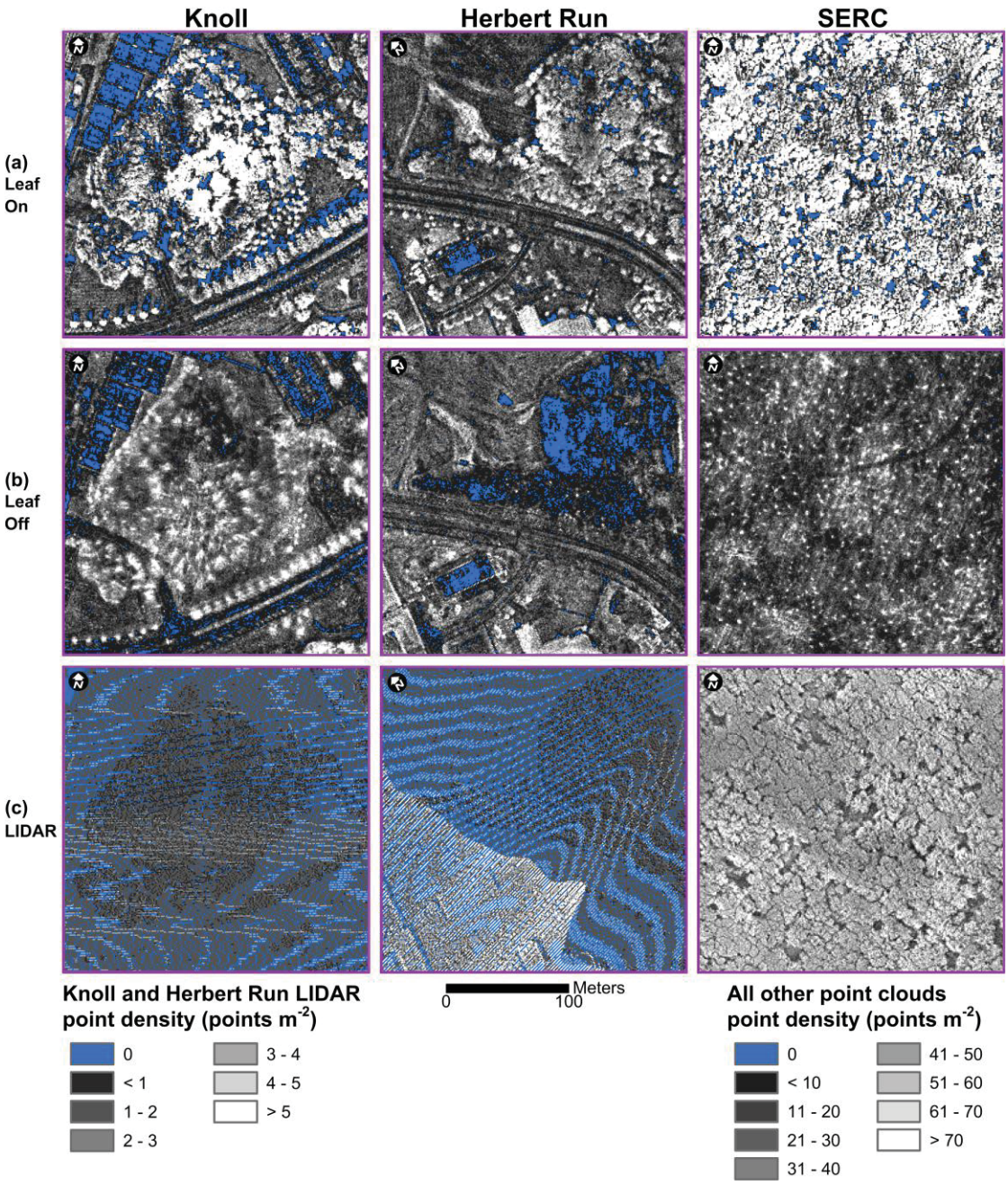


Figure 3: Point cloud density maps from Ecosynth under (a) leaf-on and (b) leaf-off conditions compared with (c) LIDAR across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). LIDAR densities combine first and last returns. Commercial LIDAR at Knoll and Herbert Run sites have lower density map legend as indicated.

Point cloud georeferencing accuracies are reported in Table 2. For the Knoll and Herbert Run sites, horizontal georeferencing accuracies of 1.2 m – 2.1 m RMSEr and vertical accuracies of 0.4 m – 0.6 m RMSEz were achieved using the GCP method. Horizontal and vertical accuracies of 4.1 m and 1.2 m, RMSEr and RMSEz respectively, were achieved for the SERC leaf-off point cloud using the spline method. However, the spline method produced lower horizontal and vertical accuracies (higher RMSE) than the GCP method at the Knoll and Herbert Run sites (RMSEr = 3.5 m – 5.4 m; RMSEz = 1.7 m – 4.7 m, Appendix A1.4). Horizontal and vertical RMSE for LIDAR are generally much lower (0.15 m, 0.24 m, contractor reported).

Table 2: Point cloud georeferencing error and accuracy across study sites for the GCP method.

	Horizontal				Vertical	
	$RMSE_x$	$RMSE_y$	$RMSE_r$	95 % Radial Accuracy	$RMSE_z$	95 % Vertical Accuracy
Knoll[a]						
Leaf-on 1	0.36	1.59	1.63	2.82	0.59	1.15
Leaf-off	0.99	0.69	1.21	2.09	0.62	1.22
Herbert Run[a]						
Leaf-on	0.71	1.79	1.93	3.33	0.61	1.20
Leaf-off	1.75	1.07	2.05	3.55	0.39	0.76
SERC						
Leaf-on[b]	-	-	-	-	-	-
Leaf-off[c]	2.46	3.31	4.13	7.14	1.16	2.28

[a] Accuracy is reported for the "GCP method", table of accuracies of the "spline method" for these sites is provided in Appendix A1.2.
[b] Accuracy could not be assessed because canopy completely covered GCPs.
[c] Georeferenced using aerial GPS "spline method".

3.3 Digital terrain models

Understory DTMs generated from computer vision are compared with LIDAR bare earth DTMs in Table 3 and Figure 4. Ecosynth DTM errors were higher under forest cover than in open areas at the Knoll and Herbert Run sites (Figure 4). Ecosynth leaf-off DTMs more accurately captured understory terrain than Ecosynth leaf-on DTMs (leaf-off RMSEz = 0.89 m – 3.04 m; leaf-on RMSEz = 2.49 m – 5.69 m; Table 3). At the Knoll, DTM difference maps between Ecosynth leaf-off and LIDAR (Figure 4c) revealed large error sinks (< -5 m) in the north-west, north-east, and southern portions of the study site. Leaf-on DTMs generally overestimated understory terrain elevation at all three study sites (Figure 4c) resulting in spikes of error (> 5 m) compared to LIDAR DTMs. At all sites, DTM differences between Ecosynth and LIDAR were larger in forest compared with non-forest areas (Figure 4c and d; Table 4).

Table 3: Understory digital terrain model (DTM) errors (meters) across entire study sites compared to LIDAR Bare Earth DTM.

	Mean (SD)	Range	$RMSE_z$	95 % Vertical Accuracy
Knoll				
Leaf-on 1	1.21 (2.17)	-11.18 – 13.97	2.49	4.88
Leaf-off	1.10 (2.83)	-39.96 – 10.30	3.04	5.96
Herbert Run				
Leaf-on	2.84 (3.77)	-4.81 – 20.69	4.72	9.25
Leaf-off	0.61 (0.64)	-12.07 – 8.57	0.89	1.73
SERC				
Leaf-on	4.90 (2.90)	-5.42 – 13.25	5.69	11.15
Leaf-off	0.84 (1.28)	-9.84 – 6.47	1.53	3.00

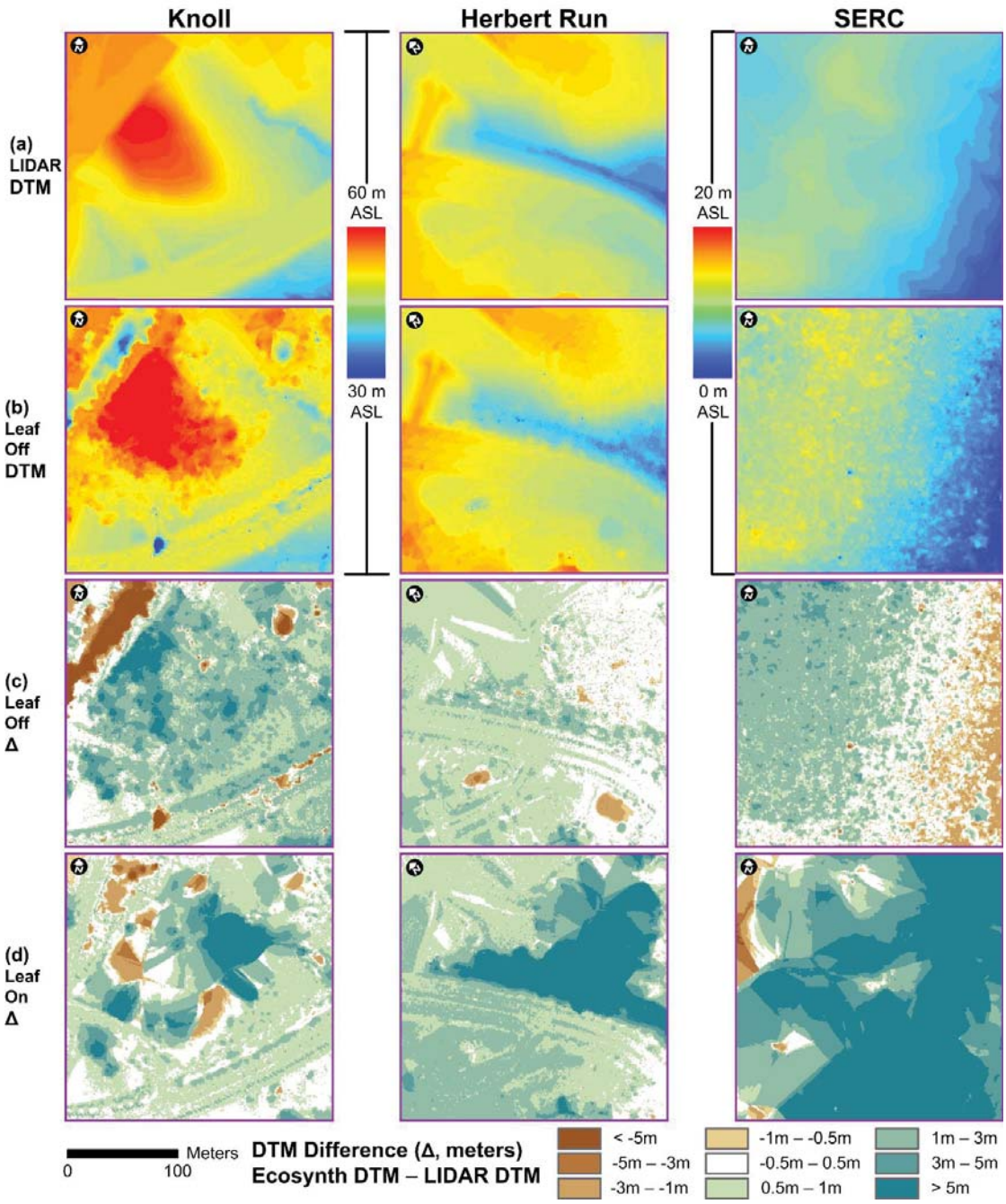


Figure 4: Digital terrain maps (DTM) from (a) LIDAR and (b) leaf-off Ecosynth and differences (Δ) between LIDAR and (c) leaf-off and (d) leaf-on Ecosynth DTMs across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). Negative differences highlight areas where Ecosynth DTM is lower than LIDAR. DTM legends for (a) and (b) differ from (c), as indicated.

Table 4: Understory DTM error compared to LIDAR bare earth DTM across different land cover types for terrain slopes ≤ 10°. Reported as Mean error ± SD (RMSE) in meters.

	Forest	Turfgrass	Brush	Buildings	Pavement	Water	Other	All
Knoll								
Leaf-on 1	1.80 ± 2.85 (3.37)	0.56 ± 0.22 (0.60)	-	1.11 ± 1.19 (1.63)	0.28 ± 1.02 (1.06)	-0.69 ± 0.30 (0.75)	-	1.09 ± 2.15 (2.41)
Leaf-off	2.27 ± 1.51 (2.72)	0.68 ± 0.69 (0.97)	-	0.68 ± 3.03 (3.11)	1.80 ± 2.85 (4.49)	-1.48 ± 0.20 (1.49)	-	0.79 ± 3.00 (3.10)
Herbert Run								
Leaf-on	4.16 ± 3.73 (5.58)	0.87 ± 0.40 (0.95)	1.62 ± 0.25 (1.64)	1.72 ± 1.02 (2.00)	0.93 ± 0.38 (1.00)	-	0.76 ± 0.17 (0.78)	2.06 ± 2.67 (3.38)
Leaf-off	0.46 ± 0.57 (0.73)	0.57 ± 0.49 (0.75)	0.71 ± 0.17 (0.73)	0.83 ± 1.20 (1.46)	0.61 ± 0.31 (0.68)	-	0.72 ± 0.08 (0.72)	0.56 ± 0.57 (0.80)
SERC								
Leaf-on	4.90 ± 2.90 (5.69)	-	-	-	-	-	-	4.90 ± 2.90 (5.69)
Leaf-off	0.84 ± 1.28 (1.53)	-	-	-	-	-	-	0.84 ± 1.28 (1.53)

3.4 Canopy height, biomass and carbon estimates

Use of Ecosynth and LIDAR CHMs to predict field measured tree heights across forest subplots at all sites is described in Table 5 and plotted for Ecosynth only in Figure 6. At the Knoll and Herbert Run sites, results demonstrate that Ecosynth CHMs adequately predicted field-measured heights of the five tallest trees per subplot (i.e., average maximum height, AvgTop5) when either Ecosynth leaf-off ($R^2 = 0.82 – 0.83$) or LIDAR DTMs ($R^2 = 0.83 – 0.84$) were used. When Ecosynth leaf-on DTMs were used, the quality of canopy height predictions was much lower ($R^2 = 0.62 – 0.67$). For the SERC site, Ecosynth predictions of field measured canopy height were very low for all DTMs ($R^2 = 0.07– 0.30$) and lower than would be expected when LIDAR was used to estimate field heights ($R^2 = 0.50$). For Ecosynth, field height prediction errors with the leaf-off

DTM (RMSE = 3.9 – 9.3 m) were generally higher than when the LIDAR DTM was used (RMSE = 3.2 – 6.8 m) but lower than when the leaf-on DTM was used (RMSE = 7.1 – 10.9 m). LIDAR CHMs at Knoll and Herbert Run showed a strong relationship to field measurements (R^2 = 0.71 & 0.77), but had larger errors (RMSE = 5.7 & 5.4 m) as expected given the 5 years elapsed between LIDAR and field measurements. At SERC, estimates of error between Ecosynth and LIDAR predictions of field canopy height were comparable (RMSE = 3.3 & 3.6 m). Direct comparison of Ecosynth and LIDAR CHMs at SERC, where data was collected only days apart, also revealed strong agreement between the two sensor systems (R = 0.87, RMSE = 2.3 m; Appendix A1.5), suggesting that the two sensors were characterizing the canopy with a similar degree of precision.

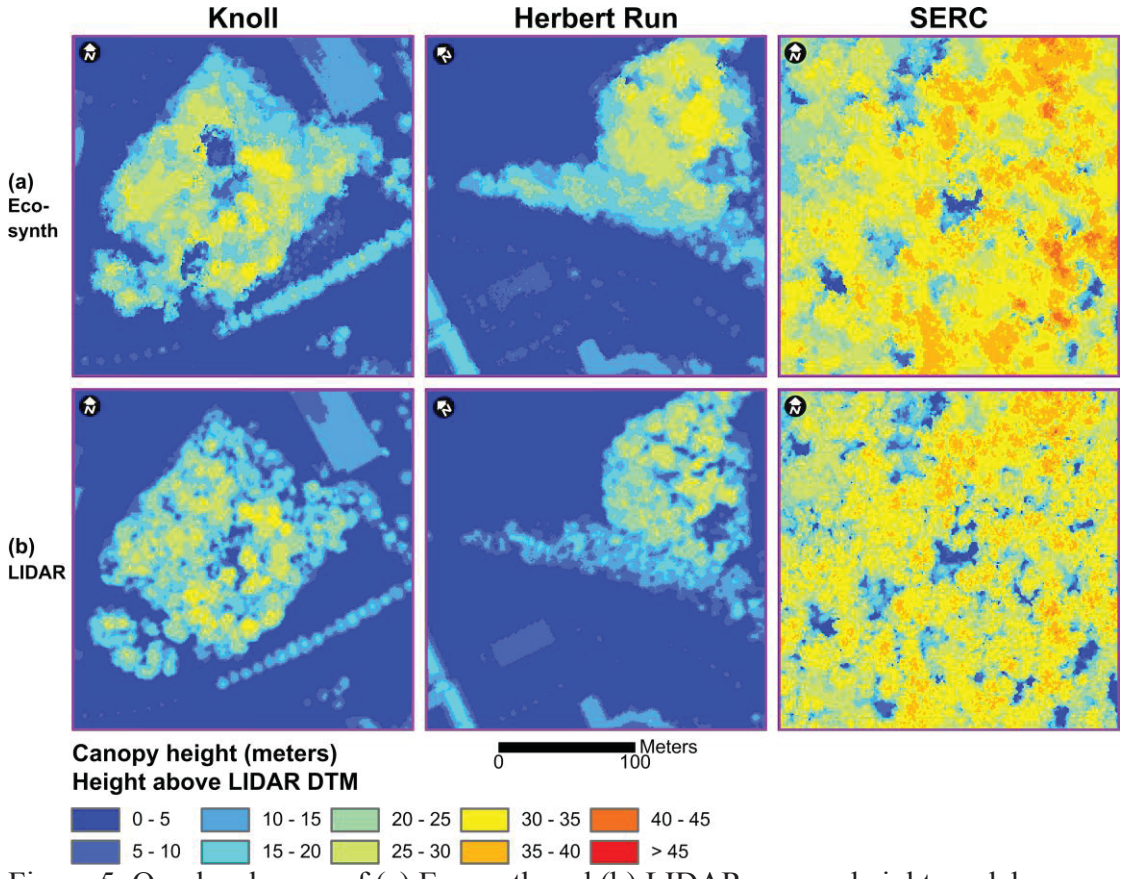


Figure 5: Overhead maps of (a) Ecosynth and (b) LIDAR canopy height models (CHM) across the Knoll, Herbert Run and SERC sites (same orientation as Figure 2a). Ecosynth CHMs produced using LIDAR DTMs.

Table 5: Best linear regression predictors (canopy height metric with the highest R^2) of field measured mean heights of the 5 tallest trees per subplot (AvgTop5) across forested areas of the Knoll, Herbert Run, and SERC sites based on Ecosynth methods with different DTMs and LIDAR alone. RMSE is deviation in meters between field measured AvgTop5 and the specified subplot canopy height metric.

	Linear Model	R^2	RMSE (m)
Ecosynth			
Knoll Leaf-on 1			
with Ecosynth leaf-on DTM	$AvgTop5 = 0.77 \times Hmed - 0.1$	0.63	6.9
with Ecosynth leaf-off DTM	$AvgTop5 = 0.67 \times Hmed + 2.0$	0.82	6.9
with LIDAR DTM	$AvgTop5 = 0.73 \times Hmed + 3.1$	0.83	4.4
Knoll Leaf -on 2			
with Ecosynth leaf-on DTM	$AvgTop5 = 0.82 \times Q\text{-}25 - 4.1$	0.67	9.4
with Ecosynth leaf-off DTM	$AvgTop5 = 0.72 \times Q\text{-}25 - 1.7$	0.83	9.3
with LIDAR DTM	$AvgTop5 = 0.81 \times Q\text{-}25 - 1.6$	0.84	6.8
Herbert Run			
with Ecosynth leaf-on DTM	$AvgTop5 = 0.78 \times Q\text{-}25 - 5.3$	0.62	10.9
with Ecosynth leaf-off DTM	$AvgTop5 = 0.88 \times Hmean - 0.3$	0.83	3.9
with LIDAR DTM	$AvgTop5 = 0.93 \times Hmed - 0.3$	0.84	3.2
SERC			
with Ecosynth leaf-on DTM	$AvgTop5 = 0.39 \times Q\text{-}90 + 15.9$	0.30	7.1
with Ecosynth leaf-off DTM	$AvgTop5 = 0.23 \times Q\text{-}90 + 26.3$	0.07	4.6
with LIDAR DTM	$AvgTop5 = 0.40 \times Q\text{-}90 + 20.7$	0.25	3.3
LIDAR			
Knoll	$AvgTop5 = 0.66 \times Hmed + 3.9$	0.71	5.7
Herbert Run	$AvgTop5 = 0.90 \times Hmed - 2.30$	0.77	5.4
SERC	$AvgTop5 = 0.50 \times Q\text{-}75 + 17.3$	0.50	3.6

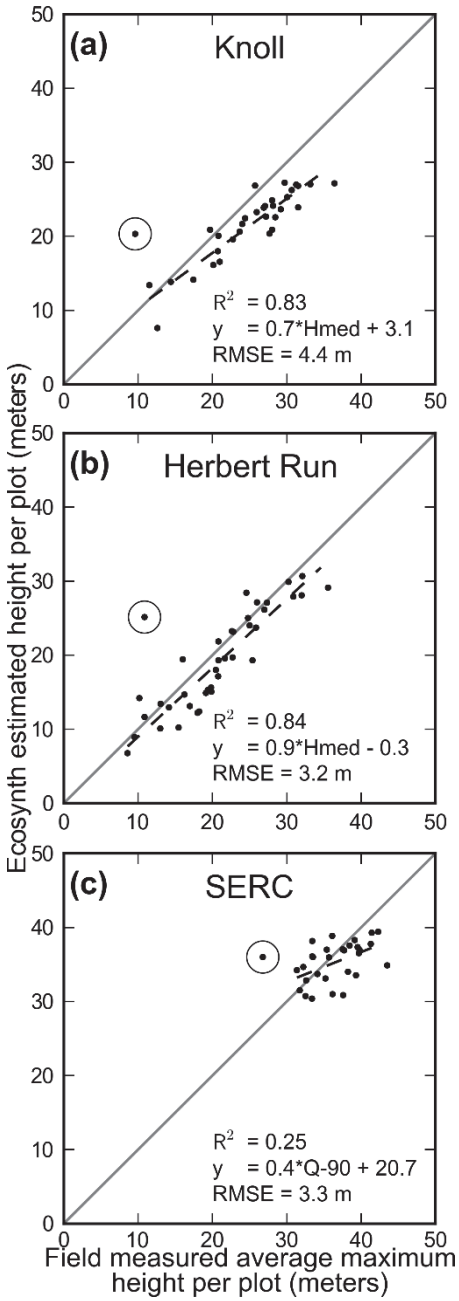


Figure 6: Ecosynth estimated best linear regression predictors (Table 5) of field measured average maximum height per subplot (AvgTop5) across forested areas of the Knoll (a), Herbert Run (b), and SERC (c) sites. Ecosynth canopy heights estimated using LIDAR DTMs. Linear regression lines (dashed), R^2, linear models, and RMSE (m) are presented for each comparison, with solid gray reference lines along the one-to-one ratio. Circled data points are outliers based on Grubb's test (> 3 SD from the mean) and are not included in regression analysis.

Aboveground biomass (AGB) predictions from Ecosynth and LIDAR CHMs at all sites are shown in Table 6. For Knoll and Herbert Run, Ecosynth predictions of field estimated AGB showed relatively strong relationships, but also relatively high error (R^2 = 0.71 & 0.73; RMSE = 94 & 87 Mg ha^{-1}, Table 6), with errors representing approximately 31 – 36 % of field estimated per subplot mean AGB densities from allometric equations. LIDAR predictions of AGB at Knoll and Herbert Run showed similar relationships to those from Ecosynth, but with 3 – 9 % more error relative to field estimated mean AGB (R^2 = 0.63 & 0.72; RMSE = 101 & 107 Mg ha^{-1}), a result that is expected given the time lag between LIDAR data acquisition and field measurements. At SERC, where Ecosynth and LIDAR data were collected at approximately the same time, the close resemblance of AGB predictions (Table 6) provides strong evidence that these systems generally yield similar estimates of AGB and aboveground carbon, which is approximated by multiplying AGB by a factor of 0.5 (Hurtt et al. 2004).

Table 6: Predicted aboveground biomass (AGB) from Ecosynth and LIDAR canopy height metrics across forested 25 m subplots at Knoll, Herbert Run, and SERC sites. Simple linear models predicting AGB are presented for the canopy height metric producing the highest R^2. Estimated AGB values are from allometric models.

				Ecosynth			LIDAR [b]		
	n (plots)	Estimated Mean AGB (Mg ha^{-1})	Estimated AGB CV	Prediction Error RMSE (Mg ha^{-1})	Linear Model	R^2	Prediction Error RMSE (Mg ha^{-1})	Linear Model	R^2
Knoll Leaf-on 1	32	299	53 %	94[a]	28.3 × *Hmed* – 313.2	0.71	101[a]	32.3× *Hmean* – 332.4	0.63
Herbert Run	40	240	65 %	87[a]	23.9 × *Q-90* – 329.9	0.73	107[a]	23.4× *Hmax* – 130.3	0.72
SERC	100	341	40 %	112[a]	18.3 × *Q-90* – 300.4	0.27	106[a]	24.9× *Q-90* – 562.1	0.34

[a] RMSE values are cross-validated (Drake et al. 2002)
[b] LIDAR at Knoll and Herbert Run is from 2005 and therefore biased, LIDAR at SERC is from 2011

3.5 Vertical canopy profiles

Vertical canopy height profiles (CHPs) of Ecosynth CHMs are shown in Figure 7 for a selected 50 m x 50 m sample area at SERC, illustrating the relative frequency of points within 1 m height bins and their mean RGB color. CHPs from the Spring 2, Summer, Fall 1 and Fall 2 time periods (Figure 7g, 7i, and 7k) showed a similar vertical density profile as the single leaf-on LIDAR acquisition at this site. However, at the same time periods, Ecosynth observed almost no points in the understory and at ground level when compared with both LIDAR and Ecosynth Winter and Spring 1 scans (Figure 7a and 7c). Mean RGB channel brightness was fairly constant across the vertical profile in each time period, except in the Spring 1 and Fall 1 time periods, with slightly lower green and higher blue levels at the top of the canopy under early senescing conditions (Fall 1, Figure 7j), and slightly higher green at the same height under early spring conditions (Spring 1, Figure 7d). Time series comparison of MODIS NDVI (MOD13Q1) for 2011 and Ecosynth Sgreen for the 38 – 39 m height bin for the corresponding day of year (DOY) is shown in Figure 8. For the observed time periods, the time series pattern of Ecosynth Sgreen closely matched that of MODIS NDVI and corresponding NDVI and Sgreen values were highly correlated ($R^2 = 0.88$).

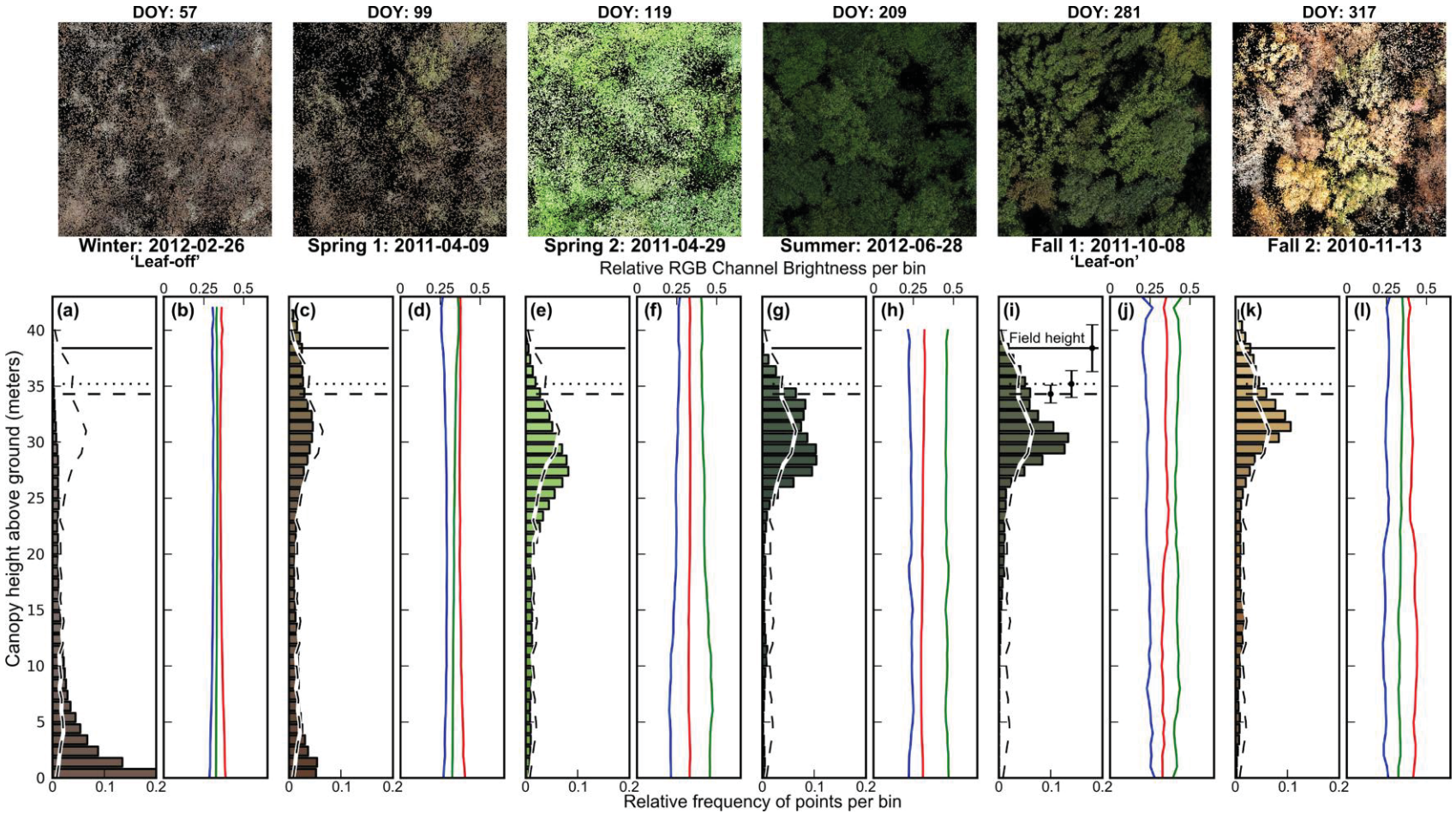


Figure 7: Time series of Ecosynth point clouds (overhead view), canopy height profiles (CHPs; (a), (c), (e), (g), (i), (k)) and mean RGB channel brightness ((b), (d), (f), (h), (j), (l)) of a 50 m x 50 m sample area of SERC site (red box in Figure 2a, 2c). Black dashed line is LIDAR CHP (all returns; 2011-10-05). Horizontal lines are estimates of mean canopy height, with standard error in (i): solid line - field height 38.4 m (2.1 m), dotted line - Ecosynth Q-90 35.2 m (1.2 m), dashed line - LIDAR Q-75 (34.3 m (0.8 m); Table 5). Overhead views (on black background) and RGB channel brightness are shown without color enhancement. CHPs show mean RGB color per 1 m bin, uniformly scaled to enhance brightness and contrast for viewing using min-max linear correction.

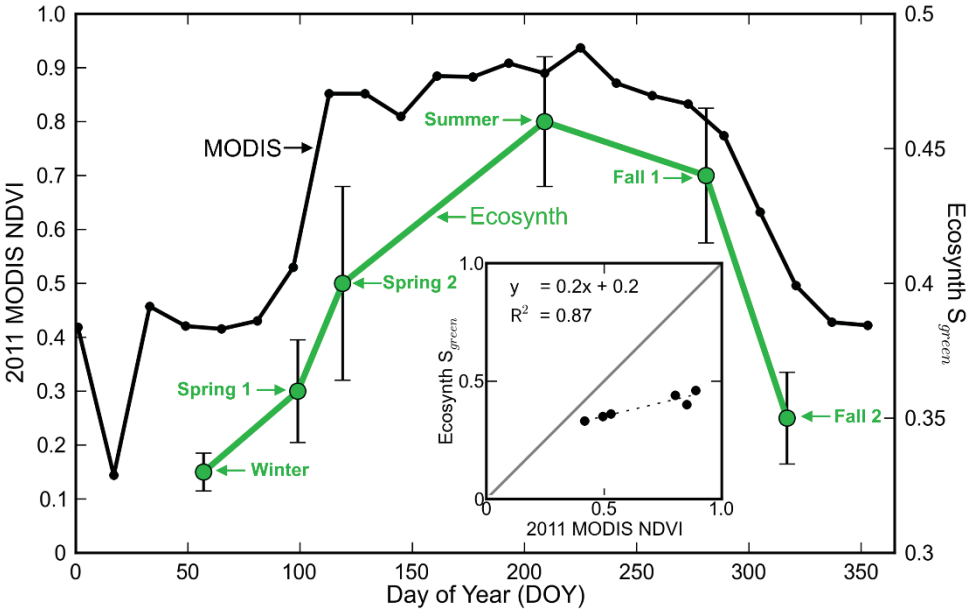


Figure 8: Time series of Ecosynth relative green brightness (green line) and MODIS NDVI (black line) across the SERC site. Relative green brightness is for each Ecosynth point cloud (Figure 7) at the 38 m – 39 m height bin; error bars are standard deviation. Linear regression line (dashed), R^2, and linear models are presented in subset, with solid gray reference line along the one-to-one ratio.

4.0 Discussion

4.1 Ecosynth canopy height models (CHMs)

Ecosynth CHMs produced strong predictions of field-measured tree heights at the Knoll and Herbert Run (R^2 = 0.82 – 0.84, Table 5, Figure 6), well within the typical range of LIDAR predictions (Andersen et al. 2006, Wang and Glenn 2008), except when Ecosynth leaf-on DTMs were used for CHM generation (R^2 = 0.62 – 0.67, Table 5). At the SERC site, Ecosynth CHM predictions of field-measured tree height were very weak (R^2 = 0.07 - 0.30) regardless of DTM used as were LIDAR CHM predictions (R^2 = 0.50, Table 5). Weaker prediction power of Ecosynth and LIDAR CHMs at SERC may be explained by the relatively lower variation in average maximum canopy height at that site (coefficient of variation, CV, 9 %) compared to Knoll or Herbert Run (CV = 23 % & 32 %; Figure 6,

Appendix A1.3). At all sites, prediction errors were lowest when the LIDAR DTM was used (RMSE = 3.2 - 4.4 m) except for the Knoll leaf-on 2 CHM where prediction errors were large regardless of DTM (RMSE = 6.8 - 9.4 m). For that dataset, relatively higher prediction errors may be explained by noting that the best linear regression model, selected based on the highest R^2 value, is based off of the 25th percentile (Q-25) canopy height metric, resulting in underestimation of field measured average maximum canopy height and high RMSE error. The next 'best' regression models based on R^2 alone show much lower RMSE error (Hmean: $R^2 = 0.83$, RMSE = 4.1 m; Hmed: $R^2 = 0.81$, RMSE = 3.5 m), results which are more similar to predictions with other CHMs.

While the practice of predicting canopy height from the 'best' of many (often highly correlated) metrics derived from 3D point clouds is common-place, either from LIDAR or computer vision structure from motion (Næsset and Gobakken 2005; Næsset and Gobakken 2008; Dandois and Ellis 2010; Lisein et al. 2013), these results suggest limitations of this approach. Estimates of canopy height from distribution-based metrics may be strongly biased by the structure of the canopy and the properties of the remote sensing instrument, which will influence the shape of the vertical height distribution from which quantile metrics are derived (Magnussen and Boudewyn 1998; Hopkinson et al. 2006). This effect was seen for both Ecosynth and LIDAR at the SERC site where both instruments under-predicted the field measured average maximum canopy height by roughly 4 meters and the distribution of points was highly skewed to the top of the canopy only, with few points observed in the understory (Figure 7i). In general, Ecosynth distribution-based canopy height metrics under-predicted field estimated average maximum canopy height at all sites (Figure 6) as did LIDAR at the SERC site

(Figure 7, Appendix A1.5). Recently, a new canopy height metric was introduced that is intended to be more comparable across different instruments and data collections and which is not as susceptible to variation in point cloud penetration into the canopy (Asner and Mascaro 2014). This top of canopy height measure (TCH) corresponds to the average maximum height of the outer surface of the canopy observed within a fixed area. Since the outer canopy surface is primarily what is being observed by remote sensing instruments (Parker and Russ 2004), especially for 3D point clouds that do not penetrate deeply into the canopy (Figure 7), such a measure may be ideally suited for estimating canopy height from Ecosynth and future research should consider this as an alternative to distribution-based metrics.

Ecosynth canopy surface heights closely resembled those mapped using LIDAR at SERC, where datasets were collected four days apart (SERC, Figure 5). At this site, Ecosynth and LIDAR CHMs were strongly correlated ($R = 0.87$), differing by < 2.3 m RMSE (Appendix A1.5), well within the error range of plot level canopy height predictions from small footprint LIDAR, which are generally between 1 and 3 m RMSE, (Clark et al. 2004, Andersen et al. 2006; Hyyppä et al. 2008). Errors in LIDAR predictions of field measured tree heights at the Knoll and Herbert Run (RMSE = 5.7 & 5.4 m) are readily explained by the 5 year time lag between LIDAR acquisition and field measurements. While comparisons of Ecosynth and LIDAR CHMs at the Knoll and Herbert Run sites are biased by the 5 year time lag, a number of ecologically relevant changes in canopy structure are observable in Figures 5a and 5b. At the Knoll, Ecosynth revealed a large tree gap just north of center in the forest where a large beech had fallen down after the LIDAR acquisition, the planting of about 30 small ornamental trees ($\approx$ 10

m height) to the south-east of the main forest area, and general increases in tree height over 5 years. At Herbert Run, rapid growth of black locust, honey locust and green ash trees is visible in a recovering riparian forest area (below road).

Errors in Ecosynth canopy height predictions are less well understood than those for LIDAR, but include some similar sources, including errors in measuring tree height and location in the field, DTM error, and errors introduced by limitations of the sensor system (Andersen et al. 2006; Hyyppä et al. 2008; Falkowski et al. 2008). With LIDAR, lower flight altitudes generally produce more accurate observations of the forest canopy, at the cost of reduced spatial coverage (Hyyppä et al. 2008). Ecosynth images were acquired at much lower altitudes than typical for LIDAR (40 m above canopy vs. > 350 m, Appendix A1.2), but it is not known if higher altitudes, which would increase the spatial coverage of observations, would also reduce the accuracy of height measurements, or even increase it. The point densities of Ecosynth were comparable to the dense point clouds produced by the NASA G-LiHT LIDAR at SERC (Table 1), but it is not known whether Ecosynth point densities are correlated with the accuracy of canopy height estimates. LIDAR studies indicate that estimates of height, biomass, and other structural attributes are relatively robust to changes in point cloud density down to 0.5 points m^{-2} (Næsset and Gobakken 2005; Næsset 2009; Treitz et al. 2012).

In Ecosynth methods, overstory occlusion limits observations and point densities lower in the canopy. Leaf-on DTMs were therefore of much lower quality than those from leaf-off conditions, lowering the accuracy of CHMs that can be produced in regions without a leaf-off period. However, repeated estimates of forest canopy heights confirm that Ecosynth methods are robust under a range of forest, terrain, weather, flight

configuration, and computational conditions. For example, at the Knoll site, two leaf-on image collections acquired with different canopy conditions and color due to autumn senescence, different lighting conditions (clear and uniformly cloudy; Appendix A1.2) and were processed using different versions of computer vision software (Photoscan v0.8.4 and v0.7.0), yet these produced canopy height estimates comparable to field measurements when LIDAR or Ecosynth leaf-off DTMs were used (Knoll leaf-on 1 R^2 = 0.83 & 0.82; leaf-on 2 R^2 = 0.84 & 0.83; Table 5). Nevertheless, the accuracy and density of Ecosynth point clouds do appear to be sensitive to a number of poorly characterized factors including camera resolution, flight altitude, and the SFM algorithm used for 3D processing, justifying further research into the influence and optimization of these factors to produce more accurate estimates of vegetation structure.

4.2 Predictions of aboveground biomass and carbon

At Knoll and Herbert Run, Ecosynth predictions of canopy height metrics and field estimated AGB (R^2 = 0.71 & 0.73; Table 6) were comparable to those common for LIDAR and field measurements, which have R^2 ranging from 0.38 to 0.80 (Lefksy et al. 2002; Popescu et al. 2003; Zhao et al. 2009). However, at SERC both Ecosynth and LIDAR predictions of field estimated AGB were lower than would be expected (R^2 = 0.27 & 0.34). When assessed using cross-validated RMSE as a measure of AGB prediction 'accuracy' (Drake et al. 2002; Goetz and Dubayah 2011), Ecosynth AGB was also less accurate than field estimates based on allometry and LIDAR (Table 6). In addition to errors in canopy height metrics, AGB error sources include field measurements along with errors in allometric modeling of AGB from field measurements which include uncertainties of 30 % - 40 % (Jenkins et al. 2003). Another limit to the

strength of AGB predictions (R^2) is the relatively low variation in canopy heights and biomass estimates across this study; higher R^2 are generally attained for models of forests across a wider range of successional states (Lefsky et al. 1999). For example at SERC, the relatively low variation in subplot AGB (coefficient of variation, CV, 40 %) relative to other sites (53 % & 65 %) may explain the low R2 and large error in LIDAR estimates of AGB (R^2 = 0.34, RMSE = 106 Mg ha^{-1}); at the Knoll and Herbert Run, 2005 LIDAR AGB predictions cannot be fairly compared with those based on 2011 field measurements. Despite their generally lower quality, Ecosynth canopy height metrics can be successfully combined with field measurements of biomass, carbon or other structural traits (e.g., canopy bulk density, rugosity) to generate useful high-resolution maps for forest carbon inventory, fire and habitat modeling and other research applications (Vierling et al. 2008; Skowronski et al. 2011; Hudak et al. 2012).

4.3 Observing canopy spectral dynamics in 3D

Vertical profiles of forest canopy density and color generated from Ecosynth point clouds reveal the tremendous potential of computer vision remote sensing for natively coupled observations of vegetation structure and spectral properties at high spatial and temporal resolutions (Figure 7, Figure 8). Canopy structure observed by LIDAR and Ecosynth 4 days apart at SERC under early fall (Fall 1) conditions yielded similar point densities across canopy height profiles with a notable peak around 30m (Figure 7i), and similar densities were observed under senescing (Figure 7k) and summer conditions (Figure 7g). As expected, under leaf off-conditions Ecosynth point density showed a completely different pattern, with the highest density observed near the ground (Figure 7a). Comparison of green and senescent canopy color profiles reveals a shift in relative

brightness from 'more green' to 'more red' (Figure 7j & 7l), caused by increasing red leaf coloration in deciduous forests during autumn senescence that has also been observed in annual time series from stationary multispectral web cameras in deciduous forests in New Hampshire, USA (Richardson et al. 2009). Under leaf-off conditions, colors were fairly constant across the vertical canopy profile with the relatively grey-brown coloration of tree trunks and the forest floor (Figure 7b). During spring green-up, the canopy profile showed a strong increase in canopy density in the upper layers, likely due to the emergence of new small leaves and buds (Figure 7c), and this is confirmed by slight increases in relative green brightness at the top of the canopy (Figure 7d).

Summer and fall canopy profiles (Summer, Fall 1, Fall 2) had greater point densities than LIDAR at the overstory peak, but few to no points below this peak (Figures 7g, 7i, and 7k), likely because dense canopy cover under these conditions occluded and shadowed understory features. When cameras cannot observe forest features, they cannot be detected or mapped using computer vision algorithms, a significant limitation to observing forest features deeper in the canopy, especially under leaf-on conditions. Conversely, the spring green-up profile (Spring 2; Figure 7e) showed a greater density of points in the understory compared to summer and fall profiles, but also a lower peak. This may be due to the fact that Spring 2 photos observed deeper into the canopy from being over-exposed (e.g., brighter but with reduced contrast) due to changes in illumination during the scan flyover, resulting in more well illuminated shadows (Cox and Booth 2009).

The strength of greenness (Sgreen) in the overstory at SERC followed a similar pattern throughout the growing season as the MODIS NDVI time series over the same

area and Sgreen was highly correlated with corresponding MODIS NDVI DOY values ($R^2 = 0.88$; Figure 8), suggesting that Ecosynth may be a useful proxy for NDVI. NDVI measured with satellite remote sensing provides strong predictions of ecosystem phenology and dynamics (Pettorelli et al. 2005; Morisette et al. 2008; Zhang and Goldberg 2011) and high spatial resolution, near surface observations obtained with regular digital cameras can provide more detailed information to help link ground and satellite based observations (Richardson et al. 2009; Graham et al. 2010; Mizunuma et al. 2013). High spatial and temporal resolution 3D-RGB Ecosynth data provides an additional level of detail for improving understanding of ecosystem dynamics by incorporating information about canopy 3D structural change along with color spectral change. The increase in Sgreen at the top of the canopy in the Spring 1 point cloud may be associated with the dominant *Liriodendron tulipifera* (tulip-poplar) tree crowns within the forest, which are expected to green-up first in the season (Appendix A1.6; Parker & Tibbs 2004).

Unlike current LIDAR image fusion techniques, Ecosynth methods natively produce multispectral 3D point clouds without the need for high precision GPS and IMU equipment, enabling data acquisition using inexpensive, lightweight, low altitude UAS, thereby facilitating routine observations of forest spectral dynamics at high spatial resolutions in 3D, a new and unprecedented observational opportunity for forest ecology and environmental management. Ecosynth methods may also complement LIDAR image fusion collections by enabling high frequency observations of forest canopy dynamics in between infrequent LIDAR acquisitions, with LIDAR DTMs enhancing Ecosynth CHMs in regions where forests do not have a leaf-off season.

4.4 General characteristics of Ecosynth 3D point clouds

Ecosynth point clouds are generated from photographs, so 3D points cannot be observed in locations that are occluded from view in multiple photos, including understory areas occluded by the overstory, or in areas masked in shadow, leading to incomplete 3D coverage in Ecosynth datasets. In contrast, LIDAR provides relatively complete observations of the entire canopy profile, from top to ground, even in leaf-on conditions, owing to the ability of laser pulses to penetrate through the canopy (Dubayah and Drake 2000). Nevertheless, 3D point clouds produced using UAS-enhanced Ecosynth methods compare favorably with those from aerial LIDAR, though positional accuracies of Ecosynth point clouds were significantly lower (horizontal error = 1.2 m – 4.1 m; vertical error = 0.4 m – 1.2 m; Table 2) than those derived from LIDAR (0.15 m, 0.24, contractor reported). While lower positional accuracies are certainly an important consideration, accuracies in the one to four meter range are generally considered adequate for most forestry applications (Clark et al. 2004), and are consistently achieved by Ecosynth methods under all conditions.

Ecosynth point cloud densities (23 – 67 points m^{-2}) were substantially higher than those common for commercial LIDAR products (1.5 points m^{-2}; UMBC sites) and were comparable with the research-grade LIDAR datasets acquired during testing of the NASA G-LiHT fusion system for remote sensing of vegetation (> 78 points m^{-2} ; SERC site; Cook et al. 2012). Point densities were higher for leaf-on point clouds (27 – 67 points m^{-2}) compared with leaf-off point clouds (20 – 24 points m^{-2}; Table 1; Figure 3) and also in forested versus non-forested areas within sites. Structurally homogenous and simple land surfaces (e.g., rooftops, open grass, pavement) produced far fewer points

when compared with structurally complex surfaces (e.g., forest canopy, riprap and low brush; Table 1). This higher point density in tree covered areas is most likely the result of high textural variation in image intensity and/or brightness, which are the basis for feature identification in computer vision (de Matías et al. 2009), though the greater height complexity of forested areas is probably also a factor. Regardless of mechanism, the density and accuracy of 3D point clouds produced by Ecosynth methods across forested landscapes are clearly sufficient for general forestry applications.

4.5 Practical challenges in producing Ecosynth point cloud measurements

4.5.1 Image acquisition using UAS

UAS image acquisition systems generally performed well, but required significant investments in training and hardware. Operator training and system building required six weeks and was accomplished using only online resources. To maintain image acquisition capabilities on demand in the face of occasional aircraft damage and other issues, it was necessary to purchase and maintain at least two, and better three, fully functional UAS imaging systems, an investment of approximately USD$4000 for the first unit and USD$3000 for additional units (some equipment was redundant). Automated UAS image acquisitions by trained operators were mostly routine (the 9 scans of this study were acquired in 11 acquisition missions), enabling repeated acquisitions on demand across 6.25 ha sites using the same flight plan with flight times < 15 min. The only major limitations to acquisition flights were precipitation and wind speeds > 4.4 m s^{-1}, which caused significant deflection of the aircraft and incomplete image acquisitions. Technological developments in hobbyist-grade UAS are very rapid and accelerating, improving capabilities, driving down prices and increasing availability, as exemplified by

the rapid growth and spread of the Ardupilot (http://copter.ardupilot.com/) open-source UAS autopilot platform and the DIYdrones online community (http://diydrones.com).

4.5.2 Computation

The commercial computer vision software used in this study required > 27 hours to produce a single 3D point cloud across a 250 m x 250 m site when run on a high-end computer graphics workstation with full utilization of all CPU and RAM resources (Appendix A1.2). The widely available open source computer vision software, Bundler (Snavely et al. 2008), would likely take more than one month to produce similar results. These computational limits are being overcome by more rapid and efficient open-source computer vision algorithms now under development, utilizing parallel processing (e.g. Agarwal et al. 2009) and Graphical Processing Units in calculations (Wang and Olano 2011), and by redesigning the computer vision processing pipeline to incorporate the sequential structure of image acquisitions (Wang and Olano 2011).

4.5.3 Georeferecing

Two different georeferencing techniques were used to produce Ecosynth point clouds: one based on ground markers visible from the air (GCP method) and one based on the aircraft GPS path (spline method). As would be expected from the relatively low precision of the inexpensive lightweight GPS in the UAS, the spline method consistently produced point clouds with lower horizontal and vertical RMSE (4.3 m, 2.5 m; Appendix A1.4) than the GCP method (1.7 m, 0.6 m; Table 2). This is likely the major source of the large georeferencing errors observed at the SERC site (4.1 m, 1.2 m), where ground markers were obscured by the tree canopy. Use of a more precise (and expensive) aircraft GPS, improving the georeferencing accuracy of the spline method algorithms,

and the development of other techniques for georeferencing without GCP markers would be useful foci for future research, as field marker placement is both time consuming and fruitless in closed canopy forests without regular canopy gaps. One possible solution may be the development of algorithms that combine aerial GPS locations directly with the camera intrinsic parameters solved for by computer vision algorithms (Xiang and Tian 2011). It might also be useful to improve the georeferencing accuracy of the GCP method by more accurately surveying GCP marker locations (mapping-grade GPS was used in this study) - a relevant consideration at field research sites where permanent GCP markers can be established to facilitate repeated data collections.

4.5.4 Terrain models

DTM accuracy fundamentally constrains the accuracy of canopy height and related measures of vegetation structure (Andersen et al. 2006; Wang and Glenn 2008). Ecosynth DTMs showed large deviations from LIDAR DTMs (Figure 4), which are expected to have elevation precisions of approximately ± 2 m RMSE depending on many factors not specifically evaluated in the current study (Kobler et al. 2007; Gatziolis et al. 2010; Tinkham et al. 2011). As would be expected, the precision of Ecosynth DTMs was highest under leaf-off conditions (RMSEz = 0.73 m to 2.72 m) compared with leaf-on acquisitions (3.37 m to 5.69 m; Table 3; Figure 4), and were also more precise in the non-forested areas of the Knoll and Herbert Run (0.60 to 4.49 m) compared with forested areas. Nevertheless, Ecosynth leaf-off DTMs accurate to within 1 – 3 m RMSE error when compared to LIDAR DTMs can be produced that are adequate for estimating and mapping forest canopy heights.

In Tropical Moist Forests and other regions without leaf-off periods, the limited leaf-on DTM accuracy of Ecosynth methods remains a significant challenge to producing accurate measurements of vegetation structure. Even under leaf-off conditions, there are multiple challenges to producing accurate Ecosynth DTMs. Leaf-off point densities in forested areas of Herbert Run (Figure 3b) were much lower than at other sites; non-forest densities were comparable (Table 1). Some inherent characteristic of Herbert Run forests might explain this, but differences in lighting conditions offer a stronger explanation. Imagery used for Herbert Run DTM generation were collected under overcast conditions, in contrast with the Knoll (partly cloudy) and SERC (clear), where brighter understory illumination may have enhanced computer vision point recognition and produced deeper and denser understory point clouds. Further study of the effects of lighting and other scene conditions may help identify more optimal strategies for Ecosynth DTM production.

A second challenge in Ecosynth DTM production is terrain filtering. Even after noise filtering to remove extreme outliers, Ecosynth DTMs tended to retain large sinks caused by low outliers in the terrain point cloud that were not removed by terrain filtering algorithms, which were designed for LIDAR point clouds (Sithole and Vosselman 2004). These sinks are clearly visible in the north-east, north-west, and southern part of the Knoll leaf-off DTM (Figures 4b and 4c). DTM accuracy is generally influenced by terrain slope, vegetation cover and by the type of filtering algorithm employed (Sithole and Vosselman 2004; Tinkham et al. 2011; Tinkham et al. 2012), with the greatest accuracies usually achieved by manual filtering (Kobler et al. 2007; Gatziolis et al. 2010). Improved terrain filtering algorithms designed specifically for Ecosynth DTM production

would likely create stronger results than those designed for LIDAR point clouds- another useful area for future study.

4.5.6 Advancing computer vision remote sensing

By combining automated UAS image acquisition with state-of-the-art computer vision algorithms, consistent and repeatable high-spatial resolution 3D point clouds of vegetation were produced across study sites with practical levels of computer resources, largely addressing the major challenges raised in prior work (Dandois and Ellis 2010). Yet substantial room remains to improve understanding of the parameter space of computer vision remote sensing systems (Table 7). With LIDAR, observational error models and the effects on accuracy of different sensor parameters including altitude and scan resolution are well understood thanks to decades of research (Glennie 2007; Næsset 2009). With Ecosynth, basic questions remain about the effects on accuracy of basic elements of the remote sensing system (e.g., the platform, camera, processing algorithms, etc.,) and the conditions of observation (e.g., wind, illumination, forest type and phenology, etc.,), and these parameters likely interact in determining the quality and accuracy of Ecosynth results. It is also not clear precisely how computer vision algorithms 'see' canopy structure to identify features in imagery (e.g., leaves, branches, gaps, etc.), and how ecologically relevant spectral information might be better acquired by these algorithms. Future investigations of these factors influencing Ecosynth data quality and accuracy across a range of different forest types should enable a more complete understanding of how Ecosynth methods can be optimized to measure forest structural and spectral traits and their dynamics.

Table 7: Key factors influencing the quality of data obtained by computer vision remote sensing.

Factor	Effects on data quality
Platform	Altitude, speed, and flight path overlap affect the detail and depth of canopy that can be observed. Camera angle and potentially camera array structure may affect point densities, detail and depth of observations into canopy.
Camera	Resolution, frame rate, overlap, exposure, color settings, spectral channels (RGB, NIR) may all affect feature identification and matching, resulting in different point cloud spectral properties and densities.
Algorithms	Algorithms for feature identification, feature matching, use of secondary densification algorithms, color assignment to features, and camera calibration may affect point cloud 3D model accuracy, density and spectral properties.
Georeferencing	UAS GPS and GCP quality affect spatial accuracy of point clouds and estimates of vegetation structure.
Post-processing, filtering	Different filtering algorithms (e.g., DTM filtering) affect accuracy in terrain and canopy height models
Wind	Route following errors can reduce image overlaps, moving leaves and branches limit feature matching and generate positional errors.
Illumination	Brighter light/full sun increase shadow, leading to decreased penetration in CHP. Diffuse lighting appears to increase penetration in CHP but also lowers contrast, reducing feature identification.
Forest: type, species, phenology	The same computer vision techniques may produce different results in different forest types (e.g. closed canopy needleleaf forests vs. open canopy deciduous woodlands), as demonstrated for the same canopies under different phenological conditions.

5.0 Conclusions

Ecosynth methods produce coupled spectral and structural observations at the high spatial and temporal resolutions required to observe vegetation phenology in 3D, portending new approaches to observing and understanding the dynamics of woodland ecosystems. Moreover, Ecosynth yields 3D forest measurements and mapping products comparable to LIDAR and field-based methods at low economic and logistical costs, facilitating multispectral 3D scanning of vegetation on demand at landscape scales (< 1 km^2) by end users of these data, heralding a new era of participatory remote sensing by field ecologists, community foresters, and even the interested public. Applications of Ecosynth

range from high spatial resolution 3D observations of vegetation phenology at the cutting edge of ecological research, to the monitoring of forest carbon stocks or habitat quality by local land managers and conservation groups (Goetz and Dubayah 2011). This is only the beginning of the transformation of remote sensing by computer vision technologies. By combining inexpensive imagery with computation for 3D canopy reconstruction, computer vision remote sensing systems can be made ever more light-weight, inexpensive and easy to use. As computing powers increase, Ecosynth and related methodologies might ultimately enable multispectral 3D remote sensing on demand by anyone with a cell-phone.

References

Agarwal, S., Snavely, N., Simon, I., Seitz, S.M., & Szeliski, R. (2009). Building Rome in a day. In, 2009 IEEE 12th International Conference on Computer Vision (ICCV) (pp. 72-79).

Andersen, H.-E., Reutebuch, S.E., & McGaughey, R.J. (2006). A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. Canadian Journal of Remote Sensing, 32, 355-366.

Andersen, H.E., McGaughey, R.J., & Reutebuch, S.E. (2005). Estimating forest canopy fuel parameters using LIDAR data. Remote Sensing of Environment, 94, 441-449.

Anderson, J., Plourde, L., Martin, M., Braswell, B., Smith, M., Dubayah, R., Hofton, M., & Blair, J. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. Remote Sensing of Environment, 112, 1856-1870.

Antonarakis, A.S., Saatchi, S.S., Chazdon, R.L., & Moorcroft, P.R. (2011). Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. Ecological Applications, 21, 1120-1137.

Asner, G.P. (2009). Tropical forest carbon assessment: integrating satellite and airborne mapping approaches. Environmental Research Letters, 4, 034009.

Asner, G.P., & Martin, R.E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Frontiers in Ecology and the Environment, 7, 269-276.

Asner, G.P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. Remote Sensing of Environment, 140, 614-624.

Clark, M.L., Clark, D.B., & Roberts, D.A. (2004). Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. Remote Sensing of Environment, 91, 68-89.

Cook, B., L. Corp, R. Nelson, D. Morton, K. J. Ranson, J. Masek, and E. Middleton. 2012. G-LiHT: Goddard's LiDAR, Hyperspectral and Thermal Airborne Imager. Silvilaser, Vancouver, Canada, 16-19 September 2012.

Cox, S.E., & Booth, D. T. (2009). Shadow attenuation with high dynamic range images. Environmental Monitoring and Assessment, 158, 231-241.

Dandois, J.P., & Ellis, E.C. (2010). Remote sensing of vegetation structure using computer vision. Remote Sensing, 2, 1157-1176.

de Matías, J., Sanjosé, J.J.d., López-Nicolás, G., Sagüés, C., & Guerrero, J.J. (2009). Photogrammetric Methodology for the Production of Geomorphologic Maps: Application to the Veleta Rock Glacier (Sierra Nevada, Granada, Spain). Remote Sensing, 1, 829-841.

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B., & de Souza Jr, C. (2007). Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. Environmental Science and Policy, 10, 385-394.

Dey, D., Mummert, L., & Sukthankar, R. (2012). Classification of plant structures from uncalibrated image sequences. In, Applications of Computer Vision (WACV), 2012 IEEE Workshop on (pp. 329-336).

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.D. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sensing of Environment, 79, 305-319.

Dubayah, R.O., & Drake, J.B. (2000). Lidar Remote Sensing for Forestry. Journal of Forestry, 98, 44-46.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., & Chazdon, R.L. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. Journal of Geophysical Research, 115, G00E09.

Erdody, T.L., & Moskal, L.M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sensing of Environment, 114, 725-737.

Evans, J.S., & Hudak, A.T. (2007). A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments. Geoscience and Remote Sensing, IEEE Transactions on, 45, 1029-1038.

Falkowski, M.J., Smith, A.M.S., Gessler, P.E., Hudak, A.T., Vierling, L.A., & Evans, J.S. (2008). The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. Canadian Journal of Remote Sensing, 34, S338-S350.

Flood, M., 2004. ASPRS guidelines. Vertical accuracy reporting for lidar data. http://www.asprs.org/a/society/committees/lidar/Downloads/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf [Accessed: 2012-03-14].

Frolking, S., Palace, M.W., Clark, D.B., Chambers, J.Q., Shugart, H.H., & Hurtt, G.C. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. Journal of Geophysical Research, 114.

Gatziolis, D., Fried, J.S., & Monleon, V.S. (2010). Challenges to Estimating Tree Height via LiDAR in Closed-Canopy Forests: A Parable from Western Oregon. Forest Science, 56, 139-155.

Geerling, G., Labrador-Garcia, M., Clevers, J., Ragas, A., & Smits, A. (2007). Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. International Journal of Remote Sensing, 28, 4263-4284.

Glennie, C. (2007). Rigorous 3D error analysis of kinematic scanning LIDAR systems. Journal of Applied Geodesy, 1, 147-157.

Goetz, S., & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management, 2, 231-244.

Gonzalez, P., Asner, G.P., Battles, J.J., Lefsky, M.A., Waring, K.M., & Palace, M. (2010). Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. Remote Sensing of Environment, 114, 1561-1575.

Graham, E.A., Riordan, E.C., Yuen, E.M., Estrin, D., & Rundel, P.W. (2010). Public Internet-connected cameras used as a cross-continental ground-based plant phenology monitoring system. Global Change Biology, 16, 3014-3023.

Grubbs, F.E. (1969). Procedures for Detecting Outlying Observations in Samples. Technometrics, 11, 1-21.

Hardiman, B.S., Bohrer, G., Gough, C.M., Vogel, C.S., & Curtis, P.S. (2011). The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. Ecology, 92, 1818-1827.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sensing, 4, 1573-1599.

Hirschmugl, M., Ofner, M., Raggam, J., & Schardt, M. (2007). Single tree detection in very high resolution remote sensing data. Remote Sensing of Environment, 110, 533-544.

Holl, K.D., Zahawi, R.A., Cole, R.J., Ostertag, R., & Cordell, S. (2011). Planting Seedlings in Tree Islands Versus Plantations as a Large-Scale Tropical Forest Restoration Strategy. Restoration Ecology, 19, 470-479.

Hopkinson, C., Chasmer, L., Lim, K., Treitz, P., & Creed, I. (2006). Towards a universal LIDAR canopy height indicator. Canadian Journal of Remote Sensing, 32, 139-152.

Houghton, R.A., Hall, F., & Goetz, S.J. (2009). Importance of biomass in the global carbon cycle. Journal of Geophysical Research, 114.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sensing of Environment, 82, 397-416.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S.n., & Falkowski, M.J. (2012). Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. Remote Sensing of Environment, 123, 25-40.

Hunt, J., E. Raymond, Hively, W.D., Fujikawa, S., Linden, D., Daughtry, C.S., & McCarty, G. (2010). Acquisition of NIR-Green-Blue Digital Photographs from Unmanned Aircraft for Crop Monitoring. Remote Sensing, 2, 290-305.

Hurtt, G., Dubayah, R., Drake, J., Moorcroft, P., Pacala, S., Blair, J., & Fearon, M. (2004). Beyond potential vegetation: combining lidar data and a height-structured model for carbon studies. Ecological Applications, 14, 873-883.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., & Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. International Journal of Remote Sensing, 29, 1339 – 1366.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., & Birdsey, R.A. (2003). National-Scale Biomass Estimators for United States Tree Species. Forest Science, 49, 12-35.

Jung, K., Kaiser, S., Böhm, S., Nieschulze, J., & Kalko, E.K.V. (2012). Moving in three dimensions: effects of structural complexity on occurrence and activity of insectivorous bats in managed forest stands. Journal of Applied Ecology, 49, 523-531.

Kampe, T.U., Johnson, B.R., Kuester, M., & Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. Journal of Applied Remote Sensing, 4, 043510.

Ke, Y., Quackenbush, L.J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. Remote Sensing of Environment, 114, 1141-1154.

Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K., & Dzeroski, S. (2007). Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. Remote Sensing of Environment, 108, 9-23.

Leberl, F., Irschara, A., Pock, T., Gruber, M., Scholz, S., & Wiechart, A. (2010). Point clouds: Lidar versus 3D vision. Photogrammetric Engineering and Remote Sensing, 76, 1123-1134.

Lefsky, M., Cohen, W., Harding, D., Parker, G., Acker, S., & Gower, S. (2002). Lidar remote sensing of above-ground biomass in three biomes. Global Ecology and Biogeography, 11, 393-399.

Lefsky, M.A., Harding, D., Cohen, W.B., Parker, G., & Shugart, H.H. (1999). Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sensing of Environment, 67, 83-98.

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., & Lejeune, P. (2013). A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. Forests, 4, 922-944.

Lin, Y., Hyyppä, J., & Jaakkola, A. (2011). Mini-UAV-Borne LIDAR for Fine-Scale Mapping. Geoscience and Remote Sensing Letters, IEEE, 8, 426-430.

Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60, 91-110.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 28, 1016-1031.

McMahon, S.M., Parker, G.G., & Miller, D.R. (2010). Evidence for a recent increase in forest growth. Proceedings of the National Academy of Sciences, 107, 3611-3615.

Mizunuma, T., Wilkinson, M., L. Eaton, E., Mencuccini, M., I. L. Morison, J., & Grace, J. (2013). The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern England. Functional Ecology, 27, 196-207.

Morisette, J., Richardson, A., Knapp, A., Fisher, J., Graham, E., Abatzoglou, J., Wilson, B., Breshears, D., Henebry, G., & Hanes, J. (2008). Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. Frontiers in Ecology and the Environment, 7, 253-260.

Mundt, J.T., Streutker, D.R., & Glenn, N.F. (2006). Mapping sagebrush distribution using fusion of hyperspectral and lidar classifications. Photogrammetric Engineering and Remote Sensing, 72, 47.

Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment, 113, 148-159.

Næsset, E., & Gobakken, T. (2005). Estimating forest growth using canopy metrics derived from airborne laser scanner data. Remote Sensing of Environment, 96, 453-465.

Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. Remote Sensing of Environment, 112, 3079-3090.

Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC). 2012. MODIS subsetted land products, Collection 5. Available on-line [http://daac.ornl.gov/MODIS/modis.html] from ORNL DAAC, Oak Ridge, Tennessee, U.S.A. Accessed March 04, 2012.

Ofner, M., Hirschmugl, M., Raggam, H., & Schardt, M. (2006). 3D Stereo Mapping by Means of UltracamD Data. In, Workshop on 3D Remote Sensing in Forestry. Vienna, Austria.

Parker, G., Harding, D., & Berger, M. (2004). A portable LIDAR system for rapid determination of forest canopy structure. Journal of Applied Ecology, 41, 755-767

Parker, G.G., & Russ, M.E. (2004). The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. Forest Ecology and Management, 189, 307-315.

Parker, G.G., & Tibbs, D.J. (2004). Structural phenology of the leaf community in the canopy of a Liriodendron tulipifera L. forest in Maryland, USA. Forest Science, 50, 387-397.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., & Stenseth, N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution, 20, 503-510.

Popescu, S., Wynne, R., & Nelson, R. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing, 29, 564-577.

Rango, A., Laliberte, A., Herrick, J.E., Winters, C., Havstad, K., Steele, C., & Browning, D. (2009). Unmanned aerial vehicle-based remote sensing for rangeland assessment, monitoring, and management. Journal of Applied Remote Sensing, 3, 15.

Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). UAV photogrammetry for mapping and 3D modeling: Current status and future perspectives. In, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Switzerland.

Richardson, A.D., Braswell, B., Hollinger, D.Y., Jenkins, J.C., & Ollinger, S.V. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications, 19, 1417-1428.

Rosnell, T., & Honkavaara, E. (2012). Point Cloud Generation from Aerial Image Data Acquired by a Quadrocopter Type Micro Unmanned Aerial Vehicle and a Digital Still Camera. Sensors, 12, 453-480.

Rousseeuw, P.J., & Leroy, A.M. (1987). Robust Regression and Outlier Detection. New York, NY, USA: John Wiley.

Schimel, D., Keller, M., Berukoff, S., Kao, B., Loescher, H., Powell, H., Kampe, T., Moore, D., & Gram, W. (2011). NEON Science Infrastructure. In, NEON 2011 Science Strategy (pp. 36-38).

Simard, M., Zhang, K., Rivera-Monroy, V.H., Ross, M.S., Ruiz, P.L., Castañeda-Moya, E., Twilley, R.R., & Rodriguez, E. (2006). Mapping height and biomass of mangrove forests in Everglades National Park with SRTM elevation data. Photogrammetric Engineering and Remote Sensing, 72, 299-311.

Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 59, 85-101.

Skowronski, N.S., Clark, K.L., Duveneck, M., & Hom, J. (2011). Three-dimensional canopy fuel loading predicted using upward and downward sensing LiDAR systems. Remote Sensing of Environment, 115, 703-714.

Snavely, N., Seitz, S., & Szeliski, R. (2008). Modeling the world from internet photo collections. International Journal of Computer Vision, 80, 189-210.

Snavely, N., Simon, I., Goesele, M., Szeliski, R., & Seitz, S.M. (2010). Scene Reconstruction and Visualization from Community Photo Collections. Proceedings of the IEEE, 98, 1370-1390.

St. Onge, B., Hu, Y., & Vega, C. (2008). Mapping the height and above-ground biomass of a mixed forest using lidar and stereo Ikonos images. International Journal of Remote Sensing, 29, 1277-1294.

Tao, W., Lei, Y., & Mooney, P. (2011). Dense point cloud extraction from UAV captured images in forest area. In, 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services. Fuzhou: June 29 2011 - July 1 2011.

Thomas, R.Q., Hurtt, G.C., Dubayah, R., & Schilz, M.H. (2008). Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain. Canadian Journal of Remote Sensing, 34, S351-S363.

Tinkham, W.T., Huang, H., Smith, A.M.S., Shrestha, R., Falkowski, M.J., Hudak, A.T., Link, T.E., Glenn, N.F., & Marks, D.G. (2011). A Comparison of Two Open Source LiDAR Surface Classification Algorithms. Remote Sensing, 3, 638-649.

Tinkham, W.T., Smith, A.M.S., Hoffman, C., Hudak, A.T., Falkowski, M.J., Swanson, M.E., & Gessler, P.E. (2012). Investigating the influence of LiDAR ground surface errors on the utility of derived forest inventories. Canadian Journal of Forest Research, 42, 413-422.

Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., & Etheridge, D. (2012). LiDAR Sampling Density for Forest Resource Inventories in Ontario, Canada. Remote Sensing, 4, 830-848.

Treuhaft, R.N., Law, B.E., & Asner, G.P. (2004). Forest Attributes from Radar Interferometric Structure and Its Fusion with Optical Remote Sensing. Bioscience, 54, 561-571.

Triggs, B., McLauchlan, P., Hartley, R., & Fitzgibbon, A. (1999). Bundle adjustment- A modern synthesis. Lecture Notes in Computer Science, 298-372.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. Trends in Ecology & Evolution, 18, 306-314.

Vauhkonen, J., Tokola, T., Maltamo, M., & Packalén, P. (2008). Effects of pulse density on predicting characteristics of individual trees of Scandinavian commercial species using alpha shape metrics based on airborne laser scanning data. Canadian Journal of Remote Sensing, 34, S441-S459.

Verhoeven, G. (2011). Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeological Prospection, 18.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., & Clawges, R.M. (2008). Lidar: shedding new light on habitat characterization and modeling. Frontiers in Ecology and the Environment, 6, 90-98.

Vitousek, P., Asner, G.P., Chadwick, O.A., & Hotchkiss, S. (2009). Landscape-level variation in forest structure and biogeochemistry across a substrate age gradient in Hawaii. Ecology, 90, 3074-3086.

Wallace, L., Lucieer, A., Watson, C., & Turner, D. (2012). Development of a UAV-LiDAR System with Application to Forest Inventory. Remote Sensing, 4, 1519-1543.

Wang, C., & Glenn, N.F. (2008). A linear regression method for tree canopy height estimation using airborne lidar data. Canadian Journal of Remote Sensing, 34, S217-S227.

Wang, Y., & Olano, M. (2011). A framework for GPS 3D model reconstruction using structure-from-motion. In, ACM SIGGRAPH '11. Vancouver, BC, Canada.

Wolf, P.R., & Dewitt, B.A. (2000). Elements of Photogrammetry with Applications in GIS 3rd Edition: McGraw-Hill Science/Engineering/Math.

Xiang, H., & Tian, L. (2011). Method for automatic georeferencing aerial remote sensing (RS) images from an unmanned aerial vehicle (UAV) platform. Biosystems Engineering, 108, 104-113.

Zhang, K. (2008). Identification of gaps in mangrove forests with airborne LIDAR. Remote Sensing of Environment, 112, 2309-2325.

Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing, 41, 872-882.

Zhang, K., & Cui, Z. (2007). ALDPAT v1.0: Airborne LiDAR Data Processing and Analysis Tools. In: Available from: http://lidar.ihrc.fiu.edu/lidartool.html.

Zhang, X. and M. D. Goldberg. (2011). Monitoring fall foliage coloration dynamics using time-series satellite data. Remote Sensing of Environment 115, 382-391.

Zhao, K., & Popescu, S. (2009). Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. Remote Sensing of Environment, 113, 1628-1645.

Zhao, K.G., Popescu, S., & Nelson, R. (2009). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment, 113, 182-196.

Chapter 3: What is a point? Analysis of computer vision image features of vegetation

Abstract

Computer vision structure from motion (SFM) algorithms are becoming increasingly popular for producing LIDAR-like three-dimensional (3D), multi-spectral point clouds of tree and forest canopy vegetation, yet much remains unknown about what the datasets produced by these algorithms are measuring, limiting potential applications and comparisons to other forms of remote sensing of vegetation. Here we examine the way in which vegetation is represented by SFM 3D-spectral fusion point clouds and evaluate the empirical quality of those observations. 3D-spectral point clouds of a single free-standing tree and three Temperate Deciduous forest patches are produced with the free and open source Bundler SFM algorithm and SIFT feature detector, based on photos collected from the ground and air by an automated hobbyist unmanned aerial system (UAS). 3D-spectral 'fusion' quality was evaluated by comparison between point and image color as well as the color of painted targets placed in free-standing trees under leaf-on and leaf-off conditions. Evaluation of what is 'seen' in SFM point clouds was by clustering of the SIFT image feature descriptors associated with points as well as manual interpretation of descriptors in a custom graphical user interface. Painted targets placed in the test tree were correctly classified (observed with the correct color) by SFM point clouds at a higher rate under leaf-off conditions than under leaf-on conditions (93 % vs. 29 %) due primarily to the fact that targets were frequently not observed by SFM under leaf-on conditions. Analysis of point cloud 'image features' by clustering of SIFT numeric descriptors revealed that points are located at places within the scene that when photographed are either brighter or darker than the immediate surroundings. SIFT

clusters showed high overlap in the frequency of manual classification tags and in general showed no strong connection to real canopy objects or the semantic content of images (e.g., leaves, branches, crowns, etc.). Combined, these results indicate that the quality of SFM 3D-fusion products and the relationship between data points and real-world objects is closely linked to the feature detector used in the structure from motion pipeline which regulates how the scene is sampled with points. In general, feature detectors have not been developed and tested with the goal of 3D scanning of vegetation in mind, suggesting potential new areas of research. Future interpretation of SFM 3D-spectral point cloud data for vegetation should more strongly consider the role of the feature detector in interpreting results and also for its potential for classification of point cloud points into distinct categories, for example pavement, grass, trees, or even leaves, fruits, and flowers.

1.0 Introduction

Three-dimensional (3D), multi-spectral observations of forest canopies by automated computer vision structure from motion (SFM) algorithms represent a transformative technological advance for ecological remote sensing. Relatively easy-to-use SFM algorithms and increased access to automated hobbyist unmanned aerial systems (UASs) has lowered the barrier to entry for remote sensing of vegetation structural and spectral attributes (Snavely et al. 2010; Anderson and Gaston 2013; Dandois and Ellis 2013). SFM algorithms produce a 'point cloud' dataset that is similar in appearance to a discrete return LIDAR (Light Detection and Ranging) point cloud but with RGB (red-green-blue) color associated with each point from the images. This new form of SFM UAS-enabled remote sensing enables on-demand, high-resolution observation of vegetation for estimating canopy structure and biomass (Dandois and Ellis 2010; Lisein et al. 2013), the

topography of stream channels and bare geologic substrates (Harwin and Lucieer 2012; Westoby et al. 2012; Javernick et al. 2014), the structure of single trees (Morgenroth and Gomez 2014), and even the seasonal dynamics of forest canopies at the scale of individual trees based on repeat observations (Dandois and Ellis 2013).

The inherent 'fusion' of structural 3D and optical color-spectral information in computer vision SFM remote sensing opens the door for advanced forms of analysis of forest canopies that had previously been out of reach. Fusion remote sensing enables improved tree crown classification and even biodiversity mapping, among many applications (Asner et al. 2007; Asner and Martin 2009; Baldeck and Asner 2013, 2014). To date, remote sensing fusion has only been possible by combining 3D and spectral measurements from different sensors, which is costly do to well, may not be accessible as-needed or for frequent monitoring, and presents its own challenges due to the differences in the way in which different sensors observe the canopy (Hudak et al. 2002; Popescu and Wynne 2004; Mundt et al. 2006; Geerling et al. 2007; Packalén et al. 2009; Kampe et al. 2010; Schimel et al. 2011). Computer vision SFM remote sensing offers the potential to overcome these challenges by producing an inherently fused 3D-spectral remote sensing dataset from the same sensor. In addition, because of the low cost and ease with which UAS can be deployed from image acquisition, individual researchers can not only access but produce on their own 3D-spectral fusion measurements of vegetation canopies when and where needed, a feat not previously possible with existing forms of remote sensing (Dandois and Ellis 2013).

Even so, much remains unknown about the empirical quality of the 3D-spectral fusion datasets produced by computer vision SFM remote sensing and how the observations are actually representing canopy objects (Chapter 2: Section 4.5.6, Table 7; Dandois and Ellis 2013). What are SFM point cloud points and how are they related to canopy objects? Does an individual point represent a leaf, branch, part of a crown, or something entirely unrelated to the canopy? What is the quality of the inherent 3D-spectral fusion of an SFM point cloud? Is the color of a point in the correct place? Such uncertainties limit interpretability of the measurements relative to more well understood remote sensing and field data and also limit the widespread application of SFM remote sensing to ecological applications. We aim to address these questions and in doing so improve understanding of how vegetation is observed or 'scanned' by a computer vision structure from motion remote sensing system.

1.1 SFM for remote sensing 3D-spectral fusion

SFM remote sensing produces a LIDAR-like 3D point cloud dataset where color is inherently assigned ('fused') to points as part of the overall processing pipeline. The fusion of optical color-spectral imagery with three-dimensional structural remote sensing data, primarily from Light Detection and Ranging (LIDAR), represents the state-of-the-art for characterizing ecosystem vegetation. 3D-spectral fusion products improve understanding of ecosystem vegetation beyond what can be achieved with either system alone, for example in describing the spatial heterogeneity of forest canopy biochemistry (Vitousek et al. 2009), fuel loading (Erdody and Moskal 2010), land cover types (Tooke et al. 2009; Zhou et al. 2009), and even in the discrimination of individual forest canopy tree species (Asner and Martin 2009; Ke et al. 2010; Baldeck and Asner 2013, 2014).

Despite this potential, the actual practice of fusing data from two separate sensor systems is challenging due to mismatches in spatial coverage and misalignment, different scales of observation (e.g., pixel size vs. LIDAR foot print size), and the inherent difficulties in attempting to collect co-synchronous data from two systems (Hudak et al. 2002; Mundt et al. 2006; Geerling et al. 2007; Anderson et al. 2008). Precision integrated LIDAR-spectral fusion systems address these challenges (Asner and Martin 2009) and are currently being deployed as part of national scale remote sensing campaigns, yet these systems are too costly to deploy at high frequencies or at study sites on an as-needed basis (Kampe et al. 2010; Schimel et al. 2011) and so remain out of reach for most field scientists. By enabling inherently fused structural and spectral measurements of vegetation from a single sensor, SFM may overcome many of the existing technical challenges of remote sensing fusion that occur when structure is measured from one sensor and spectral traits are measured from another. Despite this potential, SFM 3D-spectral fusion data quality has only been evaluated at coarse spatial scales (250 m x 250 m) relative to satellite remote sensing observations (Dandois and Ellis 2013) and not to the precision of individual points or at fine scale (e.g., ca. 0.1 m).

1.2 Bundler 'Photo-Tourism' SFM data model

Prior Ecosynth research used two different SFM algorithms for producing 3D point clouds of forests from images. Dandois and Ellis (2010) used the free and open source Bundler algorithm (Snavely et al. 2006), but later the commercial, closed source software Agisoft Photoscan (http://www.agisoft.ru) was used because it enabled faster reconstruction of large image datasets compared to Bundler (Chapter 2, Dandois and Ellis 2013). The current research uses Bundler because it enables access to information about

the SFM reconstruction process that cannot be accessed in Photoscan due to the commercial and proprietary nature of the later. Bundler was not developed with the specific intent of being applied to vegetation and the image datasets used in its development did not specifically contain vegetation as the focus of the reconstruction, (e.g., the Trevi Fountain and Coliseum in Rome, the Great Wall of China, and the Statue of Liberty). Bundler uses the Scale Invariant Feature Transform (SIFT; Lowe 2004) feature detector to identify potential features to match across images. Figure 1 diagrams the Bundler data model showing the link between point cloud 3D-RGB points, feature views on each point, and the SIFT feature descriptor associated with each view. While the terms are often used interchangeably, in this work the term 'detector' is used to refer to the feature detection algorithm and 'descriptor' is used to refer to that which is produced by the algorithm.

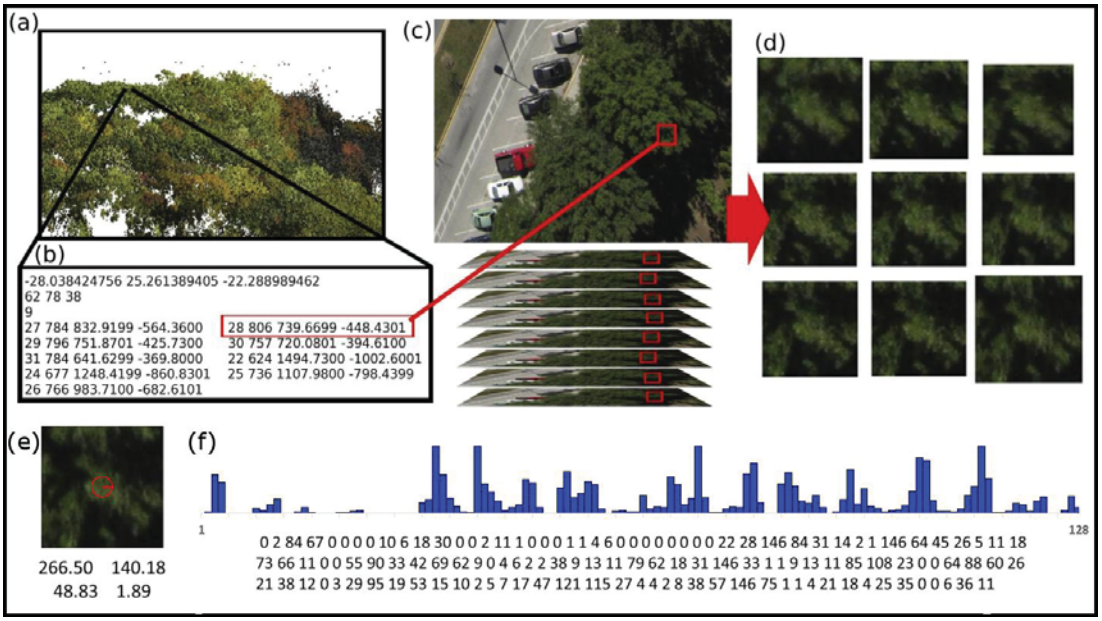


Figure 1: Diagram of Bundler data model: (a) a 3D-RGB point cloud, (b) a point 'view list', (c) a set of point 'views', (d) 'image feature tiles' associated with each view, (e) a single 'image feature tile' with a circle indicating the scale of the original feature, and (f) the 128-D SIFT image feature descriptor as a histogram and numerical vector.

In the Bundler data model, each point cloud point (Figure 1a) is represented by a vector of numbers that describe the point's XYZ location, RGB color, and the list of views on that point from multiple images (Figure 1b). In the example in Figure 1, the selected point was determined by the feature detector and feature matching to have been viewed in 9 images (Figure 1b & 1c), at a location in 3D space and the corresponding 2D locations in the image spaces by bundle adjustment. Each SIFT feature descriptor can be represented in two forms: as an 'image feature tile' (Figure 1e) and as a 128-dimensional (128D) numeric feature descriptor vector (Figure 1f). When viewed in reverse, Figure 1 also diagrams the SFM pipeline from the perspective of a single image feature from feature identification (1e & 1f) and matching (1d & 1c), to the representation of the feature as a point in the 3D point cloud produced by the bundle adjustment stage with color associated with each point (1b & 1a).

1.3 SIFT Scale Invariant Feature Transform feature detector

The image feature tile (Figure 1e) contains information about the XY location of the feature in the image, its scale in terms of pixel size, and a measure of its relative orientation based on the dominant gradient of intensities within each tile. The SIFT feature detector determines the location of feature points within gray-scale versions of the original images as locations of local minima or maxima of the Difference of Gaussian representation of the image within a scale-space pyramid (Lowe 2004). In other words, SIFT first produces a scale-space representation of the original image by iteratively blurring the image using a Gaussian filter and then resampling to a reduced resolution to simulate the effect of viewing the same scene from a greater distance (Lindeberg 1994; Hay et al. 2002). At each resampled resolution (scale octave) each Gaussian blurred

image is subtracted from its neighboring blurred image to produce a stack of Difference of Gaussian (DOG) images which represent distinct edges (Figure 1 in Lowe 2004). A feature point is located at any pixel in a DOG image that is a maximum or minimum relative to the pixels immediately around it and the pixels above or below in the adjacent DOG images (Figure 2 in Lowe 2004). The location and scale of the point identified in this step is represented by the red circle in Figure 1e. The image feature tile is much larger than the size of the point (roughly 6x larger) as SIFT computes the numeric feature descriptor describing image gradients across a square area around each point (Lowe 2004). The numeric feature descriptor, often displayed as a histogram (Figure 1f), represents the sum of the magnitude of gradients in 8 primary directions within a 4x4 grid of subregions overlaid on the image feature tile (Figure 7 in Lowe 2004).

SIFT has been used for many applications, including in remote sensing for automated image registration and multiple image fusion (de Matías et al. 2009; Huang et al. 2009; Lingua et al. 2009; Zhang et al. 2009; Schwind et al. 2010; Oh et al. 2011). Although it was not designed specifically for SFM reconstruction of vegetation, its initial testing included images of '…outdoor scenes, human faces, aerial photographs, and industrial images…' (p. 8, Lowe 2004). Recently, Yang and Newsam (2013) used SIFT to classify different land cover types in high-resolution aerial imagery. By applying multiple levels of k-means clustering to SIFT numeric descriptors, they found that groups of SIFT features within a given spatial area can be used to accurately discriminate among different landcover types (e.g., urban areas, suburbs, water, coastline, etc.). Similarly, SIFT feature descriptors linked to Bundler point cloud points were used to find the location of a particular object (the face of a clock tower) within a 3D point cloud of a city

plaza based on matching of the SIFT feature of the clock face in a reference image to the SIFT features of each point in the point cloud (Li et al. 2011). These studies highlight the important role that the feature detector plays in the SFM process with regard to what is being represented by the point cloud. In order to understand what SFM point cloud points represent, it is necessary to examine the features that were identified and matched at the point locations in images.

1.4 Research objectives

This research aims to improve understanding of the way in which forest canopies are represented by SFM 3D-spectral point clouds and to evaluate the empirical quality of those observations by focusing on two primary research questions: (1) What are SFM point cloud points? and (2) What is the quality of SFM 3D-spectral data fusion?. The quality of 3D-spectral fusion was evaluated by measuring the classification accuracy of painted targets placed in the tree under leaf-on and leaf-off conditions by SFM and also by a terrestrial laser scanner (TLS) with an attached calibrated camera for image fusion. To evaluate what SFM point cloud datasets are observing, i.e., what are points / what is 'seen', a custom graphical user interface was developed that prompts users to supply semantic tags (e.g., tree, grass, car) to the SIFT 'image features' identified at each point cloud point, which were then compared to clusters of SIFT features based on their numeric feature descriptors. It was hypothesized that if point cloud points represented real canopy objects like leaves, branches, and crowns, then clusters of SIFT image features would contain a relatively higher proportion of the semantic tags associated with those objects.

2.0 Materials and methods

2.1 Data collection

2.1.1 Ground-based scanning of a single tree 'test subject'

To evaluate the quality of SFM fusion 3D-color point clouds, a single free-standing tree was used as a 'test subject' to facilitate repeated imaging and to allow for placement of painted targets to guide analysis. A red maple (*Acer rubrum*; height: 5.2 m; DBH: 0.13 m) was scanned under leaf-on (2012-08-20) and leaf-off (2013-03-05) conditions using an Olympus E-PL2 DSLR digital camera on a 2.0 m pole and a RIEGL VZ-400 terrestrial LIDAR scanner (TLS) equipped with an Nikon D700 DLSR digital camera for built in LIDAR-color fusion. TLS LIDAR height was 1.6 m and the accessory camera mount was 0.4 m above this, or 2.0 m above ground. Prior to scanning, foam balls (diameters: 0.05 m, 0.07 m, 0.15 m) were painted matte red and hung throughout the tree (leaf-on: n = 11; leaf-off: n = 14). In each season, SFM digital images and TLS scans were collected at the same time at roughly mid-day (10:00 - 14:00) to minimize the effect of shadows, but changes in lighting based on relative location of camera and sun were unavoidable. Ten replicates of SFM digital images were taken at 2.5° / 0.3 m intervals around the entire tree at a 7 m radius. Four TLS laser scans plus digital images were collected from orthogonal positions at 7 m distance from the tree (i.e., north, east, south, and west). Single tree scan configuration is diagrammed in Appendix A2.1.

2.1.2 Canopy aerial imagery from hobbyist UAS

Aerial image data from prior research (Chapter 1; Dandois and Ellis 2013) collected using a point and shoot digital camera and a hobbyist multirotor unmanned aerial system

(UAS), plus one additional flight, was used to evaluate the relationship between SFM image features, vegetation objects, and other types of landcover. Aerial image datasets represent three Temperate Deciduous forest areas, captured under varying conditions of flight altitude, levels of cloud cover (clear and overcast), wind, and canopy phenological state (leaf-on, senescing, leaf-off), described in Appendix A2.2. For each flight a subset of approximately 1/3 of the original image set, or roughly 600 photos, covering a 250 m x 100 m area was used for the current study.

2.2 Data processing

2.2.1 TLS data processing

Leaf-on and leaf-off TLS scans were processed separately within the RISCAN-Pro software package (v1.7.3 release 6034) following manufacturer instructions. Reflective targets placed on tripods around the tree were used to automatically co-register individual TLS scans into a single 360° point cloud model. The same reflective targets were also used to manually refine the calibration between TLS scans and digital images collected from the on-board DSLR. RGB color from the digital images was then automatically 'fused' to the 3D point cloud based on the pixel color at the projection of each point into the corresponding image. TLS point clouds were then manually trimmed to only include points of the tree itself and exported into ASCII text files containing the XYZ location and RGB color of points. A set of 10 random samples of 90,000 points each was generated for each TLS point cloud as a form of replication to match the number and approximate average point count of SFM point cloud replicates. A detailed description of TLS data-processing is provided in Appendix A2.3.

2.2.2 SFM 3D-RGB point clouds from digital images

Three-dimensional (3D) multi-spectral (red-green-blue; RGB) point clouds were generated from ground and aerial image datasets using the open source Bundler SFM algorithm (v0.4, Snavely et al. 2010, available online: https://www.cs.cornell.edu/~snavely/bundler/, accessed: 2012-08-23). As distributed, Bundler uses of the Scale Invariant Feature Transform feature detector algorithm to identify keypoints for matching across images (SIFT; Lowe 2004; available online: http://www.cs.ubc.ca/~lowe/keypoints/, accessed: 2012-08-23). Computation required on average 12 - 24 hours for ground image datasets and 27 - 552 hours for aerial image datasets on a dual Intel Xeon X5670 workstation (12 compute cores) with 48GB of RAM running 64-bit Ubuntu 12.04. Aerial point cloud datasets were georeferenced to the WGS84 UTM Zone 18N projected coordinate system based on UAS GPS telemetry and filtered using the Python-based free and open source Ecosynth online tools (http://code.ecosynth.org/EcosynthAerial, accessed: 2014-05-01) following Dandois and Ellis (2013). Single tree point cloud datasets were manually trimmed to include only the points of the tree. To compare the 3D structure derived from TLS and SFM point clouds, it was necessary to spatially align the point clouds into the same coordinate system (Henning and Radtke 2006; Hosoi et al. 2010). All SFM and TLS point clouds were spatially co-registered by the iterative closest point (ICP) alignment method (Besl and McKay 1992) within the open-source software Meshlab (http://meshlab.sourceforge.net/, v1.3.2 64-bit, accessed: 2014-05-01). A detailed description of the Meshlab ICP co-registration workflow is provided in Appendix A2.4.

2.2.3 Extracting points at painted targets

Threshold filtering based on the amplitude of returned laser energy was used to identify TLS points at painted targets (Dassot et al. 2011; Seielstad et al. 2011). Reflected laser energy at targets was relatively higher than that of the surrounding foliage, but similar to large branches, so threshold-filtered point clouds were manually trimmed to remove any non-target points. For each target, a 'target area' was defined based on the TLS points associated with that target as a cube centered on the average of XYZ coordinates with a side length equal to the average range in each XYZ dimension. Points in each TLS and SFM replicate falling inside the target areas were set aside for additional analysis.

2.2.4 Extracting image features for SFM aerial point clouds

To evaluate whether SFM points can be linked to discrete vegetation objects (e.g., leaves, branches, crowns, etc.), a collection of point 'image features' was extracted for a simple random sample of 250 points from each SFM aerial point cloud for a total of 2000 points. For each sampled point, the 128D SIFT feature descriptor and a thumbnail image feature tile were created based on the first view of the point from the Bundler view list (Figure 1b). The thumbnail image feature tile was centered on the XY location of the point within the original image and of an approximate size to match the region over which the SIFT descriptor was evaluated (Li et al. 2011). Summary RGB color statistics were computed for the entire image feature tile as well as the smaller region in the center of the tile where the SIFT feature was located in scale space (Figure 1e).

2.3 Data analysis

2.3.1. Evaluation of 3D-RGB fusion quality

The accuracy of SFM fusion was evaluated by calculating the dominant hue color value of points observed in the target area (2.2.3) relative to the red color of the painted targets, as defined by the range of hue color values (330° – 20°) manually extracted from images of the targets (Manjunath et al. 2001; Lee et al. 2003; Aptoula and Lefèvre 2009). A target was classified as 'observed' if the point cloud replicate contained more than 1 point within the target 3D search area. A 'correct' 3D-color fusion observation was one in which on average > 50 % of the points observed at a target location had a hue color value within the red range. It was not logistically possible to place painted targets in the forest canopy to evaluate SFM fusion quality in aerial images in the same way as was done for the tree test subject. Instead, several empirical measures of color quality were evaluated directly from samples of aerial SFM points by comparing point color to the mean color of corresponding image pixels for the first image feature tile associated with each point (i.e., the smaller point 'blob' and not the larger support region, Figure 1e). Measures of point to image feature tile color quality included the difference in the hue color value as well as the correlation and ratio between RGB channel values, measures of spectral quality applied in evaluation of remote sensing imagery (Teillet et al. 2001; Biggar 2003). The same measures were also calculated for ground-based SFM point clouds of the tree test subject.

2.3.2 Evaluation of SFM image features

To examine what SIFT features are observing, an interactive graphical user interface (GUI) was developed that allows the user to see each image feature tile, its relative

location within an original image, and tag the feature with semantic tags (Appendix A2.5). A set of common tags spanning five descriptive categories (color, shape, surface, vegetation objects, and other objects; Appendix A2.6) was produced after two rounds of 'training' in which multiple user groups were asked to supply semantic tags of image feature tiles based on their own interpretation. The sample of image features from SFM aerial point clouds (2.2.4) was then presented to a group of users, who then supplied tags for each feature. K-means clustering was then carried out on the image feature 128D SIFT keys only at multiple levels of k. Clustering stability analysis was used to determine the most stable number of clusters (Manjunath et al. 2001; Lange et al. 2004; Jain 2010). The frequency of manually identified tags for each cluster was then plotted to estimate the association of clusters to distinct semantic classes.

3.0 Results

3.1 Evaluation of SFM point cloud fusion quality

SFM and TLS fusion accuracy is reported in Table 1 in the form of classification error matrices for leaf-on and leaf-off ground-based point clouds of the tree test subject (Congalton 1991). Average TLS and SFM fusion classifications are reported in the same error matrix for either leaf-on or leaf-off datasets to facilitate comparisons between the two methods, but the TLS data is not considered as ground-truth reference. Instead, the reference classification is simply that all targets (n = 11 leaf-on; n = 14 leaf-off) should be classified as red.

Table 1: Target classification accuracy as error matrices between leaf-on and leaf-off TLS and SFM point clouds as the average observation rates per target across repetitions. A target was classified as red if more than 50 % of the points found within the target 3D search area had a red hue (330° – 20°). Breakdown of classification of points per target and per replicate are provided in Appendix A2.7.

(a)

SFM Leaf-on \ TLS Leaf-on	Red	Not red	Sum
Red	1	2	3
Not red	4	4	8
Sum	5	6	11

SFM Accuracy = 3/11 = 27 %
TLS Accuracy = 5/11 = 45 %
Overall Agreement = 5/11 = 45 %
Kappa = -0.14

(b)

SFM Leaf-off \ TLS Leaf-off	Red	Not red	Sum
Red	4	9	13
Not red	0	1	1
Sum	4	10	14

SFM Accuracy = 13/14 = 93 %
TLS Accuracy = 4/14 = 29 %
Overall Agreement = 5/14 = 36 %
Kappa = 0.06

Classification accuracies differed by fusion method and by season based on observation rates of target points and the colors assigned to those points. SFM classification accuracy was highest under leaf-off conditions (93 %; Table 1b) where all targets were observed across all replicates, the average number of points per target was relatively high (29 points per target, range 3 - 105), and on average 60 % of points had a red hue (Figure 2b). SFM classification accuracy was much lower under leaf-on conditions (27 %; Table 1a) where on average 59 % of targets were observed, the average number of points per target was relatively low (3 points per target, range 0 - 27), and only 40% of points were classified with a red hue (Figure 2a). TLS classification accuracies were higher than SFM in leaf-on conditions and lower in leaf-off conditions (45 % and 29 %, respectively; Table 1). TLS observed all targets under leaf-on and leaf-off conditions and with relatively higher density than SFM under the same conditions (leaf-on: 11 points per target, range 0 - 45; leaf-off: 41 points per target, range 12 - 77), but points were less likely to have a red hue (39 % and 37 %, respectively; Figure 2c and 2d). A single target in a single TLS leaf-on sub-sampled replicate was not observed by any

points. Under leaf-on conditions TLS points were no more likely than SFM points to be red at the right location (TLS: 39 % vs. SFM: 40 %, Figure 2a and 2c). Under leaf-off conditions TLS and SFM observed the same number of targets, but SFM points were almost twice as likely to be red (TLS: 37 % vs. SFM: 60 %). The result is that SFM had higher classification accuracy than TLS under leaf-off conditions and lower accuracy under leaf-on conditions. There was low agreement between TLS and SFM classification of painted targets (45 % and 36 %) and the very small kappa values indicate that any agreement is likely by chance alone, suggesting that the two systems 'see' the tree and targets differently.

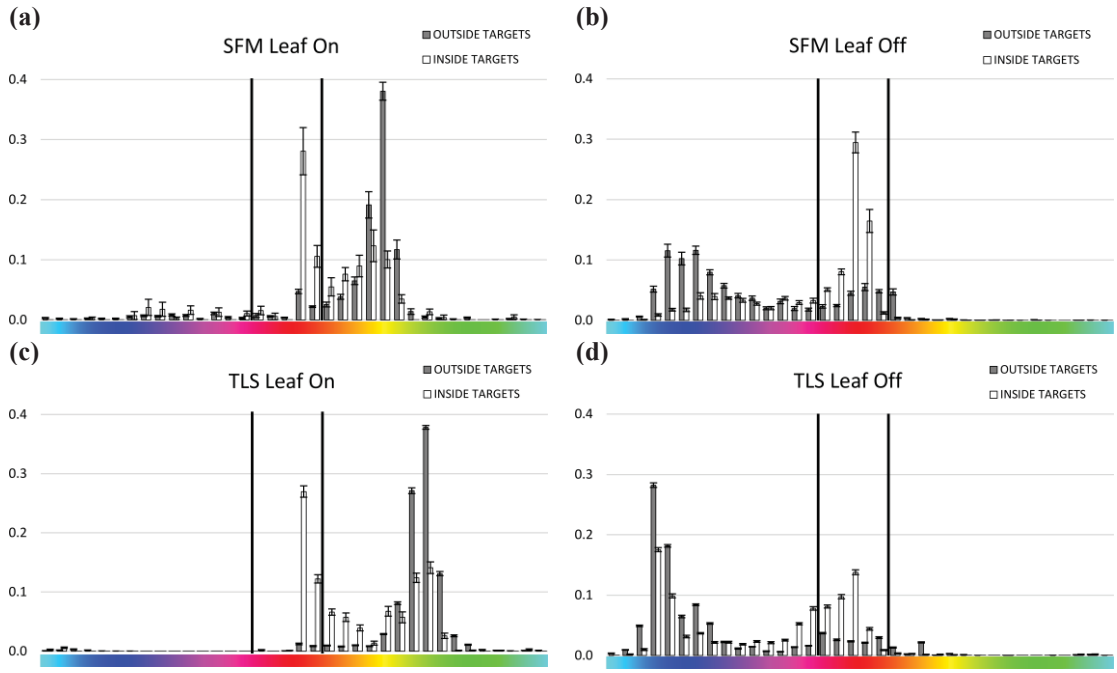


Figure 2: Mean histograms of point hue for points that fell 'inside' the target zone (white bars) and points that fell 'outside' the target zone (gray bars) for SFM (top row) and TLS (bottom row) point clouds under leaf-on and leaf-off conditions. Error bars are standard error (n=10). Black vertical lines are red hue cutoff region (330° – 20°).

Differences in fusion classification accuracy due to the color assigned to points can been be seen as peaks in the hue histogram (Figure 2) which provide an indication of dominant colors within a sample (Aptoula and Lefèvre 2009). Under leaf-on and leaf-off

conditions, the color of TLS points inside and outside the target areas were similar, with a large number of points inside the target areas showing non-target, non-red colors, contributing to low overall accuracy (45 % and 29 %, respectively). Similarly SFM leaf-on point clouds also had a large number of non-red points inside the target areas, similar to the color of points outside target areas, resulting in relatively low accuracy (27 %). The highest accuracy in target classification was obtained by SFM under leaf-off conditions (93 %) where there were few non-red points inside target areas.

Examples of low observation rates of targets in the SFM leaf-on point clouds are diagrammed in Figure 3, which shows the location of all SIFT feature points assigned to a single image view on the tree and three targets that were not assigned 3D point cloud points for this view (Figure 3a). Green triangles indicate the location of all SIFT features located within the image and pink circles highlight only those points that were included in the 3D point cloud from this 'view', representing approximately 7 % of all points identified in the image. The insets show 3 examples of painted red targets where no point cloud points were identified for this image and highlight the complex role of the feature detector in 3D-spectral fusion. Insets show a zoomed-in region around the targets in (3a) in RGB overlaid with SIFT feature points and in gray-scale without feature points. Gray-scale version is provided for visual reference since SIFT identifies features in scale-space based on a gray-scale version of the image. Inset (b) shows where SIFT identified a point at a fully visible target, but the point was not used in the point cloud for this view. Inset (c) shows a target that was completely visible, but where points were only identified around the edge of the target. Inset (d) shows a target that was partially visible but where no points were generated. Inset (d) also highlights a target where no points were

observed in the target area under any SFM leaf-on replicate. Targets that were not observed in SFM leaf-on point clouds were > 1 m from the outer hull of the tree and there was a weak negative relationship between the average rate at which a target was observed and distance to the outer hull of the tree ($R^2 = 0.34$ p-value < 0.1).

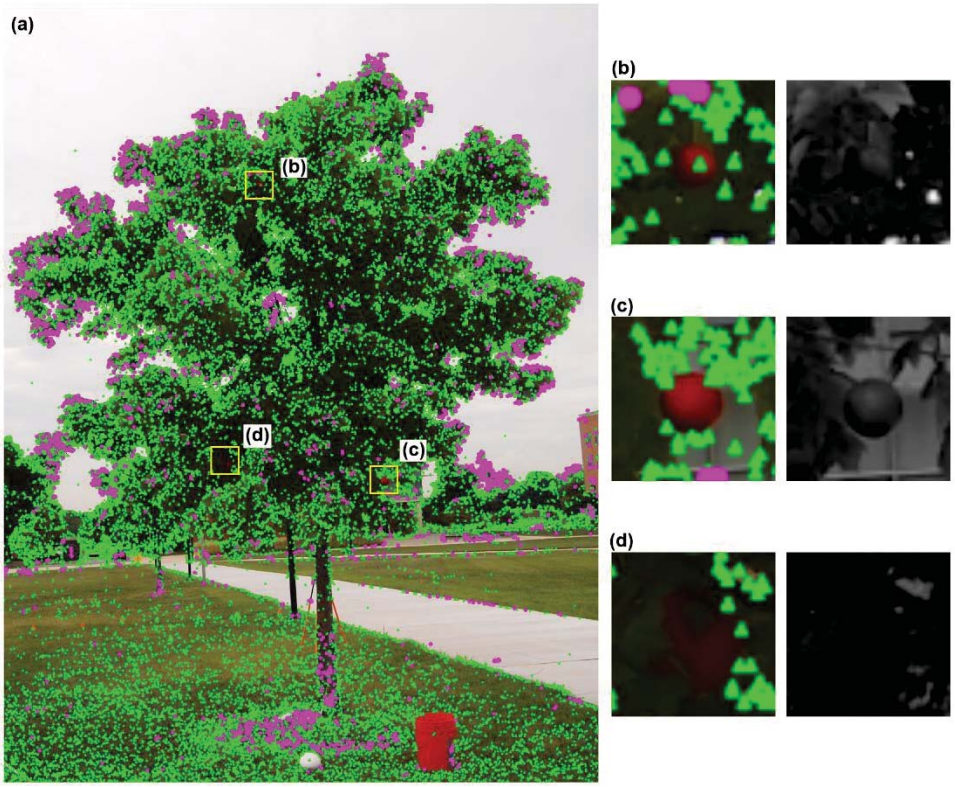


Figure 3: Example of the role of the SIFT detector in the placement of point cloud points. (a) original image overlaid with all SIFT points (green triangles) and point cloud points (pink circles), (b) – (d) insets show targets that were not assigned a point cloud point for this view with the associated gray-scale version of each inset at right.

Mean absolute difference between point hue and average hue image feature tile hue (Figure 1e) ranged from 1.8 % - 8.4 % with an average of 5.1 % across all aerial and ground SFM point clouds (Table 2). RGB channel values were highly correlated between points and image feature tiles (average $R^2 = 0.83$), with the exception of the leaf-off ground target set, where correlations were much lower ($R^2 = 0.42$ - 0.43). Average RGB channel ratio between point and image feature tile color, a measure for spectral

calibration of remote sensing (Teillet et al. 2001; Biggar et al. 2003), showed close agreement (average 1.06) but also large range +/- 25 % (ratio 0.77 - 1.28).

Table 2: SFM point color precision as measured by the difference between point color and average color of pixels at the corresponding point location within original images for random samples of points from aerial image sets and points located within target areas for ground sets.

Sample Dataset	Point vs. image feature hue Absolute Difference Mean (SD) (as percent)	Point vs. image feature RGB Channel Correlation (R^2)			Point vs. image feature RGB Channel Average ratio		
		R	G	B	R	G	B
Ground Leaf-on targets [a]	5.1 % (8 %)	0.80	0.82	0.84	1.02	1.03	1.02
Ground Leaf-off targets [a]	8.4 % (10 %)	0.43	0.42	0.43	1.12	1.27	1.24
HR_2010_10_06 [b]	4.3 % (6 %)	0.96	0.95	0.96	1.01	1.01	1.02
HR_2011_03_05 [b]	7.5 % (11 %)	0.85	0.87	0.87	1.00	1.01	1.01
HR_2012_08_24 [b]	1.8 % (2 %)	0.86	0.84	0.83	1.05	1.03	1.18
KN_2010_10_08 [b]	4.3 % (7 %)	0.90	0.88	0.89	1.10	1.08	1.14
KN_2010_10_24 [b]	4.3 % (6 %)	0.96	0.95	0.96	1.01	1.01	1.02
KN_2011_03_11 [b]	7 % (11 %)	0.93	0.93	0.93	1.01	1.01	1.01
SE_2011_10_08 [b]	3.4 % (6 %)	0.74	0.74	0.57	1.10	1.08	1.28
SE_2012_02_26 [b]	5.2 % (9 %)	0.93	0.93	0.91	1.01	1.01	1.02
Mean	5.1 %	0.84	0.83	0.82	1.04	1.05	1.09
SD	2 %	0.16	0.16	0.18	0.05	0.08	0.11

[a] Reported values are for the aggregate of all target points from all replicates

[b] Reported values are for all points within the simple random of sample of 250 points extracted from each point cloud.

Evaluation of the 3D structure quality of SFM compared to TLS is diagrammed in Figures 4 and 5. Figure 4 shows a comparison of 3D voxel (volumetric pixels, 0.001 m^3 resolution) models of the tree for a single SFM and TLS replicate under leaf-on (a) and leaf-off conditions (b). Panels show a single 0.1 m 'vertical' slice at the middle of the tree along the same axes where voxels contained at least one point, similar to a medical 'tomographic' image, and characterize where tree structure (e.g., foliage, branches) was observed by the scanning system (Holden et al. 2000; Parker et al. 2004).

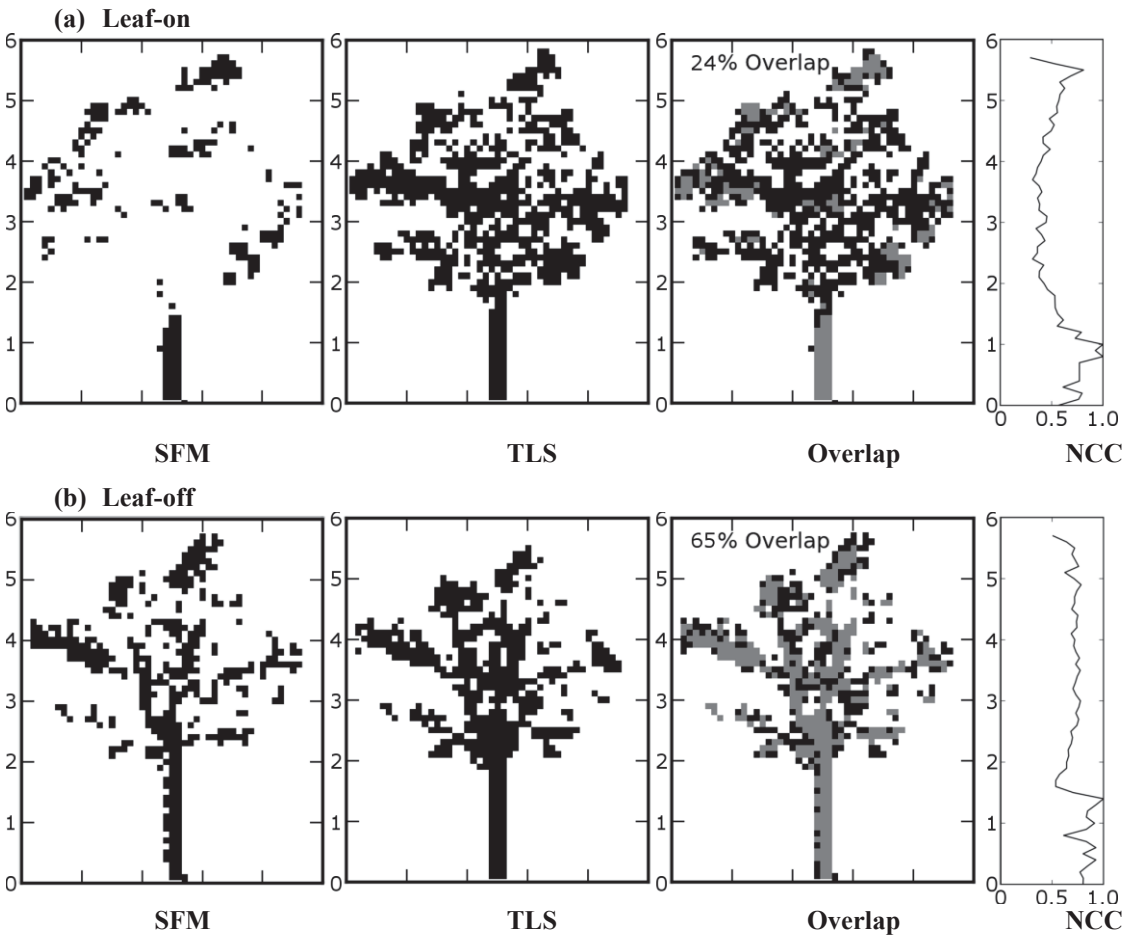


Figure 4: Tomographic slices of a 0.1 m section of leaf-on (a) and leaf-off (b) SFM and TLS point clouds. Pixels represent a 0.1 m cube within which at least 1 point was located. All slices are co-registered to the same coordinate system and viewpoint.

Right panels shows the vertical profile of the normalized cross-correlation (NCC; Holden et al. 2000) as a measure of similarity between SFM and TLS for each 0.1 m 'horizontal' slice along the height of the tree. Under leaf-on conditions, SFM observed little of the interior the tree compared to TLS (Figure 4a, NCC < 0.5), with low voxel overlap (24 %) predominantly along the trunk and exterior of the tree. A much different pattern was observed for leaf-off models (Figure 4b) where SFM generated a more comparable 3D profile of the tree compared to TLS and with greater overlap (NCC > 0.5, 65 %). Combining leaf-on and leaf-off models improved NCC and overlap relative to leaf-on models alone (Appendix A2.8, NCC > 0.5, overlap 54 %).

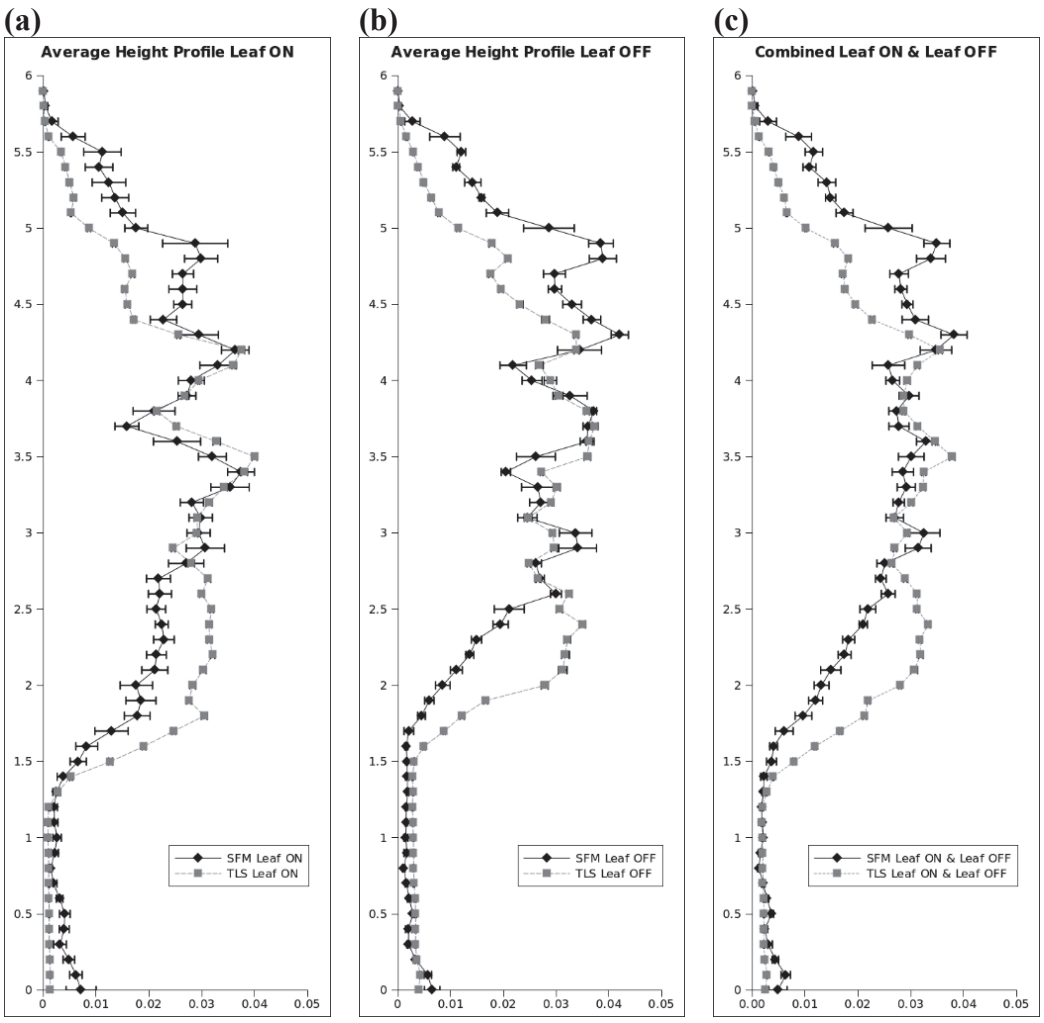


Figure 5: Mean vertical foliage profiles (VFPs) for SFM and TLS-subsampled point clouds under leaf-on and leaf-off conditions, and with leaf-on and leaf-off point clouds combined. Differences between SFM and TLS VFPs were not statistically significant based on Kolmogorov-Smirnov (K-S) test of distributions ($p < 0.0001$).

Despite differences in the way that the interior of the tree is observed by SFM and TLS, both methods produced similar vertical foliage profiles (VFPs) under leaf-on and leaf-off conditions and when leaf-on and leaf-off datasets were combined (Figure 5). VFPs were not significantly different based on Kolmogorov-Smirnov (K-S) tests of distributions (Parker and Tibbs 2004) on pairs of SFM and TLS replicates (average p-values < 0.0001), but showed distinct differences at upper and lower parts of the test tree across comparisons.

Although TLS and SFM VFPs were not significantly different based on K-S tests of distributions, the profiles show a consistent pattern of over and under estimation of vertical density. This does not immediately appear related to large differences in 3D structure of each point cloud, NCC between leaf-off TLS and SFM replicates was relatively high, indicating close agreement in overall 3D structure. The differences may be related to the way in which the feature detector governs the way that 3D space is 'sampled' in a SFM system. SIFT appeared to favor the outer edges of the tree, placing more points per image and more points that were matched to the point cloud at the upper edges of the crown compared to images of the core of the crown and lower half (Figures 2 and 6).

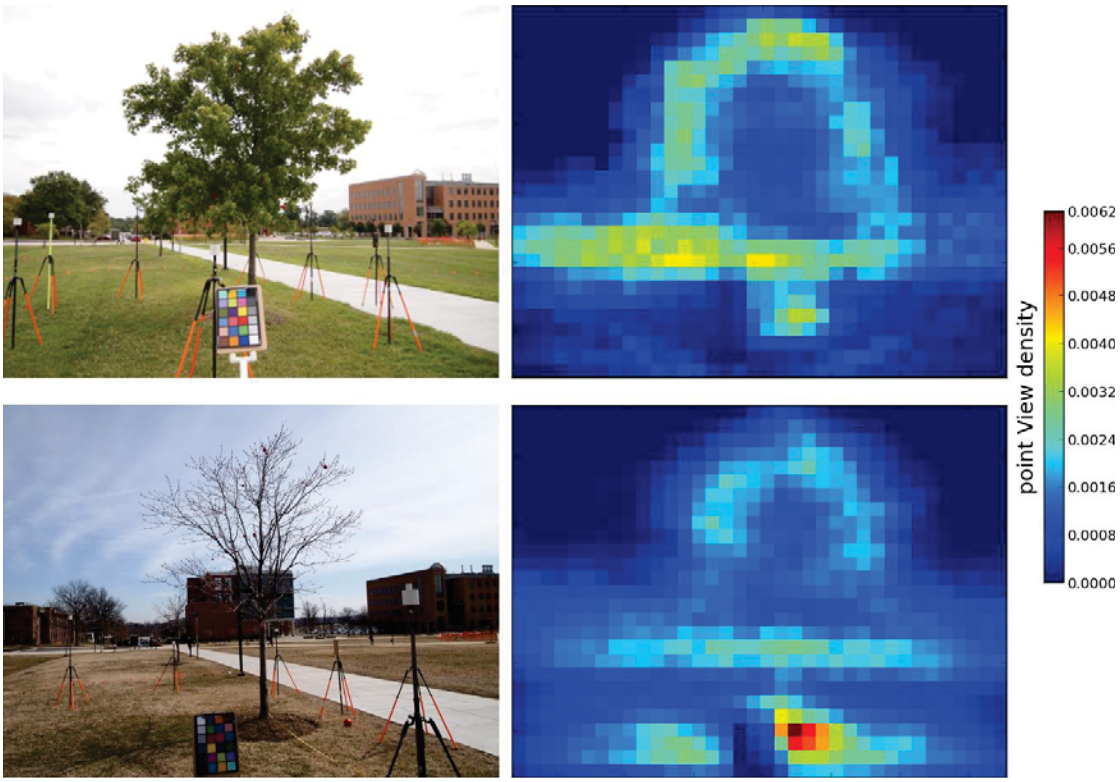


Figure 6: Example images used in single tree SFM reconstruction under leaf on (top) and leaf off (bottom) conditions and image view maps showing the density of where point cloud points were viewed from the perspective of the images. This corresponds to the density of points across all images for a single SFM replicate as diagrammed by pink circles in Figure 3.

3.2 Evaluation of point cloud image features

K-means clustering of the 128D SIFT key descriptors from SFM point cloud features consistently showed that image feature tiles cluster into two main groupings: points that are relatively brighter than the surroundings and points that are relatively darker than the immediate surroundings (Figure 7). Right panels in Figure 7 highlight this distinctive pattern by showing the average gray-scale intensity of all image feature tiles associated with a cluster, after tiles were resized to the same dimensions. The point is located at the center of each image tile and the 128D feature descriptor is determined based on image content around that point (Figure 1).

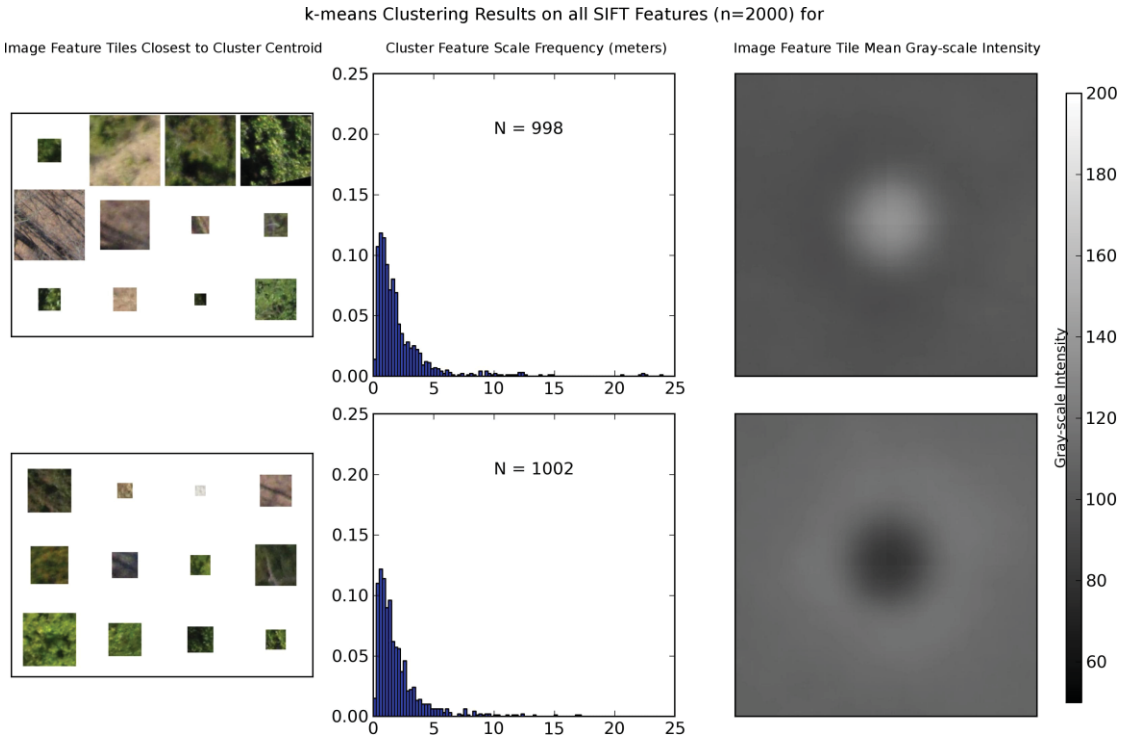


Figure 7: Clustering results (k =2) on SIFT 128D descriptor for 2000 points from aerial SFM point clouds. Left panels show 12 image feature tiles closest to cluster centroid. Center panels show the frequency distribution of the scale (meters) of cluster features. Right panels show the 'mean gray-scale intensity' of all cluster image feature tiles resized to the same size (200 pixels square).

There was no significant difference between the metric scale of image feature tiles associated with each cluster based on Kruskal-Wallis (K-W) non-parametric test of samples (Figure 7, center panels). Visual inspection of exemplar image feature tiles per cluster, features closest to the cluster centroid, revealed no distinct patterns in image content (Figure 7, left panels). These patterns were observed when clustering was performed on the entire set of 2000 features pooled from all datasets, for each individual set of 250 features sampled from each aerial dataset, and also for samples of 250 features from ground-based leaf-on and leaf-off point clouds of the single test tree (Appendix A 2.10). Comparison of the gray-scale intensity of point color to the mean gray-scale intensity of the entire image feature tile around a point revealed the same pattern, with the sign of the difference between point and tile intensity being directly related to cluster association for > 99 % of points (Figure 8).

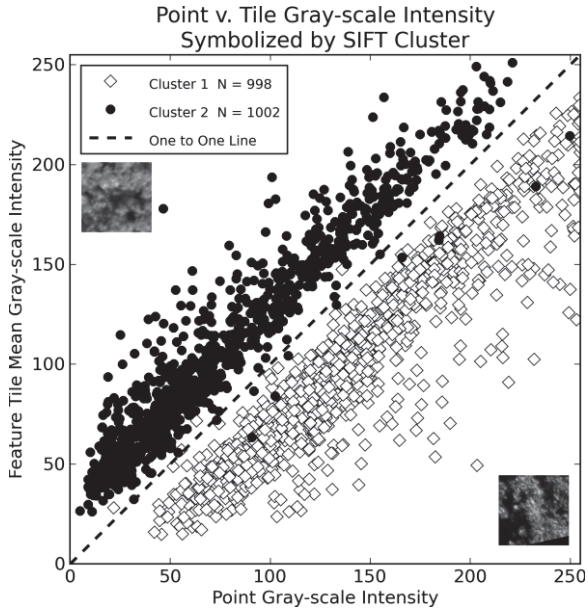


Figure 8: Scatter plot of the gray-scale intensity of points versus the mean gray-scale intensity of the image feature tile around each point, symbolized by the same SIFT clusters as in Figure 7. Points located below the one-to-one line have a point that is brighter than its surroundings, points above the line are darker than the surroundings, as shown in the example images.

For all clustering results, the number of clusters k was determined by cluster stability analysis (Manjunath et al. 2001; Lange et al. 2004; Jain 2010), which consistently identified only the two main stable clusters per sample dataset when all sample points were supplied for clustering (Figure 7, Appendix A2.10). Cluster association of SIFT descriptors followed the pattern of representing either bright or dark spots as k was increased (e.g., k=10), but it was also observed that K-W tests showed that the distributions of feature scale within clusters was significantly different, potentially due to the presence of few very large features > 10 m in size (Appendix A2.11).

Results of manual identification ('tagging') of image feature tiles are shown in Figure 9. Clustering was performed separately on bright and dark points (Figure 8) for leaf-on and leaf-off sets. In general, the frequency of tags per cluster did not reveal a strong association between clusters of SIFT image feature tiles and semantic categories of objects (e.g., leaves in one cluster, branches in another, etc.) Instead, many of the same tags appeared frequently within the top 50 % and 85 % of all tags associated with a single cluster. Tags describing feature color (e.g., green, white, brown, etc.) appeared frequently across all clusters, which is to be expected since the SIFT algorithm identifies features within a gray-scale version of the image only and SIFT feature descriptors would not contain color information (Figure 3). There were several unique tags within the top 50 % of each cluster, revealing two broad cluster groupings across leaf-on and leaf-off point clouds: points of leafy green foliage from trees with the unique tag 'single tree/crown' and points of gray pavement and linear features with the unique tags of 'pavement/road/asphalt' and 'line'. Even so, these tags occurred with relatively low frequency overall (< 10 %), which suggests that they are not highly representative of the

semantic identity of all features within that cluster. Clusters of bright and dark points from leaf-off aerial point clouds also showed a significant difference in the scale of features within clusters, which appeared to be associated with the presence of very large features > 10 m in size within a single cluster per group (Appendix A2.12). In general, the results of comparing the manual classification of image feature semantic content with clustering of SIFT 128D feature descriptors did not support the hypothesis that feature clusters are associated with distinct groups of canopy objects.

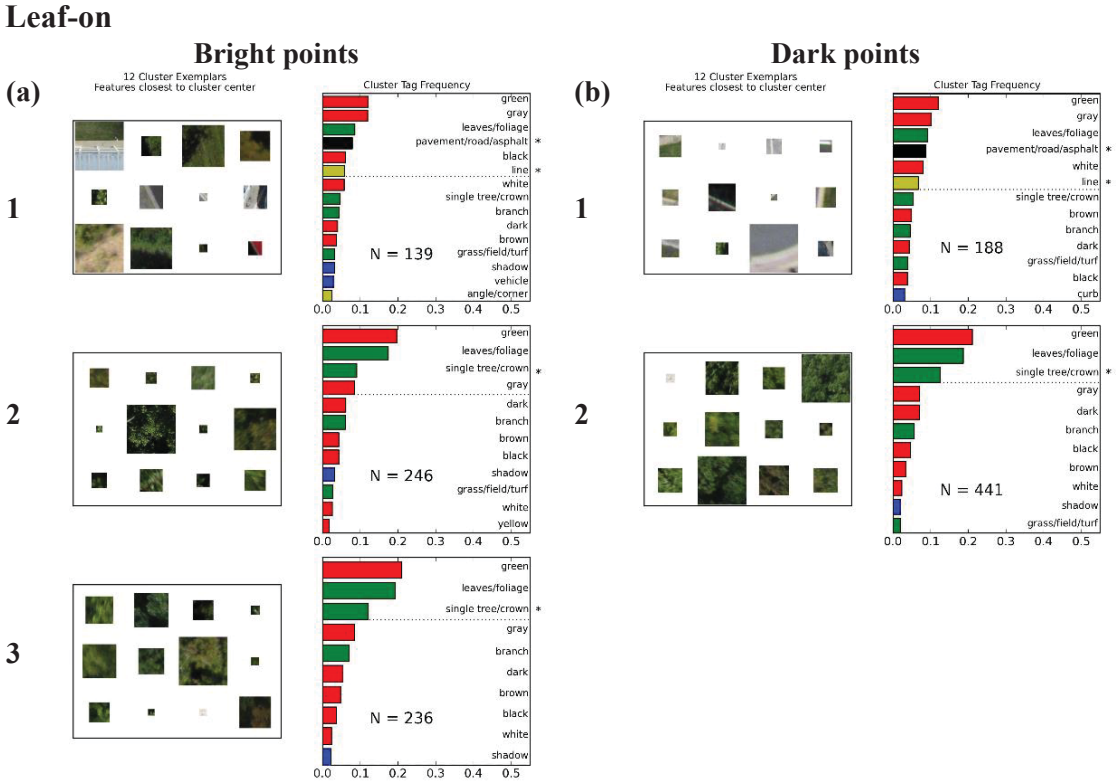


Figure 9: Sorted frequency distributions of manually identified 'tags' per image feature cluster for leaf-on aerial point cloud sample datasets. Clustering was performed separately on bright and dark points (Figure 8). Cluster frequency plots are truncated to the top 85 % to aid readability. Dotted horizontal lines indicate the 50th percentile. Starred tags ('*') are unique non-color tags in the top 50 % within the clustered dataset.

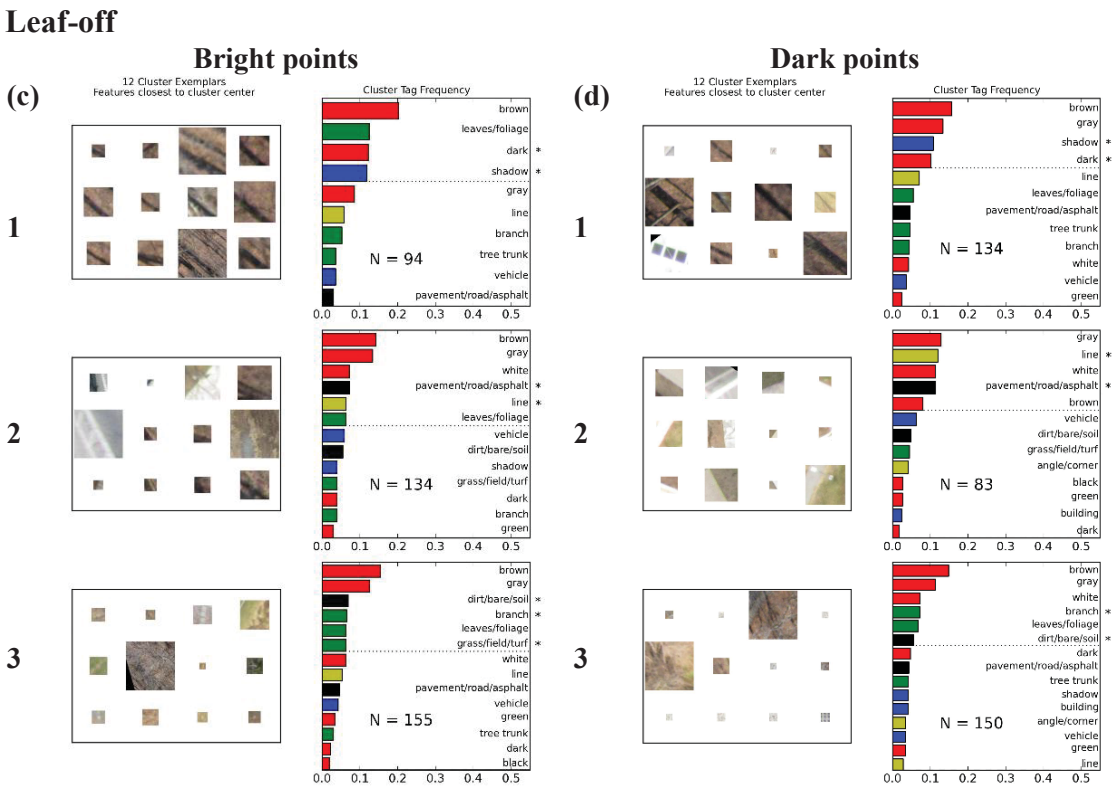


Figure 9: Continued for features from leaf-off point clouds.

4.0 Discussion

The objectives of this research were to evaluate the empirical quality of SFM 3D-spectral observations and to improve understanding of the way in which forest canopies are represented by SFM point cloud points and was motivated by two overarching questions: What is a point cloud point? and What is the quality of the inherent 3D-spectral fusion of a point cloud point? Two related research questions were posed to guide this research: Does an individual point cloud point represent a distinct canopy object like leaf, branch or part of a crown? and Is the color of a point in the correct location in 3D space?

4.1 What is a point? Whatever the feature detector wants it to be

The results indicate that what the points represent in the observed scene and how well they do so in terms of the fusion of structural and spectral information is largely dependent up on the feature detector that is used for matching points across images. Findings of this research are summarized in Table 3 as a collection of the three primary traits of a SFM point cloud point.

Table 3: What is a point? Three primary traits of a SIFT-based SFM point cloud point

	Image Sample	**Numeric Feature Descriptor**	**3D coordinate and RGB color**
Description	A portion or sample of the original image, with a specific XY coordinate, scale and rotation with respect to the image, as determined by the feature detector.	A vector of numeric values generated on a support region of the image around the point which is used for matching, as defined by the feature detector.	Point with XYZ coordinates and RGB color corresponding to where the point was located in the image and obtained by an interpolation around point location
Quality	SIFT only locates points at parts of the image that are brighter or darker than the surroundings.	Individually, describe that a bright or dark spot was located, but do not contain sufficient information to be linked to semantic category.	Point RGB color is closely matched to image color in the same location (5 – 6 % difference). Point RGB color is also descriptive of object color, if the object can be 'seen' by the detector (93 % target classification accuracy when all targets were seen).
Implications	3D sampling of a scene is determined by what the detector will 'see' which controls where points are located.	Classification of points and point clouds by 'feature' category based on descriptors may require that features are grouped first into collections or 'bags'.	SFM color and structure fusion does not have occlusion effects that occur when two different sensors are used (e.g., a LIDAR and a camera), but instead will have greater errors of omission where 3D space is not sampled because it was not seen in the images or was not sampled due to the feature detector.
Future Research	Examine the use of other detectors, perhaps ones designed for sampling vegetation. Examine how sampling of vegetation or parts of vegetation varies with different detectors.	Image features combined with object delineations from point clouds could be used for species or canopy object mapping (e.g., fruits, flowers, etc.).	Examine relationship between SFM color fusion and types of features or other fine-scale measurements of vegetation color (e.g., high resolution aerial imagery, field spectrometer).

In the current research the SIFT (Scale invariant feature transform; Lowe 2004) was used by the Bundler SFM algorithm for feature detection and matching. In many ways, the results of this research are strongly dependent upon the fact that SIFT was used instead of another algorithm and it should be expected that summary of conclusions in Table 3 would possibly be very different if a different feature detector were used. For example, the sampling of the image space by the SIFT detector appeared strongly dependent upon the fact that SIFT searches for features within a scale-space of a gray-scale version of the image (Figure 3), with the result that some painted targets, although clearly visible in RGB space, were not as visible in gray-scale and therefore were not sampled. Other feature detectors may have produced different and even better results for finding the painted targets, for example an implementation of SIFT that makes use of color information like hue or a combination of red and green (Van de Sande et al. 2010). The discussion below reviews the results that are summarized in Table 3 with particular attention to the role of the SIFT feature in understanding the results. Rather than being a weakness of the current research, the fact that the observation of tree and canopy vegetation by SFM is so strongly dependent upon the choice of feature detector presents exciting new opportunities for computer vision ecology that are discussed in the sections below.

4.2 Point as an image sample

The results suggest that the way in which vegetation in a scene is sampled or 'seen' by SFM is strongly linked to the feature detector. Analysis of the image features associated with each point based on image feature tiles and numeric descriptors (Figure 1) revealed that the SIFT algorithm placed points within the scene at places that were either brighter

or darker than the surroundings (Figures 7, 8). SIFT features were observed across a range of spatial scales (1 cm - > 20 m) and in general there did not appear to be any relationship between clusters of SIFT 128D numeric descriptors and the scale of the feature represented by the descriptor. However, statistically significant differences in the scale of features associated with clusters was observed among bright and dark points from leaf-off aerial point cloud sets and also when clustering was applied to all aerial point cloud points (N = 2000) at larger levels of k than determined by cluster stability analysis. It was unclear whether this was due to the presence of very large features (> 10 m in size) within certain clusters or some other factor (Appendices A2.11 and A2.12). Even so, these results are to be expected given the intent and implementation of the SIFT detector. SIFT is designed to find 'scale-invariant' features that in theory provide a representation of an image feature that will be similar across a range of viewing distances by searching for the most stable features within a scale-space (Lindeberg 1994; Lowe 2004). To do this, SIFT identifies local minima and maxima within the scale-space representation of the image, with the result that it responds most strongly to places within the image with relatively high contrast that are either brighter or darker than the surroundings. This means that points are only placed at the location where SIFT was able to identify the specific pattern that it was programmed to identify. Parts of the image where no feature was produced, or which were unable to be successfully matched to the same feature in other images will not be 'sampled' in the point cloud, regardless of whether something was visible in an image (Figure 3). Similar situations arise in other forms of remote sensing. For example, spectral information of the land surface in optical satellite imagery is absent or significantly altered in areas that are heavily shadowed due

to clouds (Huang et al. 2010). With infrared LIDAR, no returns are recorded over areas of water because the laser pulse was absorbed instead of reflected back to the recording instrument (McKean et al. 2009), and LIDAR will sample 3D space at different rates due to distance to the sensor and the angle of incidence between the land surface and laser pulse, among many factors (Glennie 2007).

4.3 Point as a feature descriptor

Clustering analysis of SIFT numeric feature descriptors revealed that points are located at places in the image that are either brighter or darker than the immediate surroundings. The 128D numeric descriptors consistently clustered into either of these two categories across all datasets and for data collected under a variety of conditions. SIFT numeric descriptors were also evaluated to assess the degree to which a single point descriptor contains sufficient information to be classified as a particular type of landscape feature (car, grass, canopy, etc.,). It was hypothesized that if SIFT feature descriptors contained a unique pattern for different types of objects or landcovers, then the tags describing those objects or landcovers should occur with greater frequency in the related 128D descriptor cluster. Comparison of the frequency with which manually identified tags occurred within clusters of SIFT features revealed no strong connection to the semantic content of images (e.g., leaves vs. branches vs. cars, Figure 9). This suggests that similarities among image feature numeric descriptors that are represented by k-means clusters are more strongly associated with patterns of bright or dark spots than the actual content of the images (e.g., tree cover, leaves, pavement, etc.). In other words, there was no strong evidence to suggest that SFM point cloud points are related actual canopy

objects, instead the results (Figures 7, 8 and 9) point strongly to the conclusion that points only represent the pattern that the feature descriptor was programmed to identify.

These results are in contrast with other studies that showed a strong relationship between SIFT image feature descriptors and different types of land cover observed in high resolution aerial images (Newsam and Yang 2007; Yang and Newsam 2008, 2013). However, there was a significant difference between the methods applied in that study compared to ours. Instead of evaluating the classification of features individually, Yang and Newsam (2013) applied a 'bag of features' approach (Lazebnik et al. 2006; Zhang et al. 2006) to classify specific spatial regions of high-resolution images based on the descriptors found within each region and found strong agreement with a high degree of separation across different types of landcover. The conceptual difference between that study and ours was that we began with the hypothesis that a single image feature contained information that could be linked to a particular kind of landcover. Yang and Newsam (2013) begin with the hypothesis that collection of features together in a 'feature bag' contain sufficient information to be classified or segmented into a particular landcover.

4.4 Point as a 3D-RGB point

Results of evaluating SFM 3D fusion quality indicate that there are two important aspects to determining how well color and 3D structure are represented together in a 3D point cloud: whether or not a given 3D space was 'sampled' and whether the color of what was sampled in a space was represented correctly. The accuracy with which a 3D point cloud point was assigned the correct color for the corresponding 3D space in the real world (i.e., 'Is the color of a point in the correct place?') depended strongly upon the way

in which SFM sampled the space. Whether or not a space was sampled depended strongly on the feature detector (4.2). In this case, the SIFT feature detector only placed points at locations that were identified as brighter or darker than the surroundings (Figures 7, 8) with the result that sometimes points were not placed at painted targets in the 3D point cloud even when the targets were clearly or partially visible in the images (Figure 3). In examination of the quality of SFM 3D-spectral fusion based on the classification accuracy of painted targets, it was found that fusion accuracy was quite good (93 %, leaf-off) when all targets were 'seen' in the point cloud, but fusion accuracy was much lower (27 %, leaf-on) when targets were not seen as frequently (Table 1). There was however a strong agreement between the color of points and the average color (hue, R, G, B) of point image feature tile (Figure 1e). Values of the percentage of hue difference, as well as RGB channel correlation and ratio all showed strong agreement between point color and average image feature tile at the location where SIFT identified a feature point (means: 5.1 % hue difference, 0.83 RGB channel R^2, 1.06 RGB channel ratio). There is one notable exception where point vs. image feature RGB channel correlation was quite low (mean $R^2 = 0.42$) for the ground leaf-off datasets. Visual inspection of the image feature tiles (Figure 1e) for these points as compared to leaf-on points, which showed correlations comparable to aerial scans (mean $R^2 = 0.82$), reveals that the difference may be due to the way in which the feature detector was able to 'see' painted targets. Leaf-on points at target areas were often observed against a background of foliage while leaf-off points were often observed against a background of sky which was much brighter than the target and which may have caused low correlation values (Appendix A2.9). Combined, these results suggest that the color of points at a given point in 3D space will

be in close agreement (mean 5 - 6 %) to where the same point was observed in images and that in general points will be located in the right place if they were seen by the feature detector.

4.5 SFM and TLS see vegetation differently

TLS and SFM showed very different fusion quality when evaluated based on classification accuracy of painted targets owing to the fact that the two systems 'see' vegetation in different ways. For example, TLS always observed all targets under leaf-on and leaf-off conditions, but classification accuracy tended to be quite low (45 % and 29 %, respectively) as points inside the target areas were just as likely to be red as they were to match the color of points outside the target area (Figure 2). In addition, overall agreement between TLS and SFM was also low (45 % and 36 %) and the negative and small kappa values (-0.14 and 0.06) indicate that agreement between TLS and SFM is not significantly different than agreement by chance alone (Congalton 1991). This difference may be due to the way in which the fusion of 3D color information is carried by each system. In SFM, 3D-fusion is an inherent part of the reconstruction process where the location and color of points comes from the same data source, a sampling of the images by the feature detector. For TLS, 3D comes from the laser scanner and color comes from the camera sensor each with a different optical center and field of view of the scene. This means that the laser points might see something the camera cannot see, for example at a location that is occluded from the point of view of the camera (Popescu and Wynne 2004; Packalén et al. 2009) or locations that the camera only sees as shadow or black pixels but for which the active laser beam is able to 'see' the structure that is obscured from view by the camera. Differences in the way in which SFM and TLS see vegetation were also

apparent when comparing tomographic slices of 3D voxel models of the single tree test-subject under leaf-on and leaf-off conditions. In general, SFM and TLS observed the outer 3D structure of the tree in a similar manner (Figures 4 & 5) and this result is in line with prior work that examined the relationship between the 3D structure of forest canopies observed by SFM from aerial images as compared to airborne LIDAR and field-based measurements (Dandois and Ellis 2010, 2013; Lisein et al. 2013). However, there were also distinct differences in the representation of tree structure by each sensor system. TLS tended to see more of the interior structure of the tree than SFM, especially under leaf-on conditions. This may be due to the fact that the active infrared laser beam is unaffected by shadows or areas of low contrast that caused reduced point sampling for SFM and the high sampling rate of the TLS also resulted in a greater number of points that were not occluded by the outer leaf surfaces (Hosoi et al. 2010). Differences in vertical foliage profiles (VFPs, Figure 5), are not as easily explained. One potential explanation for the large mismatch between TLS and SFM VFPs at the lower and upper parts of the tree crown (Figure 5) may be due to the fact that SFM tended to place more on the outer edges of the tree and the interior, relative to the point of view of the camera (Figures 2 & 6), with the result that the bottom parts of the tree crown were sampled with a lower than average point cloud density.

4.6 The fusion of computer vision and ecological remote sensing

The data collected in this research provides a unique test-set that could be used for future research aimed at developing and evaluating the quality of SFM products of vegetation. Future research could use the image datasets collected here to determine how changing the feature detector or modifying images (e.g., through contrast or brightness

enhancement) will change how vegetation and targets are 'seen' by SFM. In addition, future work should also examine the effects that additional computer vision processing will have on data processing. As implemented in this study, Bundler produces only a 'sparse' point cloud dataset consisting only of those points that were identified by SIFT, matched across images, and contributed to the bundle adjustment optimization (Snavely et al. 2006). Snavely et al. (2006) suggest that 'denser' point clouds can be obtained with additional levels of 'dense-matching' processing that require information about camera internal calibration and external location and orientation that are obtained from the sparse bundle adjustment stage. Examples of such algorithms include the Patch-based or Cluster-based Multi-View Stereo algorithms (PMVS, CMVS; Furukawa et al. 2010; Furukawa and Ponce 2010). Denser point clouds obtained from such processing may contribute to an improved sampling of 3D space and observation of vegetation and painted targets, but the quality of fusion measurements from this new processing still needs to be evaluated and the test datasets developed in this study could be used for such development.

The graphical user interface and test datasets may also be useful for exploring a whole new area of computer vision for ecological remote sensing: computer vision feature based remote sensing. Future research should evaluate the potential of a 'bag of features' approach for classifying SFM point clouds to landcover or tree functional type through pre-segmentation of groups of features into 2D pixels or 3D volume-pixels (voxels), or even to individual tree crown species through tree crown delineation based on point cloud 3D structural and spectral traits (Alexander 2009; Ke and Quackenbush 2011). The datasets and graphical user interface developed in this study can be used to

help calibrate or evaluate such feature descriptor-based classifications. There are also many feature detectors available for identifying potential points for matching across images. From the point of view of computer vision, they are often grouped based on the type of shape or object that they detect and represent in the image, for example points, patches or blobs, lines, and edges (Schmid et al. 2000; Mikolajczyk and Schmid 2004; Szeliski 2011). Future work should also consider the potential of dedicated vegetation feature detectors based around properties of plant, leaf, and flower color and structure (Nilsback 2009; Kendal et al. 2013). Along with providing the necessary match information needed for SFM 3D reconstructions, such vegetation-specific detectors may also be useful for automatically mapping and quantifying the amount of canopy fruits and flowers, which is a significant factor for understanding the relationship between forest plant and animal communities (Wright et al. 1999), but which is challenging to measure in practice by manual methods alone due to difficulties with accessing the forest canopy.

There is also potential to explore the relationship between landscape pattern and the scale-space representation of the landscape that is made possible by the SIFT feature descriptor. The application of scale-space theory to the study of aerial remote sensing images of trees and forests may be particularly useful for improving understanding of the multi-scale patterns of these types of landcovers within a landscape (Hay et al. 2002). While the current research was unable to show a strong relationship between the scale of features within k-means clusters of SIFT 128D numeric descriptors, future research should consider the potential hierarchical nature of SIFT feature scale as it pertains to components of the forest canopy. For, example, from a physiological point of view a forest canopy is composed of all the leaves, branches, twigs and crowns of the forest trees

(Parker 1995) and although k-means clustering of all features based on the 128D descriptor alone did not reveal multi-scale patterns, perhaps grouping features into scale-specific 'bags' first would reveal patterns of vegetation unique to different scales. Such scale-based features bags may reveal 'optimal' scales for observing different aspects or patterns of vegetation (Hay et al. 2002).

5.0 Conclusions

The results from this analysis reveal that the properties of a computer vision structure from motion point cloud are constrained at the moment of observation by the feature detector used for matching and bundle adjustment and in this study, results were dependent upon the SIFT feature detector. While it may be possible for a human operator to indicate a semantic meaning or classification of the context within which a point is located, e.g., forest, pavement, tree, bush, roof, etc., ultimately the feature detector is merely responding to a very specific type of signal within the image domain, bright spots surrounded by dark or dark spots surrounded by bright, which are identified by the algorithm independent of any semantic or ecological meaning. It follows then that the results of the current work may be different if a different feature detector is used in the place of SIFT. Indeed, future work should consider conducting structure from motion using feature detectors specifically designed for identifying and tracking ecological entities. For example, detectors could be developed to identify individual leaves, flowers, or fruits (e.g., cones, acorns, samara). In addition, future work should also consider a 'bag-of-features' approach to improve classification results of feature descriptors by first grouping features based on pre-defined spatial boundaries. For example, to what extent can individual tree species by classified based on groups of feature descriptors within

delineated tree crowns? The fusion of computer vision with ecological research represents an exciting new frontier for remote sensing of the structural, spectral, and taxonomic complexity of terrestrial ecosystems.

References

Alexander, C. (2009). Delineating tree crowns from airborne laser scanning point cloud data using Delaunay triangulation. International Journal of Remote Sensing, 30, 3843-3848.

Anderson, J., Plourde, L., Martin, M., Braswell, B., Smith, M., Dubayah, R., Hofton, M., & Blair, J. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. Remote Sensing of Environment, 112, 1856-1870.

Anderson, K., & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. Frontiers in Ecology and the Environment, 11, 138-146.

Aptoula, E., & Lefèvre, S. (2009). Morphological description of color images for content-based image retrieval. Image Processing, IEEE Transactions on, 18, 2505-2517.

Asner, G.P., & Martin, R.E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Frontiers in Ecology and the Environment, 7, 269-276.

Baldeck, C., & Asner, G. (2013). Estimating Vegetation Beta Diversity from Airborne Imaging Spectroscopy and Unsupervised Clustering. Remote Sensing, 5, 2057-2071.

Baldeck, C., & Asner, G. (2014). Improving Remote Species Identification through Efficient Training Data Collection. Remote Sensing, 6, 2682-2698.

Besl, P.J., & McKay, H.D. (1992). A method for registration of 3-D shapes. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 14, 239-256.

Biggar, S.F., Thome, K.J., & Wisniewski, W. (2003). Vicarious radiometric calibration of EO-1 sensors by reference to high-reflectance ground targets. Geoscience and Remote Sensing, IEEE Transactions on, 41, 1174-1179.

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, 35-46.

Dandois, J.P., & Ellis, E.C. (2010). Remote sensing of vegetation structure using computer vision. Remote Sensing, 2, 1157-1176.

Dandois, J.P., & Ellis, E.C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.

Dassot, M., Constant, T., & Fournier, M. (2011). The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Annals of Forest Science, 68, 959-974.

Erdody, T.L., & Moskal, L.M. (2010). Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sensing of Environment, 114, 725-737.

Furukawa, Y., Curless, B., Seitz, S.M., & Szeliski, R. (2010). Towards Internet-scale multi-view stereo. In, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 1434-1441).

Geerling, G., Labrador-Garcia, M., Clevers, J., Ragas, A., & Smits, A. (2007). Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. International Journal of Remote Sensing, 28, 4263-4284.

Glennie, C. (2007). Rigorous 3D error analysis of kinematic scanning LIDAR systems. Journal of Applied Geodesy, 1, 147-157.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sensing, 4, 1573-1599.

Hay, G., DubÈ, P., Bouchard, A., & Marceau, D. (2002). A scale-space primer for exploring and quantifying complex landscapes. Ecological Modelling, 153, 27-49.

Henning, J.G., & Radtke, P.J. (2006). Ground-based Laser Imaging for Assessing Three Dimensional Forest Canopy Structure. Photogrammetric Engineering and Remote Sensing, 72, 1349.

Holden, M., Hill, D.L.G., Denton, E.R.E., Jarosz, J.M., Cox, T.C.S., Rohlfing, T., Goodey, J., & Hawkes, D.J. (2000). Voxel similarity measures for 3-D serial MR brain image registration. Medical Imaging, IEEE Transactions on, 19, 94-102.

Hosoi, F., Nakai, Y., & Omasa, K. (2010). Estimation and Error Analysis of Woody Canopy Leaf Area Density Profiles Using 3-D Airborne and Ground-Based Scanning Lidar Remote-Sensing Techniques. Geoscience and Remote Sensing, IEEE Transactions on, 48, 2215-2223.

Huang, C., Thomas, N., Goward, S.N., Masek, J.G., Zhu, Z., Townshend, J.R.G., & Vogelmann, J.E. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. International Journal of Remote Sensing, 31, 5449-5464.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sensing of Environment, 82, 397-416.

Jain, A.K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31, 651-666.

Javernick, L., Brasington, J., & Caruso, B. (2014). Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. Geomorphology, 213, 166-182.

Kampe, T.U., Johnson, B.R., Kuester, M., & Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. Journal of Applied Remote Sensing, 4, 043510.

Ke, Y., & Quackenbush, L.J. (2011). A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. International Journal of Remote Sensing, 32, 4725-4747.

Ke, Y., Quackenbush, L.J., & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. Remote Sensing of Environment, 114, 1141-1154.

Kendal, D., Hauser, C.E., Garrard, G.E., Jellinek, S., Giljohann, K.M., & Moore, J.L. (2013). Quantifying Plant Colour and Colour Difference as Perceived by Humans Using Digital Images. PLoS ONE, 8, e72296.

Lange, T., Roth, V., Braun, M.L., & Buhmann, J.M. (2004). Stability-based validation of clustering solutions. Neural Computation, 16, 1299-1323.

Lazebnik, S., Schmid, C., & Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In, CVPR (pp. 2169-2178): IEEE.

Lee, H.Y., Lee, H.K., & Ha, Y.H. (2003). Spatial color descriptor for image retrieval and video segmentation. Multimedia, IEEE Transactions on, 5, 358-367.

Lindeberg, T. (1994). Scale-space theory: A basic tool for analysing structures at different scales,". Journal ofapplied statistics, 21, 225-270.

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., & Lejeune, P. (2013). A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. Forests, 4, 922-944.

Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60, 91-110.

Manjunath, B.S., Ohm, J.R., Vasudevan, V.V., & Yamada, A. (2001). Color and texture descriptors. Circuits and Systems for Video Technology, IEEE Transactions on, 11, 703-715.

McKean, J., Isaak, D., & Wright, W. (2009). Improving Stream Studies With a Small-Footprint Green Lidar. Eos, Transactions American Geophysical Union, 90, 341-342.

Mikolajczyk, K., & Schmid, C. (2004). Scale & affine invariant interest point detectors. International Journal of Computer Vision, 60, 63-86.

Morgenroth, J., & Gomez, C. (2014). Assessment of tree structure using a 3D image analysis technique—A proof of concept. Urban Forestry & Urban Greening, 13, 198-203.

Mundt, J.T., Streutker, D.R., & Glenn, N.F. (2006). Mapping sagebrush distribution using fusion of hyperspectral and lidar classifications. Photogrammetric Engineering and Remote Sensing, 72, 47.

Newsam, S., & Yang, Y. (2007). Comparing global and interest point descriptors for similarity retrieval in remote sensed imagery. In, Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems. Seattle, Washington: ACM.

Nilsback, M.-E. (2009). An automatic visual Flora - segmentation and classication of flower images. In, Robotics Research Group, Department of Engineering Science: University of Oxford.

Packalén, P., Suvanto, A., & Maltamo, M. (2009). A two stage method to estimate speciesspecific growing stock. Photogrammetric Engineering and Remote Sensing, 75, 1451-1460.

Parker, G., Harding, D., & Berger, M. (2004). A portable LIDAR system for rapid determination of forest canopy structure. Journal of Applied Ecology, 41, 755-767.

Parker, G.G. (1995). Structure and microclimate of forest canopies. In M. Lowman, & N. Nadkarni (Eds.), Forest canopies: A Review of Research on a Biological Frontier (pp. 73-106). San Diego: Academic Press.

Parker, G.G., & Tibbs, D.J. (2004). Structural phenology of the leaf community in the canopy of a Liriodendron tulipifera L. forest in Maryland, USA. Forest Science, 50, 387-397.

Popescu, S., & Wynne, R. (2004). Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering and Remote Sensing, 70, 589-604.

Schimel, D., Keller, M., Berukoff, S., Kao, B., Loescher, H., Powell, H., Kampe, T., Moore, D., & Gram, W. (2011). NEON Science Infrastructure. In, NEON 2011 Science Strategy (pp. 36-38).

Schmid, C., Mohr, R., & Bauckhage, C. (2000). Evaluation of Interest Point Detectors. International Journal of Computer Vision, 37, 151-172.

Seielstad, C., Stonesifer, C., Rowell, E., & Queen, L. (2011). Deriving Fuel Mass by Size Class in Douglas-fir (Pseudotsuga menziesii) Using Terrestrial Laser Scanning. Remote Sensing, 3, 1691-1709.

Snavely, N., Seitz, S., & Szeliski, R. (2006). Photo tourism: exploring photo collections in 3D. In (pp. 835-846): ACM New York, NY, USA.

Snavely, N., Simon, I., Goesele, M., Szeliski, R., & Seitz, S.M. (2010). Scene Reconstruction and Visualization from Community Photo Collections. Proceedings of the IEEE, 98, 1370-1390.

Szeliski, R. (2011). Computer Vision. Springer Algorithms and Applications.

Teillet, P., Fedosejevs, G., Gauthier, R., O'Neill, N., Thome, K.J., Biggar, S.F., Ripley, H., & Meygret, A. (2001). A generalized approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data. Remote Sensing of Environment, 77, 304-327.

Tooke, T., Coops, N., Goodwin, N., & Voogt, J. (2009). Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. Remote Sensing of Environment, 113, 398-407.

Van de Sande, K.E.A., Gevers, T., & Snoek, C.G.M. (2010). Evaluating Color Descriptors for Object and Scene Recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32*, 1582-1596.

Vitousek, P., Asner, G.P., Chadwick, O.A., & Hotchkiss, S. (2009). Landscape-level variation in forest structure and biogeochemistry across a substrate age gradient in Hawaii. Ecology, 90, 3074-3086.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., & Reynolds, J.M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314.

Wright, S.J., Carrasco, C., Calderón, O., & Paton, S. (1999). The El Niño southern oscillation, variable fruit production, and famine in a tropical forest. Ecology, 80, 1632-1647.

Yang, Y., & Newsam, S. (2008). Comparing SIFT descriptors and gabor texture features for classification of remote sensed imagery. In, Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on (pp. 1852-1855).

Yang, Y., & Newsam, S. (2013). Geographic Image Retrieval Using Local Invariant Features. Geoscience and Remote Sensing, IEEE Transactions on, 51, 818-832.

Zhang, J., Marszalek, M., Lazebnik, S., & Schmid, C. (2006). Local features and kernels for classification of texture and object categories: A comprehensive study. In (p. 13): IEEE.

Zhou, W., Huang, G., Troy, A., & Cadenasso, M.L. (2009). Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. Remote Sensing of Environment, 113, 1769-1777.

Chapter 4: Optimal conditions for aerial measurements of forest canopy structure using computer vision

Abstract

Three-dimensional (3D), multi-spectral measurement of forest canopies by consumer unmanned aerial systems (UAS) and computer vision structure from motion (SFM) algorithms is transforming remote sensing for ecological research. UAS-SFM remote sensing systems (e.g. Ecosynth) enable LIDAR-like, high spatial resolution measurements of vegetation structural and spectral traits on-demand. Yet application of this emerging technology has out-paced understanding of the relationship between data quality and the conditions of data collection. In this research, Ecosynth UAS-SFM remote sensing (Dandois and Ellis 2013) was used to produce 3D multispectral point clouds of Temperate Deciduous forest canopies at multiple levels of UAS flight altitude and photographic overlap, scene weather conditions, and with multiple image processing methods. Variation in measures of point cloud quality traits and metrics was evaluated within and across data collection configuration levels. Ecosynth estimates of canopy height were robust to large variation in data collection conditions and were highly correlated with field and LIDAR estimates of canopy height (R^2 = 0.87 & 0.99). Under optimal scanning conditions (clear lighting, high altitude, and high photographic side overlap) Ecosynth produced accurate estimates of canopy height compared to field measurements (RMSE = 4.5 m) and LIDAR (RMSE = 2.4 m). Optimal Ecosynth point clouds had horizontal and vertical geometric accuracy < 1.7 m RMSE, an average point cloud density of 35 points m^{-2} in forested areas, and an average canopy penetration of roughly 20 % or 5 m. Ecosynth point cloud quality was generally improved by increasing photographic overlap, yet this also resulted in significant increases in SFM

computation time. Ecosynth canopy height error was higher on cloudy days when image contrast was low compared to clear days when image contrast was relatively higher. Differences between quality and overlap, lighting, and image contrast may be related to the behavior of computer vision 'image features' within the SFM processing pipeline, but access to such features was restricted due to the use of a commercial, closed-source SFM algorithm. Recommendations are made for optimal UAS-SFM remote sensing data collection strategies. Future research should use free, open-source SFM algorithms to access computer vision image features to improve understanding of the relationship between point cloud quality and scanning conditions.

1.0 Introduction

The combination of hobbyist unmanned aerial systems (UASs) combined with automated computer vision structure from motion (SFM) 3D image processing represents a transformative shift in remote sensing capabilities for the evaluation of forest canopies. Such 'personal remote sensing systems' (e.g., Ecosynth) enable on-demand observations of vegetation structural and spectral traits at high spatial resolution over small extents and frequencies not possible or practical with existing forms of airborne or satellite remote sensing (Chapter 2, Dandois and Ellis 2010, 2013). Personal remote sensing systems have enabled accurate mapping of canopy height (Tao et al. 2011, Lisein et al. 2013) and biomass density as well as the discrimination of structural and color-spectral phenology traits at the scale of individual trees (Dandois and Ellis 2010, 2013). Similar systems have also been used for mapping stream channel geomorphology (Castillo et al. 2012, Javernick et al. 2014), mapping vineyard and orchard plant structure (Dey et al. 2012, Mathews and Jensen 2013), the topography of bare substrates (de Matías et al. 2009,

Harwin and Lucieer 2012, James et al. 2012, James and Robson 2012, Rosnell and Honkavaara 2012, Westoby et al. 2012), and mapping lichen and moss extent and coverage (Turner et al. 2012). However, rapid application of UAS and SFM for research has out-paced understanding of the quality of the products derived from these technologies. While automated hobbyist UAS reduces the barrier to entry for remote sensing data collection (Anderson and Gaston 2013), little is known about the relationship between observations of vegetation structural and spectral metrics and the conditions under which observations are obtained. This research aims to address these uncertainties by defining relevant measures of SFM point cloud quality and then characterizing how these measures change under different observation conditions.

1.1 UAS-SFM remote sensing

Use of SFM and UAS for remote sensing has increased rapidly thanks to the relative ease with which these technologies can be deployed for research applications. Recent research spans a range of applications, algorithms, data collection methods (including UAS, manned aircraft, and ground-based strategies), and flight configurations, summarized in Table 1 of Chapter 1. SFM research that used UAS was conducted at a range of flight altitudes (30 m – 225 m above ground level, AGL) and parameters of photographic overlap (from 40 % to > 90 % forward and side overlap). These studies also used different SFM applications (both free, open-source and commercial, closed-source), as well as different data collection platforms, including fixed-wing and multirotor, autonomous and manually operated UAS, kites, full-scale manned aircraft, and even photos collected by individuals walking on the ground around scenes. While these studies arrive at the similar conclusion that accurate 3D reconstructions of

landscapes (including vegetation and topography) can be produced with SFM remote sensing, the diversity of methods with which the research was carried out highlights a significant challenge and research opportunity for this burgeoning remote sensing field.

Many of these challenges were also presented at the end of Chapter 2 (Section 4.5.6, Table 7; Dandois and Ellis 2013). Despite the tremendous potential of new areas of research made possible by Ecosynth UAS-SFM remote sensing, that research highlight many uncertainties about the way in which the data could be collected. It is not clear how changes in UAS flying altitude, photographic overlap, and resolution as well as wind, cloud cover, and light will influence the quality of SFM point clouds or even what the relevant measures of quality might be for such a system. It is also unclear how such changes in SFM point cloud quality will influence vegetation measurements, in particular metrics of canopy structure like height and biomass. Næsset (2009a) highlighted similar challenges in the use of LIDAR for remote sensing of forest canopies, noting the large diversity in data collection strategies, including differences in sensor, flying altitude, and pulse repetition frequency. They argued that lack of understanding of how such differences would influence canopy metrics could potentially limit future applications, in particular when multiple datasets are combined to assess change. They demonstrated significant differences in LIDAR point cloud traits as a function of scan configuration, in particular point cloud density, but noted that canopy structure measurements were relatively robust to a range of collection strategies. Current research on the relationship between attributes of LIDAR data quality and derived products (e.g., Cook et al. 2009; Næsset 2009b; Næsset 2009a; Hudak et al. 2012) are made possible by over 25 years of LIDAR remote sensing of vegetation, from the earliest research on using LIDAR for

estimating forest biomass (Nelson et al. 1988) that seems extremely simple compared to modern remote sensing capable of accurately mapping canopy biomass at large scales (Asner and Mascaro 2014). Similar bounding understanding of the quality of UAS-SFM canopy products is required to advance this growing remote sensing field from a novel concept to a practical research tool.

1.3 Research objectives

The objectives of this research are to improve understanding of the quality of UAS-SFM forest canopy products and the relationship between quality and the conditions under which data are collected using empirical and applied methodological approaches (Chapter 1, Section 1.1.6). Based on relevant measures used in LIDAR and optical image remote sensing, three categories of Ecosynth point cloud quality measurements are defined that describe point cloud geometric positioning accuracy, canopy sampling as a function of point cloud density and canopy penetration, estimates of canopy structure, and point cloud color radiometric quality. Empirical traits and forest canopy metrics of Ecosynth SFM 3D point cloud models are evaluated across a range of different data collection conditions. A replicated set of Ecosynth UAS-based image acquisitions were carried out under crossed treatments of lighting, flight altitude, and image overlap. Variation in traits and metrics were compared within and across treatment levels as well as to wind speed measured from a nearby weather station and to post-processing based variations in image datasets. Forest canopy metrics derived from Ecosynth products were compared to field based measurements of canopy height and also to a high-resolution ($\approx$ 10 points m^{-2}) discrete-return LIDAR point cloud collected during the same period as UAS collections. Radiometric quality of Ecosynth 3D-RGB (red-green-blue) point cloud color values were

evaluated by measurement of variation or contrast within different land covers, including forest, pavement, grass, and rooftops.

2.0 Materials and methods

2.1 Data collection

2.1.1 Study area and field data

Research was carried out at the same three 6.25 ha (250 m x 250 m) Temperate Deciduous forest research study sites in Maryland USA used in prior Ecosynth UAS SFM remote sensing research (Chapter 2, Dandois and Ellis 2013). Two sites ('Herbert Run' and 'Knoll') are located on the campus of the University of Maryland Baltimore County (UMBC: 39°15′18″N 76°42′32″W) and one at the Smithsonian Environmental Research Center in Edgewater Maryland (SERC: 38°53'10"N 76°33'51"W). In the current study, the majority of UAS flights were carried out at the Herbert Run study site on the UMBC campus. Each site was divided into 25 m x 25 m plots and the average height of the five largest (by DBH) trees plot was used to estimate the average maximum canopy height (Dandois and Ellis 2013). Above ground biomass density (Mg ha^{-1}) was estimated for each plot following Dandois and Ellis (2013) by allometric modeling of the DBH of all stems greater than 1 cm DBH within each plot (Jenkins et al. 2003). Average field estimated AGB across all plots (n= 49) at Herbert Run was 204 Mg ha^{-1} with a standard deviation of 156 Mg ha^{-1}.

2.1.2 UAS image acquisition under a controlled experimental design

Image data were collected with a Canon ELPH 520 HS digital camera attached to a hobbyist, commercial multirotor UAS consisting of Mikropkopter frames (HiSystems

GmbH, http://www.mikrokopter.de) and Arducopter flight electronics (3D Robotics, Inc., http://copter.ardupilot.com) in 'octo' and 'hexa' configurations of 8 and 6 propellers (http://wiki.ecosynth.org) and following Dandois and Ellis (2013). Images were collected at roughly 2 frames per second (fps) at 10 megapixels (MP) resolution and "Infinity" focus (≈ 4 mm focal length) using fixed white balance and exposure settings calibrated to an 18 % gray card with an unobstructed view of the sky. Image datasets were collected following a crossed factorial experimental design based on combinations of light, altitude above the canopy, and photographic side overlap, described in Table 1a. Two levels of lighting, uniformly clear and uniformly cloudy (diffuse) were controlled for by selecting specific days for flight based on desired lighting conditions. Four levels of flight altitude above the canopy (20 m, 40 m, 60 m, 80 m) and four levels of photographic side overlap (0 %, 25 %, 50 %, 75 %) were controlled for by pre-programming of automated UAS waypoint flight paths based on a designated flight altitude above the launch location and by the spacing between parallel flight tracks, respectively. Flight configurations that include each of the three variables are outlined in Table 1b.

UAS were programmed to fly at an average of 6 m s^{-1} (13.4 mph). Due to maximum flight time limits the hexacopter configuration was used on flights less than 15 minutes in flight time or 4 km in flight distance and the octocopter configuration was used for longer flights, < 25 minutes flight time and < 8 km flight distance. Five replicates were planned for each treatment which were flown from 2013-06-21 to 2013-10-21. Automated waypoint control, take-off, and landing modes were used to facilitate rapid collection of replicate and in general replicates were collected between 09:00 -

16:00 each day. All treatments with five replicates were flown at the Herbert Run site on the campus of UMBC (Dandois and Ellis 2013), and several treatments were flown with replication at the Knoll and SERC sites. Due to rapidly changing lighting conditions on some days or human error, five flight configurations were collected with between 2 - 4 replicates, instead of the planned 5 replicates. Several flight configurations at low altitudes and low levels of side overlap were not collected due to onset of autumn leaf senescence as the rapidly changing condition of the forest canopy resulted in a loss of experimental control on the canopy itself, as a result, a total of 82 replicates were collected. Based on flight design, all UAS flights conducted in this research were carried out below 120 m (≈ 400 feet) above the surface. Wind speed during UAS flights over the Herbert Run area was measured by eddy covariance instrumentation at approximately 90 m MSL on top of a building on the UMBC campus. Wind speeds were extracted from eddy covariance data logs by matching time stamps with UAS GPS telemetry time stamps for each flight. Average wind speeds during flights ranged from 0.6 – 5.9 m s^{-1} (1.3 – 13 mph) and were converted to values of Beaufort scale for comparison across datasets (http://www.spc.noaa.gov/faq/tornado/beaufort.html, accessed: 2014-05-20).

2.1.3 Aerial LIDAR

Discrete-return LIDAR data were acquired over all three sites (Knoll, Herbert Run, and SERC) on 2013-10-25 by the contractor Watershed Sciences, Inc. LIDAR was collected with a nominal point density of 10.06 points m^{-2} (0.028 m and 0.017 m horizontal and vertical accuracy). While some seasonal change was already underway at the time of acquisition, the data represents the only comparable leaf-on LIDAR for the UMBC campus that can be used for analysis of Ecosynth point cloud canopy quality measures

(Dandois and Ellis 2013). The contractor-provided 1 m x 1 m gridded bare earth filtered product was used as a digital terrain model (DTM) for extracting heights from LIDAR and Ecosynth point clouds. The points corresponding only to the LIDAR first return were extracted for each 6.25 ha study site to serve as a 'gold-standard' measure of the overall surface structure (Habib et al. 2009).

Table 1: Experimental design of Ecosynth flights (a). Boxes with "X" indicate treatments that were successfully flown, gray shaded boxes indicate treatments that were not covered by data collection. (b) shows configurations of individual flights.

(a)

	L: CLEAR	L: CLOUDY	A: 20 m	A: 40 m	A: 60 m	A: 80 m	S: 0 %	S: 25 %	S: 50 %	S: 75 %
L: CLEAR										
L: CLOUDY										
A: 20 m	X	X								
A: 40 m	X	X								
A: 60 m	X	X								
A: 80 m	X	X								
S: 0 %	X	X				X				
S: 25 %	X	X				X				
S: 50 %	X	X		X	X	X				
S: 75 %	X	X	X	X	X	X				

L = Lighting: uniformly CLEAR or CLOUDY conditions
A = Nominal flight altitude above highest canopy point
S = Photographic side overlap (sidelap)
N = Number of replicates actually collected

(b) Flight configurations

L	A	S	N
Clear	20 m	75 %	4
Clear	40 m	50 %	5
Clear	40 m	75 %	2
Clear	60 m	50 %	5
Clear	60 m	75 %	5
Clear	80 m	0 %	5
Clear	80 m	25 %	5
Clear	80 m	50 %	5
Clear	80 m	75 %	7
Cloudy	20 m	75 %	5
Cloudy	40 m	50 %	5
Cloudy	40 m	75 %	3
Cloudy	60 m	50 %	5
Cloudy	60 m	75 %	2
Cloudy	80 m	0 %	5
Cloudy	80 m	25 %	5
Cloudy	80 m	50 %	4
Cloudy	80 m	75 %	5

2.2 Data processing

All Ecosynth image replicates were processed into 3D-RGB point cloud datasets using the commercial and closed-source Agisoft Photoscan SFM software package (64-bit, v0.91, build 1703) following Dandois and Ellis (2013). Point clouds were produced using the 'Align Photos' tool with settings 'high' accuracy, 'generic' pair pre-selection, and the default of maximum 40,000 point features per image (refer to Section 1.1.5, Chapter 1 for a review of image features). Point clouds produced in Photoscan were then processed following Ecosynth data processing procedures (Dandois and Ellis 2013, http://code.ecosynth.org/EcosynthAerial), which included filtering to remove stray noise points and georeferencing into the WGS84 UTM Zone 18N projected coordinate system by optimized 'spline' fitting of the SFM camera point 'path' to the UAS GPS telemetry path. This approach required an estimate of the horizontal location and elevation of the camera during flight. Horizontal location was obtained from latitude and longitude values reported from the UAS GPS telemetry which were transformed to the UTM projected coordinate system. Elevation was available in two forms: a relative elevation above the launch location from on-board pressure sensor and an absolute elevation by a built-in combination of GPS-based elevation and pressure sensor elevation. Manufacturer reported specifications of the pressure sensor estimate precision of +/- 1.7 m within the nominal operating ranges found during data collection for this study (+/- 0.2 millibar at 20° C and 1000 millibar nominal pressure). UAS GPS was enabled to use the WAAS (Wide Area Augmentation System) when available, which has a 95% horizontal and vertical precision limit of 4 m (FAA 2008). All relative elevation values were aligned to the local vertical datum by addition of the elevation of the launch location in

meters above mean sea level (MSL) as reported by the LIDAR DTM. Comparison of relative and absolute elevation values to copter elevations recorded with a Trupulse 360B handheld laser rangefinder (+/- 1 m precision to low quality targets) revealed significantly greater error and variation from absolute height measurements than from relative height measurements (RMSE = 10.5 m vs. 6.6 m, $p < 0.0001$). A calibration model was then developed to estimate the true elevation of the copter based on the relative elevation values recorded in the UAS telemetry ($R^2 = 0.999$, 1.2 m RMSE residual error, Appendix A3.1). All Ecosynth point clouds were georeferenced using calibrated relative UAS elevation values.

In addition to the experimental treatments tested through field data collection (Table 2), several additional treatments representing different forms of post-processing were tested by manipulating the original replicate images. A single set of five image replicates collected under the same conditions (clear lighting, 80 m altitude, 75 % side overlap, collected 2013-08-26) were also processed under different variations of image processing and computation to evaluate the effects of these variables on point cloud traits and metrics. These replicates were processed into 3D-RGB point clouds after images had been down-sampled from the original 10 MP to reduced resolution (7.5 MP, 5 MP, 2.5 MP, 1 MP, 0.5 MP, 0.3 MP). The same replicates, at original 10 MP resolution, were also processed into point clouds after incrementally sub-sampling the image sequences from every single image to every 10th image, corresponding to a decrease in forward photographic overlap from ≈ 96 % to 60 %. To facilitate the large amount of computation required to process the 82 replicates for the main set of treatments (Table 2), plus additional replicates of reduced resolution (n = 35) and forward overlap (n = 50),

SFM reconstructions were run on multiple computers with different configurations of OS, RAM, and CPU resources. A set of the same replicates were also run on each computer to evaluate what if any effect that variable would have on point cloud traits and metrics. In total over 4,500 compute hours (> 188 compute days) were required to process all of the image datasets used in this study. A single replicate was also run in two other versions of Photoscan, v0.84 which was used in prior research (Dandois and Ellis 2013) and version 1.04 the latest version (2014-05-25), as well as an enhanced version of the popular free and open source Bundler-PMVS pipeline (Table 1) designed to take advantage of GPU computing and UAS GPS telemetry information named 'Ecosynther' (v0.8, http://code.ecosynth.org/Ecosynther, accessed: 2014-05-26).

2.3 Data analysis

Ecosynth data quality was measured based on three main categories of empirical traits and applied forest canopy metrics that were extracted from georeferenced Ecosynth point clouds: positioning accuracy, canopy sampling, and canopy structure. Positioning accuracy traits were divided into measures of the horizontal and vertical positioning error (meters root mean square error, RMSE) and calculated as 'relative' and 'absolute' positioning error. Canopy sampling traits described point cloud density and canopy penetration. Metrics of canopy structure describe error in Ecosynth estimates of canopy height and aboveground biomass density relative to LIDAR and field measurements. The radiometric quality of Ecosynth point clouds was also measured based on the amount of variation in red, green, and blue channel brightness and grayscale intensity within different landcovers.

2.3.1 Measurements of position accuracy

Relative positioning accuracy was defined based on the goodness of fit of the SFM 'camera' path to the UAS GPS telemetry path. This was measured as meters RMSE of the XY and Z distance between a camera point and the closest segment of two GPS points along the UAS GPS track. Because SFM camera and scene points are located within the same coordinate system, this 'Path error' value measures how close the entire SFM point cloud was aligned to the UAS-GPS positioning data during the georeferencing step. Following Habib et al. (2009) the absolute positioning error of Ecosynth SFM point cloud datasets was evaluated automatically by computing the horizontal and vertical displacement required to rigidly align the SFM point cloud to the LIDAR first return point cloud based on 3D translation and rotation only and without scaling. LIDAR horizontal and vertical positioning accuracy was estimated to be less than 3 cm by the manufacturer and served as a 'gold-standard' of the 3D geometry of the scene for comparison against Ecosynth measurements. A Python implementation of the Iterative Closest Point algorithm (ICP; Besl and McKay 1992) was used for the rigid alignment and measures of 'ICP error' refer to the horizontal and vertical RMSE in meters of the location difference between same SFM point cloud points before and after alignment. ICP fitting to LIDAR was only used to measure absolute positioning accuracy of point clouds and no other metrics were computed based on these 'fitted' point clouds. To provide an absolute measure of positioning accuracy that does not require a coincident high-resolution LIDAR acquisition, the average difference between the known launch location altitude in meters above sea level from the DTM and the estimated elevation of the same location in each Ecosynth point cloud was measured. While this measure was

based on a coincident LIDAR acquisition, such a measure can be produced relative to any DTM that would be used for estimating canopy heights from Ecosynth point clouds, including when it is only possible to obtain old LIDAR DTMs that no longer accurately portray canopy height (Dandois and Ellis 2010, 2013) or potentially even satellite remote sensing based DTMs (Ni et al. 2014).

2.3.2 Measurements of canopy structure

To evaluate the accuracy with which Ecosynth SFM point clouds characterized canopy structure, gridded canopy height models (CHMs) at 1 m x 1 m pixel resolution were produced for all SFM point cloud datasets and for the first-return of LIDAR points based on the highest point elevation within each grid cell after subtraction of LIDAR digital terrain model (DTM) values from each point cloud elevation (Dandois and Ellis 2013). Within each 25 m x 25 m field plot, the top-of-canopy height (TCH; Asner and Mascaro 2014; Mascaro et al. 2014) was calculated from CHMs based on the average of all pixel values within the plot. While it is common in applications of SFM and LIDAR point clouds to estimate canopy height based on the 'best' of many, often correlated, height metrics (e.g., mean, median, minimum, maximum, every 10th percentile; Næsset 2009; Dandois and Ellis 2010, 2013; Lisein et al. 2013), TCH was selected as a measure that is more comparable across different remote sensing representations of the canopy surface and one that is less susceptible to differences in canopy penetration depth (Asner and Mascaro 2014; Mascaro et al. 2014). Measures of data quality from canopy height are defined as the RMSE error between Ecosynth, field, and LIDAR measurements at the scale of 25 m x 25 m (0.06 hectare) field plots. Aboveground biomass density (Mg ha^{-1})

was estimated for each plot at the Herbert Run site using Ecosynth TCH estimates following Dandois and Ellis (2013).

2.3.3 Measures of canopy sampling

Two additional measures of data quality were defined that characterize the way in which the canopy is sampled or 'seen' by Ecosynth SFM point clouds: point cloud density (points m^{-2}) and canopy penetration (the coefficient of variation of point cloud heights), both of which are calculated first within a raster grid of 1 m x 1 m cells and then averaged across forested areas. Previous research has suggested the strong relationship between LIDAR canopy height metrics and measures of point cloud density (Næsset 2009). Measures of canopy penetration are also related to canopy height metrics (Chasmer et al. 2006a; Chasmer et al. 2006b; Wasser et al. 2013), as well as canopy habitat (Goetz et al. 2007; Jung et al. 2012).

2.3.4 Radiometric quality of Ecosynth point clouds

The radiometric quality of Ecosynth point clouds was evaluated by quantifying the variability (standard deviation) of the color of points inside 1 m x 1 m bins, which was then averaged by landcover. Radiometric quality was measured on a per channel basis for the red, green, and blue (RGB) color of points as well as point grayscale intensity (Gray = 0.299×R + 0.587×G + 0.114×B; Yang and Newsam 2013). This measure of radiometric quality provides an indication of the amount of noise in Ecosynth point colors within a fixed area (Yang and Lo 2000; Baltsavias et al. 2001; Poli et al. 2014), but can also be interpreted as a measure of color contrast (Peli 1990; Cox and Booth 2009), which has a strong influence on the stability of computer vision image features used in SFM reconstruction (Lowe 2004).

3.0 Results

3.1 Ecosynth SFM estimates of canopy structure

Ecosynth SFM estimates of canopy height based on the top-of-canopy height (TCH) measure were robust to a large range of scanning conditions. In general, Ecosynth TCH was highly correlated with the average maximum height per plot ($R^2 = 0.87$) but with relatively high average root mean square error (RMSE range = 3.3 – 10.5 m, average 6.5 m). By comparison, LIDAR TCH was similarly correlated with field measured plot heights ($R^2 = 0.88$) but with slightly lower error (RMSE = 5.5 m. Figure 1c). Both Ecosynth and LIDAR TCH under-predicted field measured canopy height but were highly correlated with each other ($R^2 = 0.99$, RMSE range = 0.8 – 6.3 m, average 2.4 m). Under optimal scanning conditions (clear skies, 80 m altitude above the canopy, 75 % side photographic overlap), Ecosynth TCH estimated field measured canopy height with 4.5 m and 4.6 m RMSE error at the Herbert Run and SERC sites (Table 2, Figure 1a). At all sites, Ecosynth TCH estimates had relatively low error compared to LIDAR TCH (1.5 – 2.2 m RMSE) and generally high correlation values ($R^2 = 0.89 – 0.99$; Table 2, Figure 1b). Explanation of suggested optimal scanning conditions is provided in the sections below. Ecosynth TCH values showed relatively constant precision with an average standard deviation of 2.5 m, which was un-related to canopy height (Figure 1d). The correlation between Ecosynth TCH and field heights was low at SERC due to low variation in field heights (Chapter 2, Dandois and Ellis 2013). Error in tree height estimates was high at Knoll sites for both Ecosynth and LIDAR (7.9 m and 6.4 m, respectively).

Table 2: Summary of Ecosynth TCH estimates relative to field and LIDAR measurements averaged across optimal replicates at Herbert Run and for one replicate each of optimal settings at Knoll and SERC sites.

		Compared to Field Heights		Compared to LIDAR TCH	
Site	**N**	**RMSE (m)**	**R^2**	**RMSE (m)**	**R^2**
Herbert Run	7	4.5	0.86	2.2	0.99
Knoll	1	7.9	0.79	2.0	0.99
SERC	1	4.6	0.19	1.5	0.89

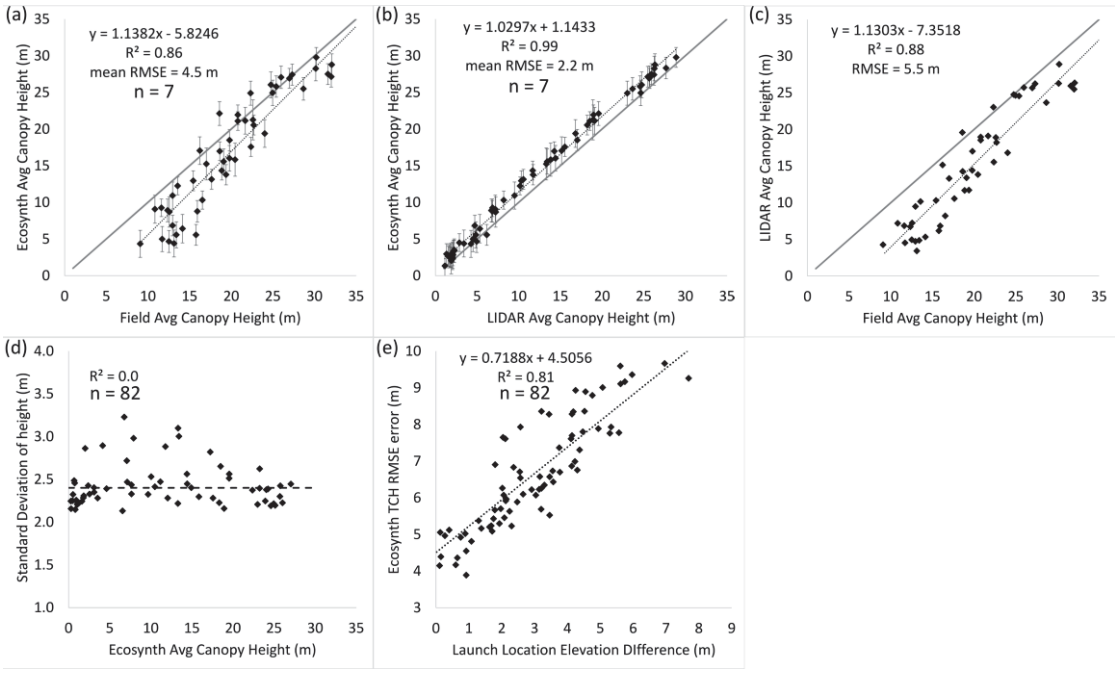


Figure 1: Regression between Ecosynth average TCH per plot across optimal replicates (n = 7) to (a) field and (b) LIDAR TCH, (c) LIDAR TCH and field measurements, (d) the standard deviation in Ecosynth TCH by height, and (e) Ecosynth TCH RMSE error relative to vertical offset of Ecosynth from LIDAR DTM (e) across all replicates (n = 82). Solid line (a - c) is one to one line, error bars in are standard deviation, dotted lines are linear model and dashed line in (d) is the mean.

Sources of error in Ecosynth TCH estimates of canopy height appear to include error in the vertical positioning of the point cloud relative to the DTM used for estimating heights, the amount of photographic overlap, and even scene lighting. A large portion of the vertical error in Ecosynth TCH predictions of height (RMSE) was explained by the

amount of vertical positioning error of the entire point cloud, as measured by the fit between Ecosynth and LIDAR point clouds (ICP-Z error, $R^2 = 0.71$, Table 3) and the difference between the estimated and known elevation of the launch location (Launch Elevation Difference, $R^2 = 0.73$, Figure 1e). Average RMSE error in Ecosynth TCH compared to both field and LIDAR estimates doubled as forward overlap was reduced from > 96 % to 60 % (RMSE = 3.7 m to 7.9 m, $R^2 = 0.99$; RMSE = 2.0 m to 3.9 m, $R^2 = 0.87$; Table 3). The strong relationship between error in Ecosynth TCH estimates of field canopy height and the launch location elevation difference (Figure 1e) across all replicates highlights the substantial effect of accuracy co-registration between surface and terrain models (Chapter 1, Section 1.1.7; Sithole and Vosselman 2004; Næsset 2009b; Ni et al. 2014). Accounting for this offset by adding the value of Launch Location Elevation Difference to the Z value of each point on per point cloud basis improved RMSE error across all datasets compared to field estimates (RMSE range = 3.2 – 6.1 m, average 4.2 m). Offset correction reduced the range of RMSE error in Ecosynth TCH estimates compared to LIDAR TCH and barely affected average error across all replicates (RMSE range = 1.2 – 4.0 m, average 2.5 m). Offset correction also improved error (reduced RMSE) in Ecosynth TCH estimates relative to field measurements for the set of datasets collected under optimal conditions (average RMSE = 3.7 m) but increased Ecosynth TCH error relative to LIDAR TCH (average RMSE = 3.1 m). The influence of accounting for DTM offset on other conditions (Table 3) was not addressed in the current study.

Table 3: Summary of the response of Ecosynth quality traits and metrics from all replicates (n = 82) across major scanning condition treatments of lighting, altitude above the canopy, and image side and forward overlap. Values are mean and standard deviation for all replicates within each level. Significant differences / trends are indicated with the p-value for lighting conditions, or correlation (R^2) for altitude and overlap. 'NS' indicates no significant trend at the $\alpha = 0.05$ level.

	Lighting condition			**Altitude above canopy (meters)**					**Image side overlap (%)**					**Image forward overlap (%)[a]**		
	CLEAR	**CLOUDY**	***p* <**	**20**	**40**	**60**	**80**	**R^2**	**0**	**25**	**50**	**75**	**R^2**	**96**	**60**	**R^2**
N	43	39		9	15	17	41		10	10	29	33		5	5	
Path-XY Error RMSE m	1.2 (0.6)	1.4 (1.3)	NS	0.61 (0.25)	1.0 (0.5)	1.2 (0.7)	1.6 (1.2)	0.98	1.9 (1.3)	2.2 (1.7)	1.2 (0.6)	0.9 (0.4)	0.74	1.2 (0.2)	1.6 (0.2)	0.64
Path-Z Error RMSE m	0.44 (0.13)	0.44 (0.12)	NS	0.4 (0.1)	0.5 (0.1)	0.4 (0.1)	0.5 (0.1)	NS	0.5 (0.1)	0.4 (0.1)	0.5 (0.1)	0.4 (0.1)	0.1	0.36 (0.1)	0.41 (0.1)	0.88
ICP-XY Error RMSE m	1.8 (0.8)	2.3 (1.2)	0.05	2.2 (1.1)	1.8 (0.7)	2.2 (1.2)	2.0 (1.2)	NS	2.1 (1.3)	2.8 (1.5)	1.7 (0.8)	2.1 (1.0)	NS	1.7 (0.3)	1.9 (0.3)	NS
ICP-Z Error RMSE m	2.0 (0.9)	3.8 (1.7)	0.01	3.4 (0.9)	2.7 (1.2)	2.3 (1.5)	3.0 (1.9)	NS	2.3 (1.5)	3.3 (2.0)	2.6 (1.2)	3.1 (1.8)	NS	1.8 (0.8)	2.2 (1.1)	NS
Launch Elevation Difference (m)	2.1 (1.5)	3.5 (2.4)	0.01	3.2 (0.9)	3.5 (1.2)	2.3 (1.8)	2.6 (2.6)	NS	0.7 (2.5)	3.4 (2.0)	2.9 (1.5)	3.1 (2.2)	0.48	1.8 (1.8)	2.4 (1.8)	NS
Ecosynth TCH to Field RMSE (m)	5.8 (1.2)	7.3 (1.8)	0.01	7.9 (0.6)	6.8 (1.2)	6.1 (1.8)	6.3 (2.6)	0.77	5.1 (1.3)	7.4 (1.6)	6.3 (1.3)	6.8 (1.8)	0.28	3.7 (0.9)	7.9 (1.6)	1.0
Ecosynth to LIDAR TCH RMSE (m)	1.8 (0.9)	3.1 (1.3)	0.01	3.3 (0.6)	2.0 (0.9)	2.0 (1.1)	2.5 (1.5)	NS	2.7 (2.1)	2.7 (1.4)	1.9 (0.9)	2.7 (1.1)	0.1	2.0 (1.3)	3.9 (1.2)	0.87
Point Density Points m^{-2}	33 (14)	43 (27)	0.05	80 (24)	53 (13)	38 (7)	23 (9)	0.98	14 (0.5)	18 (0.7)	34 (10)	54 (23)	0.92	37 (1)	0.7 (0.1)	0.67
Canopy Penetration % CV	18 (3)	16 (2)	0.01	16 (2)	16 (2)	18 (2)	15 (2)	NS	15 (3)	17 (3)	17 (3)	18 (2)	0.94	20 (1)	0 (1)	0.90
Computation Time hours	44 (38)	49 (46)	NS	104 (14)	70 (59)	53 (33)	23 (16)	0.98	8 (1)	10 (0.5)	26 (3)	87 (38)	0.79	45 (1.5)	0.5 (0.01)	0.46

[a] Image forward overlap was tested at decreasing increments of ≈ 4 % or every 2nd to 10th image, but only the largest and smallest levels are shown here to highlight the observed pattern.

This effect appeared to be explained by rapidly decreasing point cloud density as a function of forward overlap. RMSE error of Ecosynth TCH was negatively correlated with the logarithm of point cloud density for both field and LIDAR estimates ($R^2 = 0.99$ & 0.76, respectively; Figure 2). There was no significant effect of image resolution above 5 MP on Ecosynth TCH compared to field estimates and LIDAR, however, Photoscan reconstruction failed at resolutions below 5 MP. Error in Ecosynth TCH estimates increased as altitude decreased ($R^2 = 0.77$, Table 3), which may be due to reduced forward overlap at lower altitudes given constant camera shutter speed (≈ 2 fps) and UAS speed (≈ 6 m s^{-1}). Translating values of flight altitude to values of percent forward overlap: 96 % at 80 m, 95 % at 60 m, 92 % at 40 m and 84 % at 20 m, reveals the same strong trend between mean Ecosynth TCH error and forward overlap ($R^2 = 0.97$) that was observed when forward overlap was artificially reduced by sub-sampling images from the same set of replicates. Measures of absolute positioning accuracy (RMSE meters ICP-XY and ICP-Z error, Launch Elevation Difference) and TCH error relative to field measurements and LIDAR were higher on cloudy days compared to clear days ($p < 0.01$, Table 3).

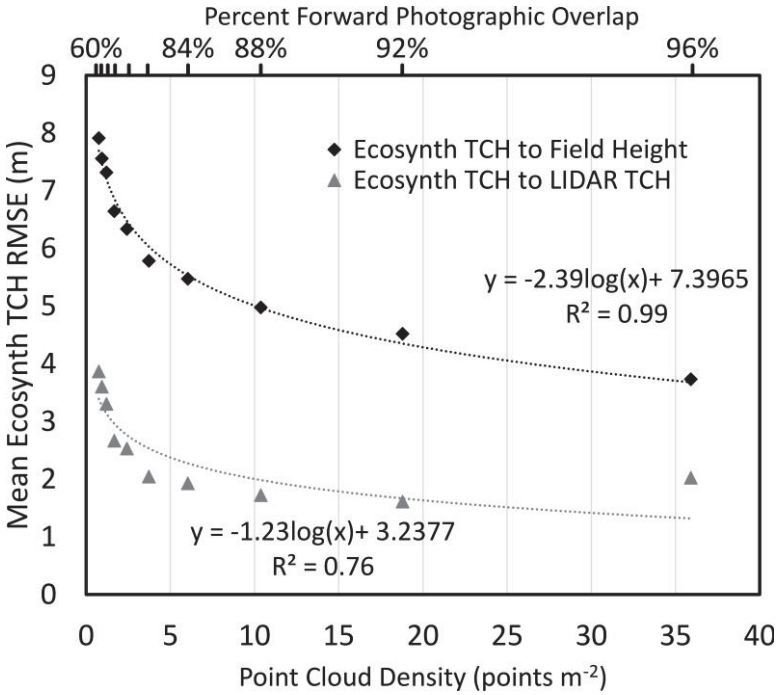


Figure 2: Scatter plots of the relationship between mean point cloud density and mean Ecosynth TCH error relative to field measurements and LIDAR TCH estimates for a single set of 5 replicates sub-sampled from every image to every 10th image to simulate decreasing forward photographic overlap. Top axis indicates corresponding forward overlap: 60%, 64%, 68%, 72%, 76%, 80%, 84%, 88%, 92%, and 96%.

Ecosynth estimates of per plot above ground biomass density (AGB) based on linear regression of TCH were highly correlated with field estimated AGB ($R^2 = 0.81$) with 64 Mg ha^{-1} RMSE error, or about 31 % of field estimated mean AGB density across all plots (Figure 3a). Ecosynth AGB estimates were relatively precise across all replicates with a 95 % CI of +/- 3.8 Mg ha^{-1} or ≈ 2% of mean AGB density. LIDAR estimates of field estimated AGB were similar to those from Ecosynth: $R^2 = 0.82$, RMSE = 59 Mg ha^{-1} or 29% of field estimated mean AGB (Figure 2b).

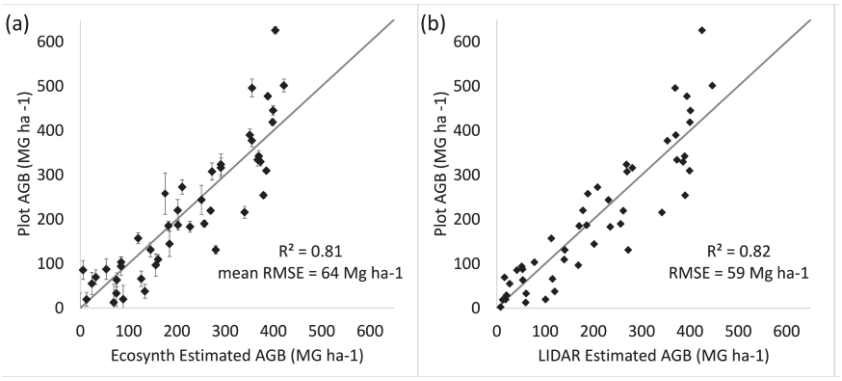


Figure 3: Average Ecosynth (a) and LIDAR (b) TCH estimated AGB (Mg ha^{-1}) relative to field estimated AGB. Solid line is one to one line, error bars are standard deviation of Ecosynth AGB per plot.

3.2 Influence of scanning conditions on point cloud positioning accuracy

Four empirical traits describing point cloud relative and absolute positioning accuracy were evaluated as a function of scanning conditions and relative to canopy height metrics. Relative horizontal positioning error (Path-XY) was strongly influenced by changes in flight altitude and photographic forward and side overlap ($R^2 = 0.64 – 0.98$, $p < 0.05$, Table 3) but in general was un-related or weakly correlated with canopy height and other metrics ($R^2 = 0.0 – 0.24$, Table 3). Relative vertical positioning error (Path-Z) was unrelated to canopy metrics and to changes in lighting and altitude, but increased with decreasing photographic overlap. However, measures of relative positioning error provide a valuable quality control metric for characterizing large errors in SFM reconstruction. At image resolutions below 5 MP, Photoscan SFM reconstructions completely failed and measures of relative positioning error reflected this in RMSE values > 30 m. Measures of absolute positioning error were only affected by changing light conditions and where un-affected by changes in all other scanning conditions. Horizontal (ICP-XY) and vertical (ICP-Z) absolute positioning error were higher under

cloudy conditions compared to clear conditions ($p < 0.05$, Table 3) and differences in lighting explained some of the variation in ICP-Z at different levels of altitude (interaction $p < 0.01$) and side overlap (interaction $p < 0.05$).

3.3 Influence of scanning conditions on measures of canopy sampling

Forest area point cloud density showed several distinctive trends within different treatments (Table 3). In general, point cloud density was highest under conditions of the highest amount of photographic forward and overlap. Point cloud density was unchanged by changing resolution (10, 7.5, to 5 MP) on the same set of images, which is counter to the understanding that increased image resolution will increase the number of image features and potentially the number of point cloud points (Lowe 2004) and is likely due to the setting for the maximum number of allowed features per image in Photoscan. However, these results are inconsistent with the observed relationship that point cloud density increased with decreasing flight altitude, ($R^2 = 0.98$), which should represent both an increase in observed detail in images and a decrease in photographic overlap at fixed UAS and camera speed. By comparison, average LIDAR first return point cloud density in forested areas was 10.8 points m^{-2}, similar to the acquisition specification of 10.06 points m^{-2}. There was a significant difference ($p < 0.05$) in mean point cloud density under different lighting conditions, with roughly 30 % more points m^{-2} observed on average on cloudy days compared to clear days (43 vs. 33 points m^{-2}).

Average canopy penetration (CP), measured as the average percent coefficient of variation of heights within 1 m x 1 m bins in forested areas, was significantly lower under cloudy conditions compared to clear lighting conditions (18 % vs. 15 %, $p < 0.01$, Table 3). This means that on average, Ecosynth point clouds observed up to 1 m deeper into

the canopy on clear days compared to cloudy days. CP was also significantly influenced by photographic overlap, increasing with increasing side overlap (15 % to 18 %, R^2 = 0.94) and decreasing with decreasing forward overlap (18 % to 2 %, R^2 = 0.90, $p < 0.001$). Average CP across all Ecosynth replicates was 17 %, by comparison average forest area CP for the LIDAR first return point cloud was 29 %. The difference between Ecosynth and LIDAR average CP in forested areas was significantly different than zero (one-sampled t-test: $t = 40.85$, $p < 0.05$) and equates to LIDAR observing roughly 2.5 m - 3.5 m deeper into the canopy than Ecosynth at canopy heights of 20 – 30 m.

3.4 Influence of wind of point cloud quality

Results of comparing mean wind speed during flight, based on the Beaufort wind force scale, to measures of Ecosynth traits and metrics are summarized in Table 4. In general, increasing wind speed led to increased error in point cloud vertical positioning and estimates of canopy height (R^2 = 0.13 – 0.99). Increased wind speed was weakly correlated with decreased relative and absolute horizontal positioning error (R^2 = 0.23 & 0.27) and was related to average forest area point cloud density. Even so, the only statistically significant trend at the $\alpha = 0.5$ level was between wind speed and average canopy penetration ($R^2 = 0.79$, $p < 0.05$), where increased wind speed led to a 12 % increase in CP (17 % to 19 %). These results suggests that the factor of wind speed alone is not enough for predicting error in Ecosynth TCH estimates, but that TCH error is likely to be greater on windier days compared to calm days.

Table 4: Variation in Ecosynth point cloud quality traits and metrics as a function of average wind speed during each flight based on the Beaufort wind force scale.

	Mean wind speed during flight Beaufort wind force ($m\ s^{-1}$) (mph)						
	1 (0.3 – 1.5) (1 – 3)	**2 (1.6 – 3.4) (4 – 7)**	**3 (3.5 – 5.4) (8 – 12)**	**4 (5.5 – 7.9) (13 – 17)**	**R**	**R^2**	**F-test**
Path-XY Error RMSE (m)	1.30	1.20	1.60	0.71	-0.48	0.23	NS
Path-Z Error RMSE (m)	0.38	0.43	0.48	0.51	0.99	0.99	NS
ICP-XY Error RMSE (m)	2.51	1.80	2.22	1.95	-0.52	0.27	NS
ICP-Z Error RMSE (m)	2.49	3.03	2.54	3.40	0.67	0.44	NS
Launch Location Elevation Difference (m)	1.63	3.01	2.90	2.91	0.73	0.54	NS
Ecosynth TCH to Field Height RMSE (m)	5.41	5.65	5.39	6.67	0.75	0.57	NS
Ecosynth TCH to LIDAR TCH RMSE (m)	2.48	2.48	2.08	2.93	0.36	0.13	NS
Average Forest Point Density (points m^{-2})	47	35	36	45	-0.11	0.01	NS
Average Forest Canopy Penetration (% CV)	17	16	18	19	0.89	0.79	$p < 0.05$

3.5 Radiometric quality of Ecosynth point clouds

Average radiometric variation across channels (RGB and grayscale intensity) ranged from 2.0 % - 5.9 % and was highly correlated with the structural complexity of the land surface, as measured by the mean rugosity or bumpiness of different landcovers (R^2 = 0.74, Figure 4, Table 5). The lowest values of average color-spectral variation (2.0 % - 3.1 %) were observed in the turf areas that were the most homogenous across the landscape and also had low structural complexity. The highest values of average color-spectral variation (3.3 % - 5.9 %) were observed in areas of forest and buildings. High variation is expected over forests as a result of the structural complexity of the forest canopy, due to relatively bumpy 3D structure (1.7 m rugosity) the color-spectral signal represents a mix of illuminated and shaded surfaces (Ogunjemiyo et al. 2005).

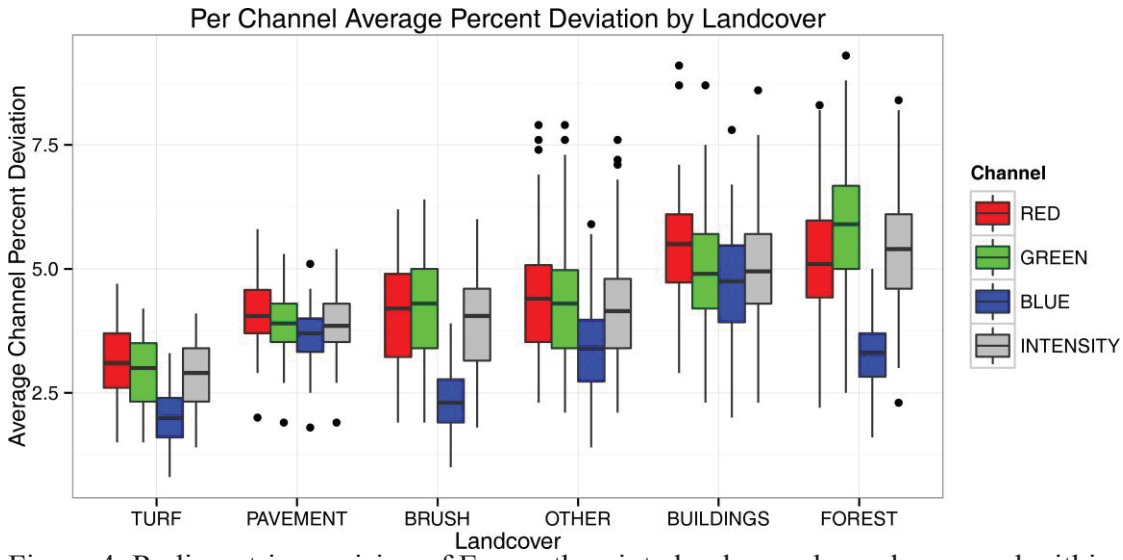


Figure 4: Radiometric precision of Ecosynth point clouds per channel averaged within different landcovers across all main replicates at the Herbert Run site (n = 82). Average channel percent deviation measures variation (standard deviation) in point color within 1 m x 1 m bins, interpreted as a percentage of the maximum potential channel brightness (255).

High color-spectral variation over buildings was not expected, as structurally simple roof-tops should be represented with a relatively homogenous signal, similar to that of pavement and turf. High variation in buildings may be the result of edge effects around building edges where due to horizontal positioning errors (section 3.2) sides of buildings were counted within the building landcover areas. Within forest areas, the average variation in point color per channel was strongly influenced by the lighting conditions of the flight (Figure 5). Color variation can also be considered a measure of contrast and results show that forest areas were observed with lower contrast on cloudy days compared to clear days ($p < 0.0001$ for all per channel and intensity differences). Relatively low image contrast is known to reduce the 'stability' or quality of computer vision feature detectors used in SFM reconstruction (Lowe 2004; Vonikakis et al. 2012) and this difference in contrast may explain the observed increase in point cloud vertical positioning and canopy height error on cloudy days compared to clear days (Table 3).

Table 5: Mean and standard deviation of the percent channel variation for red, green, and blue point colors as well as the point color gray scale intensity and mean rugosity by landcover at Herbert Run.

	Per Channel Mean (SD) Percent Deviation				Mean Rugosity (m)
	RED	GREEN	BLUE	INTENSITY	
TURF	3.1 (0.7)	2.9 (0.7)	2.0 (0.5)	2.8 (0.7)	0.45
PAVEMENT	4.1 (0.6)	3.9 (0.6)	3.6 (0.6)	3.9 (0.6)	0.41
BRUSH	4.1 (1.1)	4.1 (1.0)	2.3 (0.5)	3.9 (1.0)	1.18
OTHER	4.4 (1.2)	4.3 (1.2)	3.4 (1.0)	4.2 (1.2)	0.29
BUILDINGS	5.4 (1.1)	4.9 (1.1)	4.7 (1.0)	4.9 (1.1)	0.97
FOREST	5.3 (1.3)	5.9 (1.3)	3.3 (0.6)	5.4 (1.2)	1.73
				$R^2 = 0.74$	

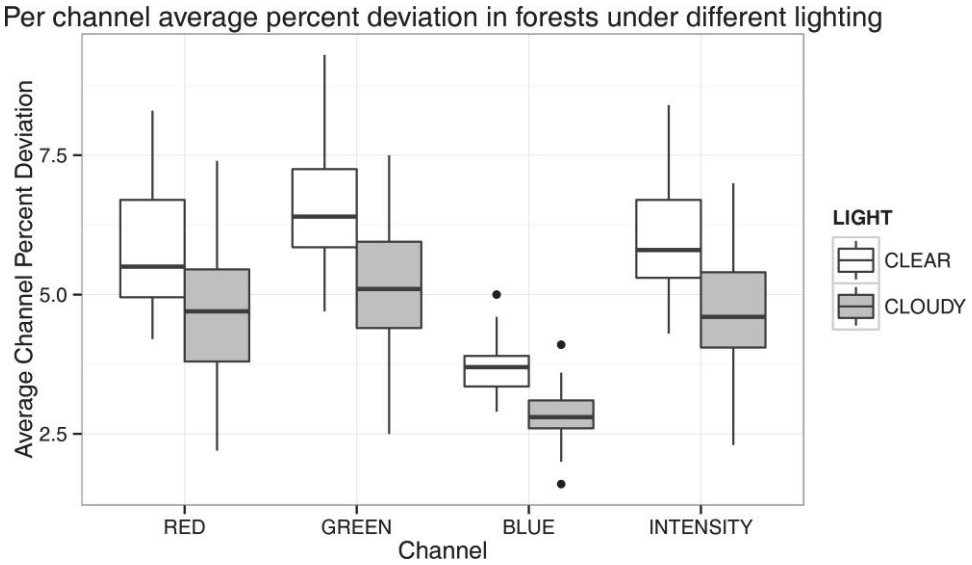


Figure 5: Average variation or contrast of point cloud point color values per channel within forest areas only under either clear or cloudy (diffuse) lighting conditions. All per channel differences in average variation under different lighting were significantly different based on analysis of variance ($p < 0.0001$).

3.6 Influence of computation resources on Ecosynth point cloud quality

Ecosynth point clouds processed in Photoscan (v0.91) required between 0.5 - 164 hours of computation, which was primarily a function of the number of photos used in the reconstruction and the resources of the computer used for processing. Across all

replicates, computation time was highly correlated with roughly the square of the number of photos, $R^2 = 0.96$, Figure 6). This is line with the understanding that in the process of identifying potential matches across images to supply to the bundle adjustment, an SFM algorithm needs to compare every image to every other image (i.e., N^2, Snavely et al. 2008). Number of photos, and therefore computation time, was also strongly dependent upon flight configuration. Flights with high overlap required more photographs, resulting in an increase in computation time (Table 3). Running the same set of photos in the same version of Photoscan on different computers resulted in little overall change in Ecosynth point cloud traits and metrics, but a large change in computation time (Table 6). The average variation in traits and metrics within replicates ranged from 0.1 – 6.7 % of the mean, with computation time being most affected by computer resources.

Table 6: Computation time, quality traits and metrics for the same 'standard' dataset run across 6 computer configurations, running version 0.91 of Photoscan.

	Average within replicate variation across 6 different computers	
Point Cloud Traits and Metrics	**SD**	**% CV**
Path-XY Error RMSE (m)	0.01 m	0.8 %
Path-Z Error RMSE (m)	0.002 m	0.7 %
ICP-XY Error RMSE (m)	0.11 m	6.7 %
ICP-Z Error RMSE (m)	0.06 m	3.9 %
Launch Location Elevation Difference (m)	0.04 m	1.5 %
Ecosynth TCH to Field Height RMSE (m)	0.01 m	0.2 %
Ecosynth TCH to LIDAR TCH RMSE (m)	0.01 m	1.2 %
Average Forest Point Density (points m^{-2})	0.02 points m^{-2}	0.1 %
Average Forest Canopy Penetration (% CV)	0.4 %	2.4 %
Computation Time (hours)	7.9 hours	18 %

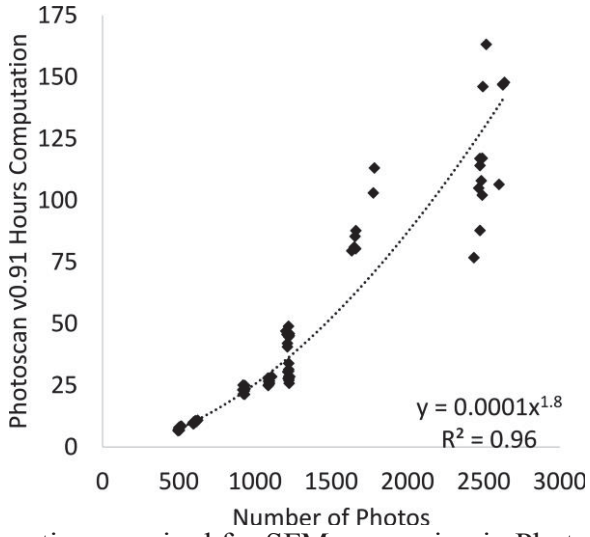


Figure 6: Computation time required for SFM processing in Photoscan v0.91 based on the number of photos.

Point clouds produced in different versions of Photoscan (v0.91 and v1.04) were nearly identical based on measures of point cloud quality traits and metrics, with the only noticeable difference being that the latest version of Photoscan required ¼ the computation time on the same computer, explained as 'improved image alignment performance' in the program's version change log (Table 7). Point clouds produced using the free and open-source Ecosynther implementation of the Bundler-PMVS pipeline were evaluated as 'sparse' and 'dense' point clouds. Sparse point clouds reflect the output of Bundler sparse bundle adjustment alone and points correspond to the image features identified and matched across images as part of the SFM reconstruction process (Snavely et al. 2008; Chapter 3). Dense models are produced by an additional level of processing that produces a dense set of points based on the 3D camera geometry produced in the sparse bundle adjustment stage (Furukawa and Ponce 2010), a step commonly used in other UAS-SFM remote sensing applications (Table 1, Chapter 1). Dense processing produced an 8x increase Ecosynther point cloud density. Ecosynther required roughly 50 % more computation time to produce a sparse model compared to Photoscan v0.91, with

dense point clouds requiring an additional 4.6 hours of computation after the sparse point clouds were produced. Ecosynther sparse point clouds were 20 % less dense than Photoscan point clouds and had half the canopy penetration (9 % vs. 18 %). Ecosynther dense point clouds were roughly 70 % more dense than Photoscan point clouds but also had lower canopy penetration (13 % vs, 18 %). TCH estimates of field measured canopy height from the sparse point cloud were similar to those from Photoscan, however the dense point cloud showed increased RMSE error (3.8 m – 5.3 m). The difference between Ecosynth and LIDAR TCH was lowest for the Ecosynther dense point cloud (RMSE = 2.0 m) compared to the sparse point cloud and Photoscan point clouds (RMSE = 2.9 – 3.0 m).

Table 7: Comparison of Ecosynth point cloud quality traits and metrics for a single replicate processed under different SFM algorithms.

Point Cloud Traits and Metrics	**Photoscan v0.84**[a]	**Photoscan v0.91** [a]	**Photoscan v1.04** [a]	**Ecosynther v0.8 Sparse** [b]	**Ecosynther v0.8 Dense** [b]
Path-XY Error RMSE (m)	1.1	1.1	1.1	1.1	1.1
Path-Z Error RMSE (m)	0.3	0.3	0.3	0.7	0.7
ICP-XY Error RMSE (m)	1.6	1.6	1.6	1.9	1.9
ICP-Z Error RMSE (m)	1.0	0.9	0.9	0.8	0.8
Launch Location Elevation Difference (m)	0.9	0.9	0.9	0.6	0.6
Ecosynth TCH to Field Height RMSE (m)	3.8	3.9	3.9	3.8	5.3
Ecosynth TCH to LIDAR TCH RMSE (m)	3.4	3.0	2.9	2.9	2.0
Average Forest Point Density (points m^{-2})	88	36	34	7	59
Average Forest Canopy Penetration (% CV)	18	18	18	9	13
Computation Time (hours)	30.4	44.9	10.7	61.8	Sparse + 4.6

[a] All versions of Photoscan run with same computer and settings, Chapter 2.
[b] Ecosynther v0.8 run on Amazon EC2 computer instance: g2.2xlarge, Intel Xeon E5-2670, 8 CPU processors, 15 GB RAM, NVIDIA GPU with 1,536 cores and 4 GB video RAM.

4.0 Discussion

4.1 Ecosynth remote sensing of forest canopies under varying conditions

Through a crossed experimental design with a relatively large number of replicates at the same site (n = 82), this research estimated field-measured average maximum canopy height based on the top-of-canopy height (TCH) of Ecosynth point clouds within 0.06 ha (25 m x 25 m) plots based on multiple data collection configurations. Analysis showed relatively high average RMSE error compared to field measured heights (mean RMSE = 6.5 m) and prior studies (RMSE = 1.7 – 4.4 m; Chapter 2; Dandois and Ellis 2010, 2013; Lisein et al. 2013). Even so, the large number of replicates collected at the same site revealed a consistent precision of Ecosynth estimates of canopy height (mean SD = 2.5 m). Under optimal conditions (clear skies, 80 m altitude above the canopy, and 75 % side photographic overlap) Ecosynth estimates of canopy height compared to both field and LIDAR measurements were significantly better. Average RMSE in Ecosynth TCH estimates of canopy height was 4.5 m compared to field estimates and 2.2 m compared to LIDAR TCH under optimal scanning conditions. Error in Ecosynth TCH measurements was strongly linked to the co-registration of Ecosynth point clouds to the DTM used for estimating heights. Accurate co-registration of DSMs and DTMs from different data sources represents a major challenge to accurately estimating canopy structure (St Onge et al. 2008; Ni et al. 2014) and remains a significant challenge for using Ecosynth to estimate canopy height. The importance of DTMs in estimates of canopy height is addressed more thoroughly in Chapter 2 (Dandois and Ellis 2013). Accounting for DTM offset based on the measure 'Launch Location Elevation Difference', substantially reduced the error in Ecosynth TCH estimates relative to field measurements: average

RMSE 4.5 to 3.7 m for the optimal set of point clouds, 6.5 m to 4.2 m average across all replicates. Interestingly, DTM offset correction actually increased RMSE error between Ecosynth and LIDAR TCH values within the set of optimal point clouds and across all replicates (average RMSE = 2.4 m to 2.5 m, 2.2 m to 3.1 m, respectively), although the reason why is unclear. This form of DTM correction can be easily applied to Ecosynth datasets collected at other sites as it only requires a single reference measure of the altitude of the ground where the copter was launched, a measure that could be obtained from a high quality DTM (e.g., LIDAR) or by mapping-grade GPS (e.g., Trimble Geo XT with sub-meter precision). While this form offset correction did show improved estimates of canopy height, the effect of this correction on the results of the factorial experimental design (Table 3) were not addressed. Future research should reevaluate the observed trends in point cloud traits and metrics after mis-registration between DTM and point cloud has been addressed.

Even with errors in co-registration between Ecosynth DSMs and LIDAR DTMs, Ecosynth TCH estimates across three different sites were highly correlated with those from LIDAR (R^2 = 0.83 - 0.99) and with relatively low error (RMSE = 1.7 - 2.5 m) compared to results from field measurements (Table 2). That both Ecosynth and LIDAR TCH underestimated field measured average maximum canopy height by on average > 4 m may be explained in part by error in field measurements, but also in the way in which in the canopy is described by 3D canopy height models (CHMs). Field based estimates of tree height have been found to have significant error (1 m to > 6 m) owing to the challenges in observing the top of trees from the ground below (Bragg 2008; Goetz and Dubayah 2011; Larjavaara and Muller-Landau 2013). In this study, field measurements

describe per plot canopy height based on the average height of the 5 tallest trees per plot (by DBH), representing the 5 tallest points of the outer canopy surface within a plot if all tree crowns reach the outer surface. The measure of top-of-canopy height (TCH), represents the average height over the entire outer surface of the canopy as observed by the remote sensing instrument, whether Ecosynth or LIDAR (Asner and Mascaro 2014). It is not surprising then that TCH underestimates field estimated height since field measures are from the tallest observed point for the largest trees in a plot and most of the canopy surface observed by LIDAR or Ecosynth should technically be lower than these tallest points, whether for a single crown or multiple crowns. Ecosynth TCH estimates of field-estimated AGB followed similar patterns relative to field and LIDAR measurements due to the fact that AGB is modeled from per plot height estimates.

4.2 Recommendations for optimal data collection

Along with enabling estimates of the precision of Ecosynth point cloud measurements, the large number of replicates collected in this research at varying levels of light, altitude, photographic overlap, and wind, makes it possible to suggest optimal data collection configurations for future UAS-SFM remote sensing. Optimal Ecosynth point cloud quality (low geometric error, low canopy height error, high point cloud density, and high canopy penetration) was obtained under configurations that maximized photographic overlap. All other factors being equal, such conditions led to reduced error in canopy height estimates, increased point cloud density, and increased penetration of points into the canopy. Within the set of replicates collected here, optimal conditions were achieved on clear days, at an altitude of 80 m above the canopy and with 75 % side photographic

overlap (Table 2, Figure 1). Table 8 summarizes the Ecosynth point cloud quality traits and metrics obtained under these optimal conditions.

Table 8: Average Ecosynth point cloud quality traits and metrics for the 'optimal' Ecosynth data collection: clear skies, 80 m altitude, 75 % side overlap processed with Photoscan v0.91, n=7.

Point Cloud Traits and Metrics	**Average Ecosynth Quality Traits and Metrics**
Path-XY Error RMSE (m)	1.1 m
Path-Z Error RMSE (m)	0.5 m
ICP-XY Error RMSE (m)	1.7 m
ICP-Z Error RMSE (m)	1.6 m
Launch Location Elevation Difference (m)	1.2 m
Ecosynth TCH to Field Height RMSE (m)	4.5 m
Ecosynth TCH to LIDAR TCH RMSE (m)	2.2 m
Average Forest Point Density (points m^{-2})	35 points m^{-2}
Average Forest Canopy Penetration (% CV)	20 %
Computation Time (hours)	45.4 hours

However, increased overlap also resulted in an increased number of photos which significantly increased computation time ($\approx N^{1.8}$, based on Photoscan v0.91 across multiple computing platforms). Flight configurations that minimize the number of photos while still providing a large amount of photographic overlap will provide an optimal trade-off between point cloud quality and computation time. Eisenbeiss (2009) also found improved results with high amounts of photographic overlap > 75 %. Computation time could be reduced by running SFM with only every other image from the data collection, at the estimated mean cost of roughly 0.5 m increased error in field height predictions.

The relationship between photographic overlap and point cloud quality may be related both to aspects of SFM reconstruction and also to the way in which the canopy is sampled by the SFM point cloud. From the point of view of SFM, decreased photographic overlap equates to an increase in the spacing between adjacent images and in the angle between views (view-angle) on the same feature from multiple images. Increasing the view-angle on the same feature leads to rapidly decreasing performance in the ability of the feature descriptor to correctly and consistently match the location, scale, and orientation of the same feature in different images (Lowe 2004). While it is not possible to know exactly what feature descriptor is used by Photoscan, it is said to be SIFT-like (Chapter 2), and the ability of SIFT to match features across images decreases rapidly after the view-angle increases beyond 20° (Lowe 2004). Based on the camera and UAS settings used here, view angles exceeded this threshold at 80 m above canopy height when photographic overlap was less than 72 %. Without additional levels of processing, SFM will only create points that have been matched as features across images and which contribute to the bundle adjustment solution (Chapter 3; Snavely et al. 2008). A decreased number of feature matches could then lead to the observed decrease in point cloud density, resulting in reduced sampling of the canopy and potentially increased error in canopy height estimates (Næsset 2009). In addition, those points that are matched under conditions of reduced overlap and view-angle may have reduced reliability of location estimation within and across images (Lowe 2004) resulting in reduced quality of 3D point locations and increases in height error.

This study found that the optimal weather conditions for Ecosynth UAS-SFM scanning were days with clear skies and light wind speeds (< 3 mph or 1.5 m s^{-1}).

Cloudy days resulted in reduced contrast in images of the canopy and increased error in both positioning accuracy traits and canopy structure metrics, potentially due to the link between image contrast and the stability of SFM feature detectors (Chapter 3, Lowe 2004, Vonikakis et al. 2012). Flights carried out on windy days showed increased error in point cloud quality traits and metrics, but differences based on wind speed alone were not significantly different than zero. Even so, high wind speeds should be avoided as they will cause the UAS to use more power to maintain position and reach waypoints during autonomous flight, resulting in reduced potential flight time that may lead to incorrect estimates of available battery life for a given mission. At average wind speeds up to 17 mph or 7.9 m s^{-1} the UAS was able to carry out missions without running out of battery power. Increased wind speed may increase error in 3D point cloud quality by introducing error into the SFM-bundle adjustment step that assumes that the only thing moving is the camera and not the features. On windy days the image features (Chapter 3) will move between image frames (leaves fluttering or branches swaying) and SFM may incorrectly estimate the 3D location of the feature because of this motion.

Further exploration of the relationship between measures of point cloud quality and the behavior of the SFM algorithm is limited by the fact that a commercial and closed source SFM algorithm was used in computation. Future SFM remote sensing research should consider the potential of using a free and open source computation algorithm to enable additional analysis of the relationship between image features and data quality. Comparisons in Ecosynth point cloud quality traits and metrics between Photoscan and a free, open-source SFM algorithm (Ecosynther) revealed similar point cloud positioning accuracy regardless of algorithm, but large differences in TCH estimates, point cloud

density, canopy penetration, and computation time depending on the level of post-processing applied in the algorithm (e.g., sparse versus dense processing). Even so, the use of a free, open source SFM algorithm provides access to the image features used in reconstruction and may make it possible to extract additional information about the relationship between point cloud quality and features as well as new canopy metrics from image features (e.g., mapping of fruits or flowers, metrics of leaf clumping or gap size). For example, access to image features would make it possible to compare the relationship between feature association (e.g., bright spots vs. dark spots, Chapter 3) and point cloud quality.

4.3 Future research on Ecosynth UAS-SFM point cloud data quality

While this research captured the relationship between Ecosynth point cloud quality and several key aspects describing data collection configurations, there are many more potential interactions that remain unexplored. Future work can take advantage of the data quality framework established here to evaluate other data collection and processing configurations, including different flight configuration levels, additional SFM post-processing, changing camera settings, or other weather conditions.

This research only considered a few major levels of photographic overlap and flight altitude and it is unclear if other levels in between those tested would provide even more optimal data quality or reduced computation time. For example, we tested levels of 50 % and 75 % side overlap and found that point cloud quality was better with more side overlap at the cost of more photographs and increased computation time. Future research could consider if similar improvements in data quality can be achieved by going from 50 % to 60 % side overlap, which would require fewer images for a given area, camera

settings, and flight altitude than 75 % side overlap. While there was a strong connection between increasing Ecosynth point cloud quality and decreasing photographic overlap (Table 3), all levels of side overlap were not fully crossed with all levels of altitude (Table 1). While capturing these additional levels within the original data collection may have provided more potential criteria for considering optimal flight configurations, all of the levels not captured represent decreases in both forward and side photographic overlap. Given the already strong connection between overlap and point cloud quality, it is expected that these additional levels would support the current results while increasing the sample size and computation requirements by roughly 40 % with full replication. Photoscan SFM alignment settings were left constant throughout all image processing (Section 2.2) and it is not known how changing these settings will influence point cloud quality or help explain observed differences in point cloud quality (Section 3.3). This research also did not explore additional levels of post-processing available in Photoscan, which includes using estimates of camera location from UAS GPS telemetry to reduce computation time by limiting the number of required image comparisons, processing that produces a more dense point cloud after the initial SFM reconstruction, or changing the maximum number of points per image to use. It is unclear then if there exists a combination of reduced image resolution and forward overlap, plus different levels of SFM computation that could produce point clouds of similar or better quality but with reduced computation requirements.

This research highlighted a strong link between the contrast with which the forest canopy is observed in images and the subsequent quality of structure measurements obtained from SFM reconstruction. The same camera with the same calibration settings

was used for all data collection, but it is unclear if different camera settings or even image post-processing could improve or change results. Varying camera exposure settings or even capturing multiple exposure settings over the same area (i.e., 'bracketed' exposure; Cox and Booth 2009) may produce images with improved contrast. In addition, image color corrections like histogram equalization may be useful for artificially enhancing image contrast to reduce the influence of scene lighting conditions during flight. This may be particularly useful for days when it is not possible to collect images under constant uniformly clear or uniformly cloudy conditions where the fixed exposure settings at the beginning of the flight and rapidly changing cloud cover result in images that are alternately over- or under- exposed (Chapter 2; Dandois and Ellis 2013). It is also possible to carry out near-infrared remote sensing with UAS using modified off-the-shelf digital cameras (Hunt et al 2005) or custom light-weight multi-spectral cameras (Berni et al. 2009; Turner et al. 2012) and future work should consider the potential of using such sensors for mapping canopy NDVI (Normalized Difference Vegetation Index) at high spatial resolution and in 3D, providing links between canopy structure, optical properties, and biophysical parameters (Dandois and Ellis 2013; Ryu et al. 2014). Finally, future research should more closely examine the relationship between point cloud quality and the features identified in images for SFM reconstruction. This link could explain the observed relationship between image contrast and measures of the 3D structure quality of forest canopies. Such questions could not be explored in this chapter in the same way as was done in Chapter 3 because the closed source nature of Photoscan prohibits access to the image features. Evaluation of the patterns observed in this research through use of point clouds generated by free and open-source SFM algorithms

like Ecosynther, may help shed light on the vital role that the image feature descriptor plays in SFM remote sensing, an element of the 'sensor system' which may be as fundamental as the image pixel in optical image remote sensing or the laser spot in LIDAR remote sensing.

5.0 Conclusions

The measurement of vegetation structural and spectral traits by automated SFM algorithms and consumer UAS is a rapidly growing part of the field of remote sensing. However, because UAS-SFM remote sensing is user-deployed and based on user-specified hardware (UAS and camera), software (SFM algorithms and post-processing), and data collection configurations (altitude and overlap), datasets can be produced under many different 'sensor' configurations and there exists no clear understanding of data quality measures or of the relationship between data quality and data collection configurations. This research defined several diagnostic quality measures based on LIDAR and optical image remote sensing, including point cloud positioning accuracy, point cloud density, canopy penetration, and radiometric noise or contrast, that can be used to describe point clouds and also to characterize sources of error in canopy height measurements. Using these quality measures, strong links were identified between vegetation structure measurements and the behavior of the SFM algorithms relative to the amount of photographic overlap and also to the quality of input images (image contrast). Recommendations were made for optimal flight configurations for future research aimed at characterizing canopy structure. UAS-SFM research is currently inspired to recreate the functionality of LIDAR for characterizing ecosystem vegetation structure, likely in large part due to the fact that the datasets produced by such systems are made available in

a similar form: a 3D point cloud. Yet the future of UAS-SFM remote sensing lies not in how it can recreate existing tasks, but how in it can be used to integrate computer vision research into ecological remote sensing, including the potential for automated mapping and identification of canopy objects like leaves, fruits, and flowers through the use of image features. If UAS provides field scientists with a birds-eye view of the landscape, computer vision will provide the 'ecologists-eye' view of the elements within that landscape.

References

Anderson, K., & Gaston, K.J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. Frontiers in Ecology and the Environment, 11, 138-146.

Asner, G.P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. Remote Sensing of Environment, 140, 614-624.

Baltsavias, E., Pateraki, M., & Zhang, L. (2001). Radiometric and geometric evaluation of Ikonos GEO images and their use for 3D building modelling. In (pp. 19-21): Citeseer.

Berni, J., Zarco-Tejada, P.J., Su•rez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. Geoscience and Remote Sensing, IEEE Transactions on, 47, 722-738.

Bragg, D. (2008). An improved tree height measurement technique tested on mature southern pines. Southern Journal of Applied Forestry, 32, 38-43.

Castillo, C., Pérez, R., James, M.R., Quinton, J.N., Taguas, E.V., & Gómez, J.A. (2012). Comparing the Accuracy of Several Field Methods for Measuring Gully Erosion. Soil Science Society of America Journal, 76, 1319-1332.

Chasmer, L., Hopkinson, C., Smith, B., & Treitz, P. (2006a). Examining the Influence of Changing Laser Pulse Repetition Frequencies on Conifer Forest Canopy Returns. Photogrammetric Engineering and Remote Sensing, 72, 1359-1367.

Chasmer, L., Hopkinson, C., & Treitz, P. (2006b). Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Canadian Journal of Remote Sensing, 32, 116.

Cook, B.D., Bolstad, P.V., NÊsset, E., Anderson, R.S., Garrigues, S., Morisette, J.T., Nickeson, J., & Davis, K.J. (2009). Using LiDAR and quickbird data to model plant production and quantify uncertainties associated with wetland detection and land cover generalizations. Remote Sensing of Environment, 113, 2366-2379.

Cox, S.E., & Booth, D.T. (2009). Shadow attenuation with high dynamic range images. Environmental Monitoring and Assessment, 158, 231-241.

Dandois, J.P., & Ellis, E.C. (2010). Remote sensing of vegetation structure using computer vision. Remote Sensing, 2, 1157-1176.

Dandois, J.P., & Ellis, E.C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.

de Matías, J., Sanjosé, J.J.d., López-Nicolás, G., Sagüés, C., & Guerrero, J.J. (2009). Photogrammetric Methodology for the Production of Geomorphologic Maps: Application

to the Veleta Rock Glacier (Sierra Nevada, Granada, Spain). Remote Sensing, 1, 829-841.

Dey, D., Mummert, L., & Sukthankar, R. (2012). Classification of plant structures from uncalibrated image sequences. In, Applications of Computer Vision (WACV), 2012 IEEE Workshop on (pp. 329-336).

FAA (2008). Global Positioning System Wide Area Augmentation System (WAAS) Performance Standard. In D.o. Transportation (Ed.).

Furukawa, Y., & Ponce, J. (2010). Accurate, Dense, and Robust Multiview Stereopsis. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32, 1362-1376.

Goetz, S., & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management, 2, 231-244.

Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sensing of Environment, 108, 254-263.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sensing, 4, 1573-1599.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S.n., & Falkowski, M.J. (2012). Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. Remote Sensing of Environment, 123, 25-40.

James, M., Applegarth, L.J., & Pinkerton, H. (2012). Lava channel roofing, overflows, breaches and switching: insights from the 2008–2009 eruption of Mt. Etna. Bulletin of Volcanology, 74, 107-117.

James, M., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of Geophysical Research: Earth Surface (2003–2012), 117.

Javernick, L., Brasington, J., & Caruso, B. (2014). Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. Geomorphology, 213, 166-182.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., & Birdsey, R.A. (2003). National-Scale Biomass Estimators for United States Tree Species. Forest Science, 49, 12-35.

Jung, K., Kaiser, S., Böhm, S., Nieschulze, J., & Kalko, E.K.V. (2012). Moving in three dimensions: effects of structural complexity on occurrence and activity of insectivorous bats in managed forest stands. Journal of Applied Ecology, 49, 523-531.

Larjavaara, M., & Muller-Landau, H.C. (2013). Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. Methods in Ecology and Evolution, 4, 793-801.

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., & Lejeune, P. (2013). A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. Forests, 4, 922-944.

Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60, 91-110.

Mascaro, J., Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Martin, R.E., Anderson, C., Higgins, M., & Chadwick, K.D. (2014). A Tale of Two "Forests": Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. PLoS ONE, 9, e85993.

Mathews, A., & Jensen, J. (2013). Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. Remote Sensing, 5, 2164-2183.

Næsset, E. (2009a). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment, 113, 148-159.

Næsset, E. (2009b). Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. Remote Sensing of Environment, 113, 2210-2223.

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating Forest Biomass and Volume Using Airborne Laser Data. Remote Sensing of Environment, 24, 247-267.

Ni, W., Ranson, K.J., Zhang, Z., & Sun, G. (2014). Features of point clouds synthesized from multi-view ALOS/PRISM data and comparisons with LiDAR data in forested areas. Remote Sensing of Environment, 149, 47-57.

Peli, E. (1990). Contrast in complex images. Journal of the Optical Society of America A, 7, 2032-2040.

Poli, D., Remondino, F., Angiuli, E., & Agugiaro, G. (2014). Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction. ISPRS Journal of Photogrammetry and Remote Sensing.

Rosnell, T., & Honkavaara, E. (2012). Point Cloud Generation from Aerial Image Data Acquired by a Quadrocopter Type Micro Unmanned Aerial Vehicle and a Digital Still Camera. Sensors, 12, 453-480.

Ryu, Y., Lee, G., Jeon, S., Song, Y., & Kimm, H. (2014). Monitoring multi-layer canopy spring phenology of temperate deciduous and evergreen forests using low-cost spectral sensors. Remote Sensing of Environment, 149, 227-238.

Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 59, 85-101.

Snavely, N., Seitz, S., & Szeliski, R. (2008). Modeling the world from internet photo collections. International Journal of Computer Vision, 80, 189-210.

Tao, W., Lei, Y., & Mooney, P. (2011). Dense point cloud extraction from UAV captured images in forest area. In, 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services. Fuzhou: June 29 2011 - July 1 2011.

Turner, D., Lucieer, A., & Watson, C. (2012). An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SFM) Point Clouds. Remote Sensing, 4, 1392-1410.

Vonikakis, V., Chrysostomou, D., Kouskouridas, R., & Gasteratos, A. (2012). Improving the robustness in feature detection by local contrast enhancement. In, Imaging Systems and Techniques (IST), 2012 IEEE International Conference on (pp. 158-163).

Wasser, L., Day, R., Chasmer, L., & Taylor, A. (2013). Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers During Leaf-Off and Leaf-On Conditions. PLoS ONE, 8, e54776.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., & Reynolds, J.M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314.

Yang, X., & Lo, C. (2000). Relative radiometric normalization performance for change detection from multi-date satellite images. Photogrammetric Engineering and Remote Sensing, 66, 967-980.

Chapter 5: Conclusions

The main objectives of this research were to define a new form of remote sensing of forest canopies (Ecosynth), to improve understanding of what is being measured by this system by addressing the question 'What is a point cloud point?', and to evaluate the quality of Ecosynth point clouds of forest canopies and how changing the way the data is collected will influence data quality. Ecosynth remote sensing with structure from motion (SFM) and unmanned aerial systems (UAS) was introduced along with a processing pipeline to produce ecological products from SFM three-dimensional (3D) point clouds. This pipeline was used to produce estimates of forest canopy height and canopy color-spectral phenology in 3D (Chapter 2). The nature of Ecosynth point cloud points was evaluated by examining the 3D-color quality of individual points at painted targets in a tree and based on comparison of point image features to manual classification of image semantic content (Chapter 3). Using a crossed experimental design, the quality of Ecosynth point clouds of forest canopies was evaluated across a range of different flight plans, weather conditions, and image processing with recommendations for optimal scanning configurations (Chapter 4). Ultimately, the computer vision 'image feature' was identified as a prominent factor in understanding the quality of both individual Ecosynth points and of the entire point cloud.

Three major research questions were posed in Chapter 1 that are arguably difficult or impossible to examine with existing forms of remote sensing, due to either technical or logistical constraints, or both. To review, those questions were: What are the spatial and temporal patterns of forest canopy phenology at the scale of individual crowns / tree species and how does this vary across within and across sites? What are the dynamics of canopy

biomass density at seasonal and longer time intervals across different forest types? How does the amount of tree flower and fruit production vary through time and as a function of climate, species, and local edaphic conditions? The major outcome of the dissertation is that researchers now have access to the necessary tools for attempting to answer these questions. Ecosynth computer vision ecological remote sensing enables on-demand observations of canopy 3D structural and spectral traits at high spatial resolution and at frequent intervals, a feat not possible or practical with other forms of remote sensing to date. This dissertation research has demonstrated that Ecosynth ecological remote sensing techniques can be used repeatedly to collect high resolution observations of canopy structural and spectral traits at the scale of individual trees (Chapter 2, Chapter 4). Research that aimed to improve understanding of what SFM remote sensing is observing with respect to the forest canopy revealed the strong role that the computer vision image feature played in how the canopy is seen, opening the door for more targeted feature based remote sensing of forest canopies, for example for tracking individual flowers on a crown (Chapter 3). The sections below summarize the major findings and contributions of each research chapter and highlight potential areas for future research. That is followed by recommendations for how to take advantage of the potential for using computer vision image features in Ecosynth UAS-SFM remote sensing of forest canopies.

1.0 Major findings and contributions of research chapters

1.1 Major findings and contributions from Chapter 2

Chapter 2 presented Ecosynth remote sensing as a combination of UAS hardware, SFM software, and post-processing steps for producing forest canopy data products over small spatial extents (< 10 ha), but at high spatial resolution and high frequency. Ecosynth

measurements of canopy height were comparable to LIDAR and field based measurements, but the quality of height measurements depended strongly on the quality of the digital terrain model (DTM) used for estimating heights from point cloud elevations. DTMs produced from Ecosynth point clouds under leaf-off forest conditions produced more accurate estimates of canopy height than those produced from leaf-on point clouds owing to the fact that leaf-off point clouds observed more of the forest floor. However, leaf-off point clouds still showed considerable errors relative to a more accurate LIDAR DTM and generally overestimated the elevation of the understory terrain, leading to underestimates of canopy height. Ecosynth point clouds were also produced over a single forest at six times across a 16 month period, capturing a full cycle of canopy phenology. Ecosynth vertical profiles of both 3D structure and color tracked the change in the canopy phenology from the winter leaf-off state through the spring greening, summer, and then autumn senescence. The pattern of canopy greenness matched that of satellite observations from the Moderate Resolution Imaging Spectrometer (MODIS; Zhang et al. 2003) and also revealed dynamics at the scale of individual trees, highlighting the use of Ecosynth techniques for observing local canopy dynamics not observed by coarse-resolution satellite sensors.

The major contributions of Chapter 2 include the first ever use of UAS and SFM for characterizing the dynamics of canopy structure and color throughout a growing season as well the Python based pipeline used for processing point clouds into products (Dandois and Ellis 2013). The 'Ecosynth Aerial' pipeline is built on this research and the initial work of the lead author, but was developed into its current state through the help of many others and is now available online as a free and open-source code repository

(http://code.ecosynth.org/ecosynthaerial, visited 2014-06-08). This chapter is already inspiring new research using Ecosynth techniques. Building on an existing forest restoration experiment (Holl et al. 2011), Ecosynth techniques are being used to characterize canopy height within different planting treatments and also to evaluate the potential of using Ecosynth canopy structure measurements (e.g., height, roughness, openness) to predict the abundance of frugivorous and insectivorous bird species with promising results. Ecosynth techniques will soon be deployed within existing forest plots of the Smithsonian Tropical Research Institute (STRI) in Panama to track the phenology of the tropical forest canopy.

1.2 Major findings and contributions of Chapter 3

Research of chapter 3 was motivated to address the question 'What is a point cloud point?' and was carried out using techniques inspired from both remote sensing and computer vision. The quality of point cloud point 3D location and RGB color fusion was evaluated by identifying painted targets placed in a single free-standing tree that was scanned by ground-based Ecosynth SFM and terrestrial LIDAR (TLS) under leaf-on and leaf-off conditions. Results showed that Ecosynth point 3D-color accuracy was highest under conditions when targets were clearly visible in images, but that this depended strongly upon what was actually 'seen' by the computer vision image feature descriptor SIFT (Scale Invariant Feature Transform; Lowe 2004). SIFT appeared to favor areas of high contrast and so painted targets that were imaged with low contrast were often not observed by SFM points. This finding was re-enforced by analysis of SIFT features as numerical feature descriptors and as image feature thumbnail tiles, which revealed that SIFT favors parts of the image that are either brighter or darker than the surroundings

following a roughly two-dimensional Gaussian shape. There was a weak relationship between SIFT feature descriptors and manually identifying image tags, indicating that points alone may not be directly linked to any real-world semantic meaning.

The first major contribution of this research was in the creation several datasets that could be used for development of computer vision tools for making observations of forests and vegetation. Results of single tree scanning were strongly dependent on the use of the SIFT feature detector, but future research can make use of the image datasets and high quality TLS 3D scans to evaluate different SFM implementations using different feature detectors. The other major contribution of this research is the Python-based graphical user interface (GUI) for tagging image features and a database of classified point cloud image features that could be used for additional computer vision ecological research. Computer vision image databases used for image feature research typically contain pictures of faces, buildings, cars, and similar objects and if forests or vegetation are contained in image datasets it is often if the form of landscape-style pictures and not the high resolution aerial images used in this research (Fei-Fei and Perona 2005). Recent research on the use of image features for 'geographic image retrieval' focused on discriminating between different types of landcovers alone and did not focus on evaluating features within forest canopies (Yang and Newsam 2013). Using a 'bag of features' approach, future research could use the datasets and GUI used here to evaluate the use of image features for classifying parts of the forest canopy, including leaf clumps, gaps, or even fruits and flowers.

1.2 Major findings and contributions of Chapter 4

The research of Chapter 4 focused on defining measures of the quality of Ecosynth point clouds of forest canopies and evaluating how those measures of quality change as the conditions of data collection change. Measures of point cloud quality included traits describing the geometric positioning accuracy, point cloud density, and canopy penetration, as well as metrics describing the estimated error of canopy height estimates relative to field and LIDAR measurements. The results showed that photographic overlap made the largest contributions to explaining variation in point cloud quality across a range of data collection configurations. Results also suggested that point clouds produced from images collected on cloudy days had increased error compared to when images were collected on clear days. Images on cloudy days had significantly less contrast than those from clear days and increased error in height measurements may have been due to reduced stability of image feature locations as a result of reduced contrast, a result which is in line with understanding of feature detector behavior (Lowe 2004).

The major contribution of the research presented in Chapter 4 is an understanding of how measures of point cloud quality will vary as the data collection conditions vary. Prior UAS-SFM research was collected across a broad range of altitudes and levels of overlap (see Chapter 1) and prior to this research there was no clear understanding of how those differences would influence results. In addition, other studies typically only focus on one measure of point cloud quality, error relative to LIDAR or field measurements, this study highlighted the importance of looking at other measures including point cloud density and canopy penetration. Finally, through the collection of a large sample of point cloud datasets (n = 82) spanning a range of scanning conditions,

this research was able to make recommendations for optimal data collection strategies, specifically the recommendation to maximize photographic overlap for a given area while minimizing the number of photos to reduce SFM computation costs.

2.0 The rise of computer vision ecological image features

Prior research has focused exclusively on how the products produced by UAS-SFM remote sensing systems (3D point clouds, canopy height models, digital surface models) are like LIDAR and can be used as a substitute for LIDAR and LIDAR-color fusion datasets. While this dissertation shows how that interpretation can lead to novel measurements of forest canopies that are not possible or practical with LIDAR, for example the observation of canopy structural and spectral phenology (Chapter 2), it does not address the full potential of computer vision remote sensing of forests. One overarching finding of this dissertation research was the recognition of the important role that computer vision image features play in the way in which forest canopies are observed by a UAS-SFM remote sensing system. The image feature was the focus of Chapter 3 and it was demonstrated that the properties of the 3D point cloud produced by SFM are strongly dependent upon the image feature detector used for matching images. Results from Chapter 4 suggest that image features also play an important role in understanding the quality of Ecosynth point clouds of forests under different scanning conditions with point cloud quality being strongly influenced by photographic overlap and image contrast, factors which will strongly influence the behavior of feature detection and matching algorithms (Lowe 2004).

Based on these findings, image features are defined here as the fundamental observation unit or minimum mapping unit of computer vision SFM remote sensing. By

comparison, a picture element or pixel is the minimum unit of optical image remote sensing, as is the laser pulse for LIDAR remote sensing (Campbell 1996; Dubayah and Drake 2000; Glennie 2007). This definition makes it possible to interpret the quality and behavior of SFM point clouds of forest canopies, as was done in Chapters 3 and 4, but also opens the door for new research applications that leverage both aerial remote sensing and computer vision research. As discussed in Chapter 1, image features are used for a number of applications, including classifying images or parts of images into different types of landcover (Yang and Newsam 2013) and even as different leaves or flowers (Nilsback 2009; Kumar et al. 2012; Kendal et al. 2013). Combining computer vision image feature research with UAS-SFM aerial remote sensing could lead to novel applications like automatically counting and mapping flower cover of tree crowns. This is diagrammed conceptually in Figure 1.

The tropical tree *Dipteryx panamensis* has bright pink flowers (Figure 1a) that are visible even when viewed from 150 m altitude by a camera on a UAS (Figure 1b). A group of *Dipteryx* flowers were identified on a crown as an image feature based potentially on distinct contrast with the green foliage (Figure 1c) and was matched across multiple images (Figure 1d) resulting in a pink 3D point on the crown within the whole point cloud (Figure 1e). Counting the number of such pink points within a single crown, for example following 3D-based crown segmentation (Fisher et al. 2013; Hu et al. 2014), may provide an estimate of the flower cover of the crown, an important but difficult to measure parameter of tropical canopies (Wright et al. 1999). Future research should explore such potential and also consider how the image features of points within a single crown ('crown bag of features') may be useful for discriminating between different genus

or species of canopy trees, similar to how time-series of per-crown spectral phenology signatures in Temperate Deciduous forests can be used for diversity mapping (Key et al. 2001).

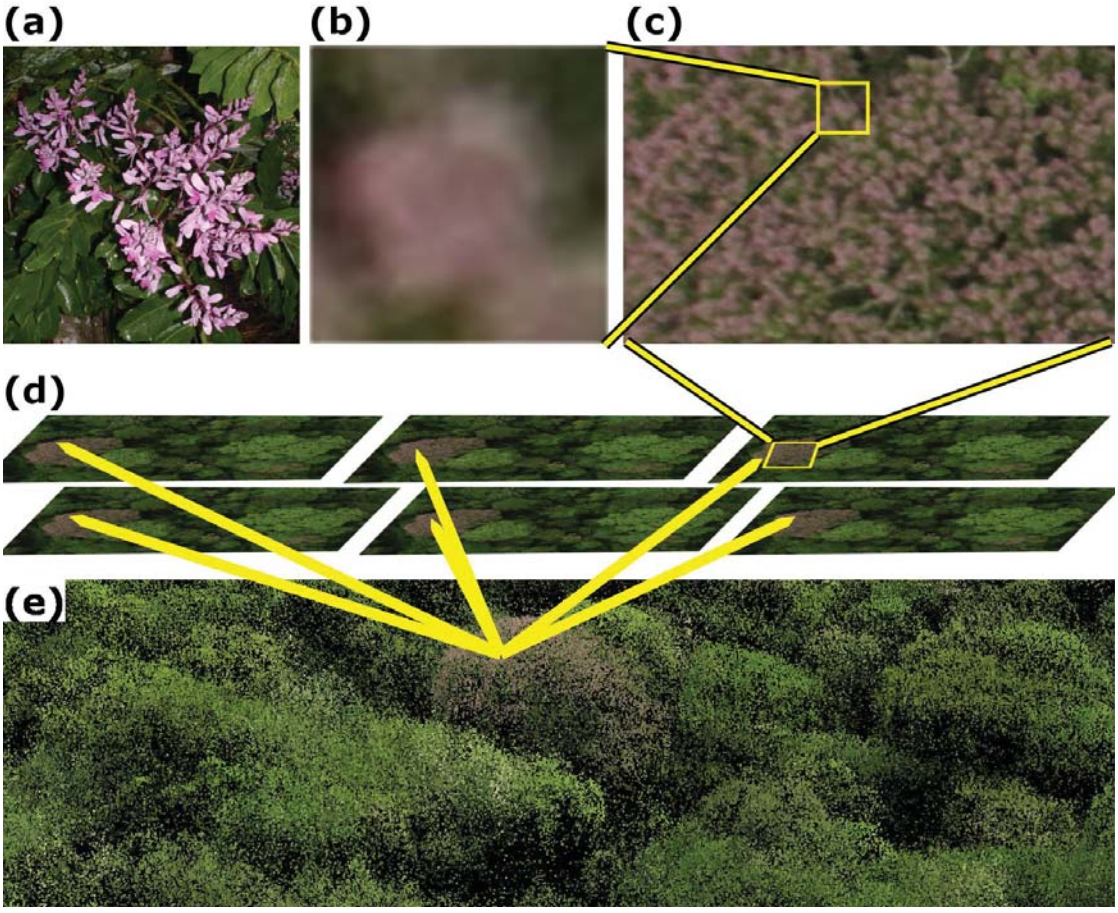


Figure 1: Conceptual diagram highlighting a single flower cluster from SFM point cloud image features. *Dipteryx panamaenis* flower cluster viewed on ground (≈ 1m, a), and by a UAS (≈ 150 m, b, c). Flower 'feature' was observed and matched across multiple images (d), and as a single point cloud point (e). Image (a) is for reference and is not the same as (b). UAS imagery and 3D point clouds from the 50 ha permanent plot on Barro Colorado Island (BCI), Panama, 2013-06-11, produced by J. P. Dandois, unpublished research, (a) courtesy Stephanie Bohlman, BCI, 2013-06-11.

While a computer vision image feature approach to ecological remote sensing presents many exciting research opportunities, future research should also consider the value of synthesis or summary research aimed at defining and describing this new approach within the context of existing remote sensing techniques. A useful example for

such synthesis research comes from the community of researchers engaged in geographic object-oriented image analysis (GEOBIA; Blaschke et al. 2014). Much like the current rise in UAS-SFM remote sensing, the area of object-oriented remote sensing was motivated by increased access to new hardware and software capabilities for remote sensing: very high resolution airborne and satellite imagery with pixel sizes smaller than the objects they represent (H-res; Strahler et al. 1986) and image processing software that enabled the creation of 'image-objects' or groups of image pixels with similar properties relative to the surroundings, also inspired by computer vision (Benz et al. 2004). Much as was done in this dissertation, research about GEOBIA has focused on describing what was being measured by image objects and how they related to geographic objects (Castilla and Hay 2008). GEOBIA research also focused on the definitions and vocabulary of object-oriented remote sensing, and such an exercise would also be useful for future research on ecological image feature remote sensing.

The term 'image features' or simply 'features' is well understood within the computer vision community, but readers from the fields of remote sensing or GIS (Geographic Information Systems) may associate different meanings to this term. Quackenbush (2004) uses the term 'features' and 'objects' interchangeably to describe shapes (points, lines, areas) that could be delineated from an image by a manual operator but instead are extracted automatically by computer algorithm. The GIS software ArcGIS (Environmental Systems Research Institute, Redlands, CA) uses the term 'features' to describe a vector-based geographic entity that provides the spatial representation of a real-world object and is linked to a database record. However, there are also similarities. Computer vision, remote sensing, and GIS all consider features to

have two main attributes: a shape or spatial characteristic and a descriptive or attribute characteristic. Synthesis of the similarities and differences between computer vision, remote sensing, and GIS 'features' would serve as a valuable point of cross-over to leverage the strengths of each discipline.

3.0 Concluding remarks

In conclusion, the rapid rise of two new technologies, autonomous, consumer unmanned aerial systems (UAS) and automated structure from motion image processing algorithms (SFM), opens the door to new ways to carry out research on forest canopies. The relatively low financial and logistical cost of combining these technologies enables remote sensing of forest canopy structural and spectral traits at spatial resolutions and frequencies not possible or practical with other forms of remote sensing, albeit at relatively small sampling extents (e.g., 10 - 50 ha). In addition, the use of SFM algorithms makes it possible to describe forest canopies based on computer vision image features, which offers the potential for providing a sematic or ecological classification along with measurements of 3D structure and color. Ultimately, this dissertation research should serve as a roadmap for this new area of ecological remote sensing, making it possible for future researchers to use computer vision to see the trees through the forest.

References:

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 58, 239-258.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., & Tiede, D. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 180-191.

Campbell, J.B. (1996). Introduction to Remote Sensing Second Edition. New York - London: Guilford Press.

Castilla, G., & Hay, G.J. (2008). Image objects and geographic objects. In T. Blaschke, S. Lang, & G. Hay (Eds.), Object-Based Image Analysis (pp. 91-110): Springer Berlin Heidelberg.

Dandois, J.P., & Ellis, E.C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sensing of Environment, 136, 259-276.

Dubayah, R.O., & Drake, J.B. (2000). Lidar Remote Sensing for Forestry. Journal of Forestry, 98, 44-46.

Fei-Fei, L., & Perona, P. (2005). A Bayesian hierarchical model for learning natural scene categories. In, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (pp. 524-531 vol. 522).

Fisher, J.T., Erasmus, B.F.N., Witkowski, E.T.F., van Aardt, J., Wessels, K.J., & Asner, G.P. (2013). Savanna woody vegetation classification – now in 3-D. Applied Vegetation Science, n/a-n/a.

Glennie, C. (2007). Rigorous 3D error analysis of kinematic scanning LIDAR systems. Journal of Applied Geodesy, 1, 147-157.

Holl, K.D., Zahawi, R.A., Cole, R.J., Ostertag, R., & Cordell, S. (2011). Planting Seedlings in Tree Islands Versus Plantations as a Large-Scale Tropical Forest Restoration Strategy. Restoration Ecology, 19, 470-479.

Hu, B., Li, J., Jing, L., & Judah, A. (2014). Improving the efficiency and accuracy of individual tree crown delineation from high-density LiDAR data. International Journal of Applied Earth Observation and Geoinformation, 26, 145-155.

Kendal, D., Hauser, C.E., Garrard, G.E., Jellinek, S., Giljohann, K.M., & Moore, J.L. (2013). Quantifying Plant Colour and Colour Difference as Perceived by Humans Using Digital Images. PLoS ONE, 8, e72296.

Key, T., Warner, T.A., McGraw, J.B., & Fajvan, M.A. (2001). A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. Remote Sensing of Environment, 75, 100-112.

Kumar, N., Belhumeur, P., Biswas, A., Jacobs, D., Kress, W.J., Lopez, I., & Soares, J.B. (2012). Leafsnap: A Computer Vision System for Automatic Plant Species Identification. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), Computer Vision – ECCV 2012 (pp. 502-516): Springer Berlin Heidelberg.

Lowe, D.G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60, 91-110.

Nilsback, M.-E. (2009). An automatic visual Flora - segmentation and classication of flower images. In, Robotics Research Group, Department of Engineering Science: University of Oxford.

Quackenbush, L.J. (2004). A Review of Techniques for Extracting Linear Features from Imagery. Photogrammetric Engineering & Remote Sensing, 70.

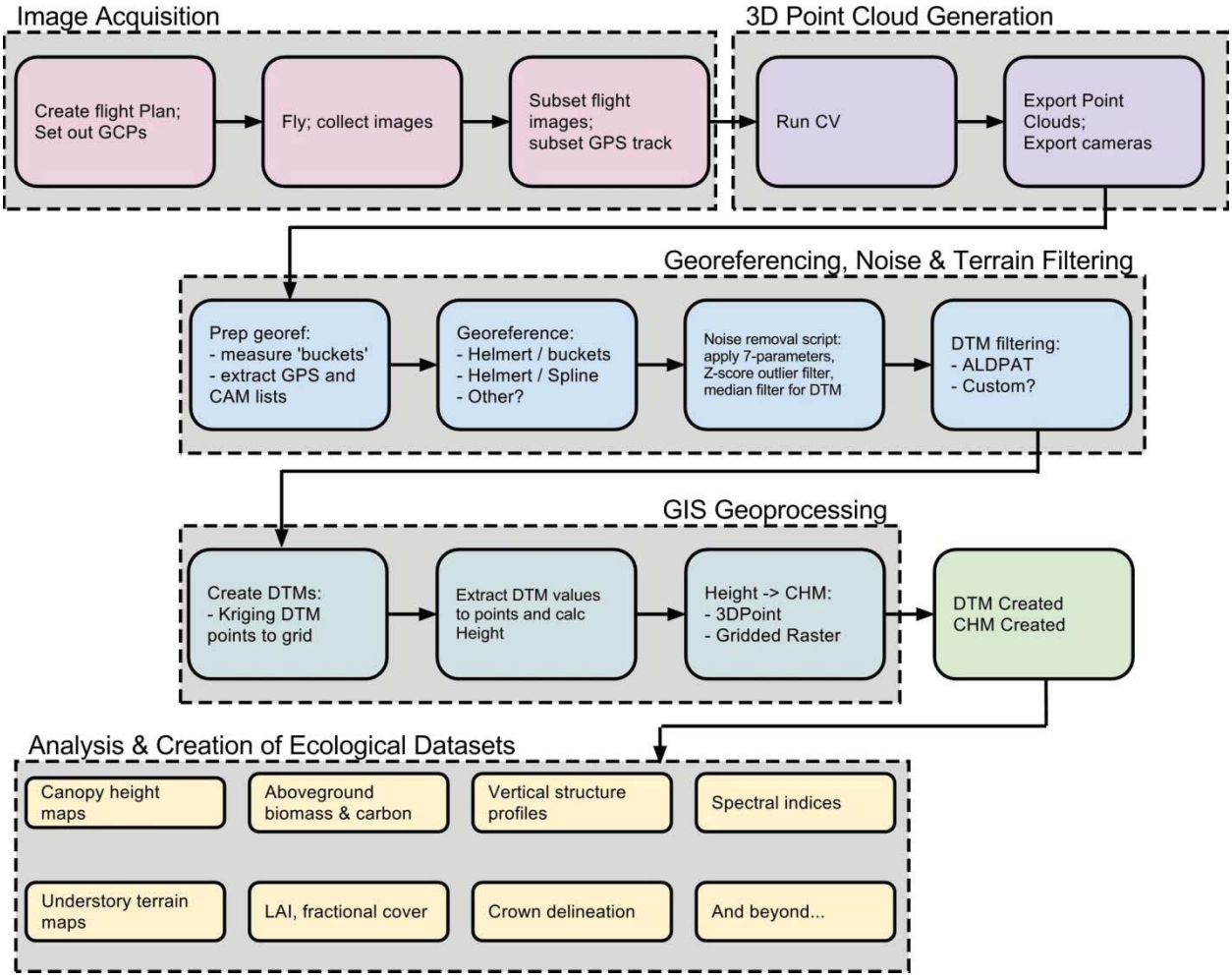
Yang, Y., & Newsam, S. (2013). Geographic Image Retrieval Using Local Invariant Features. Geoscience and Remote Sensing, IEEE Transactions on, 51, 818-832.

Zhang, X., Friedl, M., Schaaf, C., Strahler, A., Hodges, J., Gao, F., Reed, B., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. Remote Sensing of Environment, 84, 471-475.

Appendices

Appendix 1: Supplemental material for Chapter 2

A1.1: Workflow for computer vision (CV) remote sensing (Ecosynth).



A1.2: Image acquisition and computer vision data processing characteristics across study sites (6.25 ha) for Knoll and Herbert Run sites.

	Knoll			Herbert Run	
	Leaf-on 1	Leaf-on 2	Leaf-off	Leaf-on	Leaf-off
Flight date	10/8/10	10/24/10	3/11/11	10/6/10	3/5/11
Canopy condition	green	senescing	bare	green	bare
Weather	clear	overcast	partly cloudy	partly cloudy	overcast
Flight altitude (m) mean(SD) [min, max]	59 (10) [30,76]	67 (10) [39,84]	98 (10) [67,115]	61 (9) [34,77]	59 (9) [32,76]
Flight start time (local)	13:37	12:21	12:26	12:36	10:05
Flight duration (minutes)	≈ 12	≈ 12	≈ 16	≈ 11	≈ 13
Number of photos	1678	1869	2454	1808	2039
Computation time (hours)	51.3	106.3[a]	123.6	32.0	54.3
Point Cloud Statistics					
Total points generated	4,180,955	3,580,210	3,618,287	4,336,176	3,433,124
Points removed from study site by noise filter	39,312	30,873	26,606	42,822	44,790
Filtered points within study site	1,745,621	1,690,594	1,607,440	1,755,471	1,286,574
Point Density (points m^{-2}) mean(SD) [min, max]	37(57) [0,1030]	27(25) [0,222]	24(21) [0,207]	27(23) [0,251]	20(17) [0,199]

[a] Point cloud processed using earlier version of computer vision software (Agisoft Photoscan v0.7.0 build 651)

A1.2 (Continued): Image acquisition and computer vision data processing characteristics across study sites (6.25 ha) for SERC site.

	SERC					
	Winter	Spring 1	Spring 2	Summer	Fall 1	Fall 2
Flight date	2/26/12	4/9/11	4/29/11	6/28/12	10/8/11	11/13/10
Canopy condition	Bare leaf-off	Early spring	Spring green-up	Mature green	Early fall leaf-on	Fall senescing
Weather	clear	overcast	variable cloud & sun	clear	clear	clear
Flight altitude (m) mean(SD) [min, max]	72 (13) [55,120]	75 (7) [65,109]	72 (7) [58,105]	66 (8) [53,90]	40 (7) [27,74]	44 (8) [28,73]
Flight start time (local)	13:26	10:47	12:08	13:45	11:02	10:32
Flight duration (minutes)	≈ 12	≈ 12	≈ 18	≈ 13	≈ 14	≈ 12
Number of photos	1889	1842	2511	405 [b]	1952	1843
Computation time (hours)	39.6	32.8	39.5	1.5 [b]	27.3	16.2
Point Cloud Statistics						
Total points generated	2,439,999	2,401,380	7,417,442	1,795,192	5,500,149	2,168,235
Points removed from study site by noise filter	20,112	50,485	93,016	21,485	74,477	47,947
Filtered points within study site	1,476,371	1,473,194	5,021,651	1,752,542	4,167,055	1,500,547
Point Density (points m^{-2}) mean(SD) [min, max]	23(19) [0,1194]	12(13) [0,501]	40(26) [0,227]	49(34) [0,587]	67(60) [0,829]	14(16) [0,193]

[b] Short processing time and small number of photos due to camera shutting down early in flight, possibly due to high heat in excessive of 38° C at time of flight

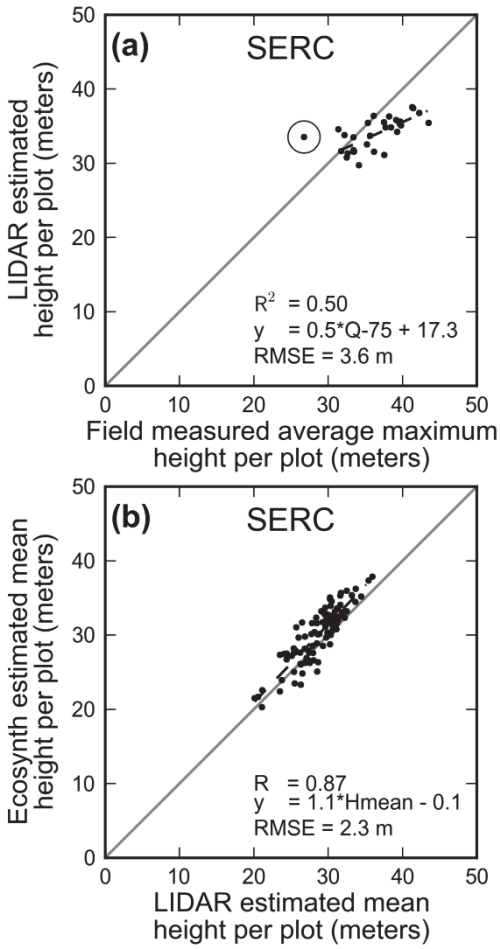
A1.3: Summary of field measured tree heights of the average of the 5 tallest trees per subplot (*AvgTop5*) at Knoll, Herbert Run, and SERC

	n (plots)	Mean (m)	Standard Deviation (m)	Coefficient of Variation (CV)	Range
Knoll	32	25.3	5.9	23%	11.5 – 36.4
Herbert Run	38	20.2	6.5	32%	8.6 – 35.5
SERC	29	36.6	3.4	9%	31.4 – 43.5

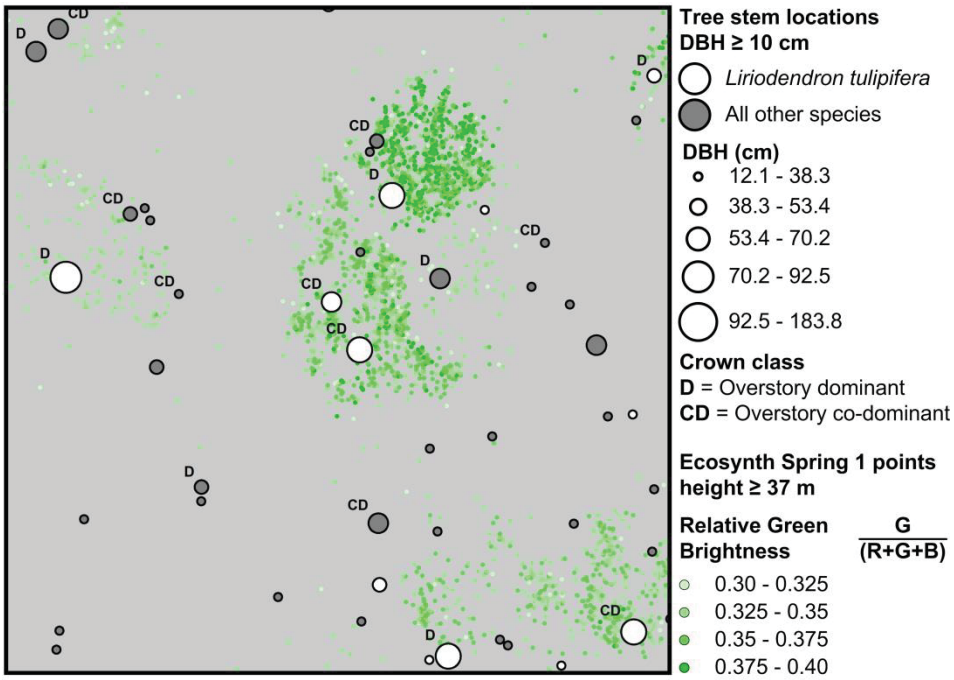
A1.4: Point cloud georeferencing error and accuracies across sites based on the spline method.

	Horizontal				Vertical	
	$RMSE_x$	$RMSE_y$	$RMSE_r$	95% Radial Accuracy	$RMSE_z$	95% Vertical Accuracy
Knoll						
Leaf On 1	2.47	3.33	4.14	7.17	2.30	4.49
Leaf Off	3.51	4.05	5.36	9.28	4.71	9.24
Herbert Run						
Leaf On	2.97	3.02	4.23	7.33	2.41	4.73
Leaf Off	2.43	2.55	3.53	6.10	1.71	3.35
SERC						
Leaf On[a]	-	-	-	-	-	-
Leaf Off	2.46	3.31	4.13	7.14	1.16	2.28

[a] Accuracy could not be assessed because canopy completely covered GCPs.



A1.5: LIDAR-estimated best linear regression predictor (Table 6) of field measured average maximum height per subplot (*AvgTop5;* a) and linear regression model of Ecosynth and LIDAR-estimated mean canopy height per plot (b) for the SERC site. Ecosynth and LIDAR canopy heights estimated using LIDAR DTM. Linear regression lines (dashed), R^2, linear models, and RMSE (m) are presented for each comparison, with solid gray reference lines along the one-to-one ratio. Circled data point in (a) is an outlier based on Grubb's test (> ±3 SD from the mean) and is not included in regression analysis.

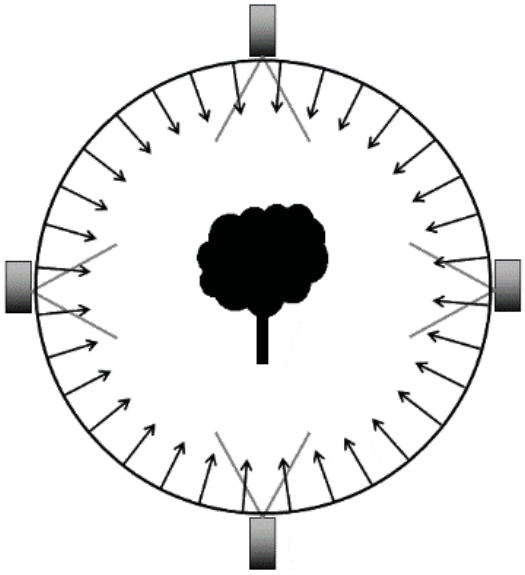


A1.6: Map showing close correspondence of Ecosynth time-series data with stem map locations of the dominant *Liriodendron tulipifera* canopy trees at a 50 m x 50 m sample area of the SERC site. Point cloud points from Ecosynth Spring 1 subset at SERC (Figure 7) with heights ≥ 37 m are shown in green colored by the point relative green brightness or 'strength of green' (i.e., $S_{green} = G/(R+G+B)$). Grey and white circles correspond to the tree stem locations (stems ≥ 10 cm DBH), with white circles indicating *Liriodendron tulipifera*, and grey circles representing all other tree species. Crown class is labeled for all dominant and co-dominant canopy trees. Map is shown on a grey background for visualization only. Blank areas of the map mean that there were no Ecosynth points in that area with heights ≥ 37 m in the Spring 1 dataset and no stems with DBH ≥ 10 cm within the stem map. Tree stems without a 'D' or 'CD' crown class label are of either an intermediate or suppressed class.

Appendix 2: Supplemental material for Chapter 3

A2.1

Diagram of computer vision structure from motion (SFM) and terrestrial LIDAR scanner (TLS) single tree scan configuration. Scanning was carried out around circle of 7 m radius from tree trunk center. TLS was positioned as indicated by gray boxes at north, east, south, west orthogonal positions around tree. SFM photos were taken 2.5° / 0.3 m intervals around circumference of scan area with camera centered on tree trunk.



A2.2

Table describing UAS aerial image datasets

Site	Knoll	Knoll	Knoll	Herbert Run	Herbert Run	Herbert Run	SERC	SERC
Flight date	10/08/10	10/24/10	3/11/11	10/06/10	8/24/12	3/5/11	10/08/11	2/26/12
Canopy condition	Early fall leaf-on	Mid fall, senescing	Leaf-off	Early fall leaf-on	Summer, green	Winter, Leaf off	Fall senescing	Winter, leaf off
Lighting	Clear	Diffuse, cloudy	Diffuse, partly cloudy	Partly cloudy	Clear	Diffuse, cloudy	Clear	Clear
Wind ($m\ s^{-1}$)	2.4	2.9	5.4	3.5	2.8	< 0.3	1.3	2.9
Average flight altitude above canopy (meters)	40	40	40	40	20	40	40	40
Number of photos	558	629	861	573	764	681	596	550
Computation time (hours)	97	27	85	85	552	52	356	105

A2.3

Detailed TLS data post-processing

TLS data were processed in the RIEGL RISCAN Pro software. The following processing tasks were carried out on each scan project (i.e., a set of four orthogonal scans) for both leaf-on and leaf-off scans. Red-green-blue (RGB) color is assigned to each point from photos collected by the DSLR camera mounted on top of the TLS using the 'Apply Color' tool. For optimal color fusion, the user must manually refine the calibration between the laser and camera optical centers by identifying matching tiepoints between an image of the laser scan and the color images at common locations, e.g., easily identifiable corners, reflective targets, etc. The software then automatically applies color from the images to each TLS laser point by using the refined calibration to map the color image pixel corresponding to the same location in the laser scan. Next, each set of four orthogonal TLS scans within a scan project were co-registered into the same 3D point cloud. During scanning in the field, an automated target searching function is used while the scanner is still set up at each scan location. Using this function, the scanner can automatically find, or be manually guided by user interaction, the location of the reflective targets placed within the scene. The scanner then repeatedly scans the targets at a very high resolution (< 0.001 m) to refine target location in 3D space. The refined target locations are then stored within the scan project for automated co-registration in the lab. The automated co-registration process attempts to align the scans together by determining the correct pairing of target locations that minimizes residual error across all target pairings based on rotating and translating the point cloud scans in 3D space relative to a single scan, i.e., applying rotation and translation factors along each X, Y, and Z axis. The point clouds are then merged into a common coordinate system based on the optimal rotation and translation factors. The merged RGB point cloud is then manually cropped by the user to include only the area of the single tree and exported as an XYZRGB ASCII text file.

The TLS point clouds were also filtered to include just the points corresponding to the red targets to facilitate further analysis. Filtering was carried out using built in tools in RiScan Pro and was based on identifying the range of returned laser energy that best corresponded to the targets. Prior research has found that the amplitude values of TLS scans can be used to classify point clouds based on the type of material that is scanned, e.g., green leaf material vs. bark and branches. This was done by first manually segmenting the point cloud to only points of a single target and calculating the range of amplitudes of reflected energy for that target. This amplitude range was then used as the filtering threshold for segmenting targets from the entire point cloud. Segmented point clouds of targets only were then exported as XYZRGB ASCII text files.

A2.4

Meshlab ICP point cloud registration methods

TLS and SFM point clouds were manually co-registered into the same coordinate system using the following method implemented in the open source 3D software Meshlab (v1.3.3 64-bit available online: http://meshlab.sourceforge.net/, downloaded 2014-04-02). All work on point clouds was carried out in the Stanford PLY file format. In practice, all models converged to a minimum AvgError within at most 2 runs of the Iterative Closest Point (ICP) algorithm, i.e., additional runs did not improve the estimate of AvgError. For leaf-off models, the mean and SD AvgErr were 0.0049 m and 0.0005 m, with a median of 0.0048 m. For leaf-on models the mean and SD AvgErr were 0.0112 m and 0.0007 m, with a median of 0.0111 m.

1. First, point clouds were manually trimmed to contain only points of the tree itself, from the base of the tree out to all branches. Some singular stray noise points persisted in both TLS and SFM models and were removed by first computing the radius or distance around each point using the 'Estimate Radius from Density' function based on the 16 nearest neighbors and removing any points with a radius larger than the 99^{th} percentile of radius values for the whole point cloud (i.e., removing 1% of the points with the lowest density or greatest distance to another point).
2. The 'Measure' tool was then used to measure distances between common points within each point cloud in order to estimate a scaling factor from SFM to TLS point clouds. The point clouds were manually rotated in two adjacent windows in Meshlab to approximately the same point of view. Then, similar points in each model were visually identified, e.g., the tips of prominent branches or the location of red Styrofoam ball targets, and the distances between these points in each point cloud were recorded.
3. Distance between the same points was measured in the TLS and SFM point clouds. Six pair distances were identified for each SFM and TLS point cloud with the same pairs identified for point clouds by both methods. A scale ratio was computed for each distance pair (D_{TLS} / D_{SFM}) and the mean scale ratio applied to each of the XYZ values of the coordinates of each corresponding SFM point cloud [$XYZ_{scaled} = (H_{TLS} / H_{CVSFM}) \times XYZ_{SFM}$] in Excel or by command line in Python using the Numpy module (v1.8.2, http://www.numpy.org/) when the point cloud was too large to be loaded into Excel.
4. Next, the SFM point cloud was roughly manually aligned to the TLS point cloud using the 'Point Based Glueing' interface within the 'Align' tool. This interface prompts the user to identify matching points between the two models, which it then uses to rotate and translate one model relative to the other. At least 4 tiepoints were identified, typically at the ends of prominent branches. Following 'Point Based Glueing', the point clouds were aligned using the Iterative Closest Point (ICP) alignment to refine the rotation and translation of the SFM point cloud to the TLS point cloud. The ICP algorithm works by taking a random sample of points from both models and attempts to minimize the distance between point pairs in each model.

This is repeated with new point sets until the minimized distance achieves a user specified threshold. At that point, the last rotation and translation values are applied to the entire point cloud. A threshold of 0.005 m (5 mm) was used here for aligning the TLS and SFM point clouds as it was found through repeated iterations that the models could not be aligned with greater precision. The new rotation and translation values were then fixed to the SFM point cloud, which was then exported to a new XYZRGB ASCII text file. This process was repeated for aligning each leaf-on and leaf-off SFM model to the corresponding TLS model for the same season. An exact workflow was developed for this task:

a. First import the TLS model into a new project in Meshlab, then import the scaled & filtered SFM model.
b. Start the 'Alignment' tool.
c. Highlight the TLS model from the list of models in the Alignment tool window and select 'Glue Here Mesh' and 'Set as Base Mesh' from the buttons.
d. Highlight the SFM model from the list of models.
e. Select 'Point Based Glueing'
f. Select tie points of approximately matching points/features in each model:
    i. At least 4 points are required.
    ii. Leave 'Allow Scaling' disabled/unchecked because a scaling factor has already been applied.
    iii. Double click a point in one model, then double click the 'same' point in the other model. Order does not matter. To remove a point hold Control and double-click.
    iv. Click 'OK' when finished.
g. Close the 'Align Tool' window and open the View > Show Layer Dialog from the main Meshlab window.
h. Remove the TLS model from the current project by right clicking on its name in the Layer window and selecting 'Delete Current Mesh.'
i. Right-click on the SFM model and select 'Freeze Current Matrix'. The 'Point Based Glueing' routine created a 4x4 rotation and translation matrix to match the alignment provided by the tiepoints. 'Freezing' the matrix applies the rotation and translation values to the point cloud.
j. Export the SFM point cloud to a new file: File > Export Mesh As. Provide a new file name and use the PLY format. When prompted, choose to include 'Colors' and 'Radius' and disable/uncheck 'Binary Encoding'. The SFM point cloud has now been scaled to match the TLS model and provided with an approximate translation and rotation to be roughly aligned with the TLS model.
k. Close the current project and open a new project in Meshlab.
l. Import the TLS model first and by itself, then import the pre-aligned SFM model from the previous step separately.
m. Open the 'Align' tool.
n. Highlight the TLS model and select 'Glue Here Mesh' and 'Set as Base Mesh'.
o. Highlight the SFM model and select 'Glue Here Mesh'.
p. Open the 'Default ICP Params' dialog and set the following parameters:
    i. Sample number: 1000 (Default)

1. Practice found no improvement in increasing this number.

ii. Minimal Starting Distance: 0.3
    1. As in 0.3 m or 30 cm and represents the minimum distance between a set of randomly selected point pairs to begin the ICP alignment.

iii. Target Distance: 0.005
    1. As in 0.005 m or 5 mm. Practice found that this is a minimum value of this parameter for this application beyond which no improvements in the measure of average error are obtained.

iv. Max iteration: 50
    1. Practice found no improvement by increasing this number.

v. Normalized Equal Sampling: ON

vi. MSD Reduce Factor: 0.8 (Default)

vii. Sample Cut High: 0.75 (Default)

viii. Rigid Matching: ON (we do not want ICP to apply a scaling factor)

ix. These parameters were determined by running multiple iterations across multiple point cloud sets and observing the visual quality of the alignment and the value of 'AvgErr' reported by the program which provides an estimate of the mean difference between a set of randomly selected point pairs in the same 3D region of space. It is not truly a measure of alignment error or accuracy, but merely a measure of fit between the two models.

q. Click 'Process' to perform the ICP alignment. In practice, this approach typically converged to the 'Target Distance' value after one iteration of ICP alignment. Further runs of alignment did not show improvement in the AvgErr parameter.

r. 'Freeze' the alignment matrix and export the SFM point cloud as in (h) – (j) above.

5. The above process was used first to align the leaf-off SFM point clouds to the leaf-off TLS point cloud. To align leaf-on and leaf-off point clouds, the TLS point clouds for the two time periods were first aligned to each other, leaf-on to leaf-off, then the leaf-on SFM point clouds were aligned to the leaf-on TLS point cloud, putting all point clouds in the same coordinate system as the TLS leaf-off point clouds. Co-registration of the leaf-on SFM and TLS models used the same process as above, but co-registration of the leaf-on and leaf-off TLS models required a slight modification of the process, as described below.

    a. ICP alignment was unable to correctly rotate and translate the two leaf-on and leaf-off TLS models due the large differences in the crown 3D geometry between the two scans.

    b. To address this, the two models were first cropped to contain points only of the main tree trunk (≈ 1.3 m above ground) and approximately 0.2 m of the lowest branches that are not obstructed by leaves in both scans. Visual inspection determined that this part of the tree was the most similar across the scans taken in different seasons.

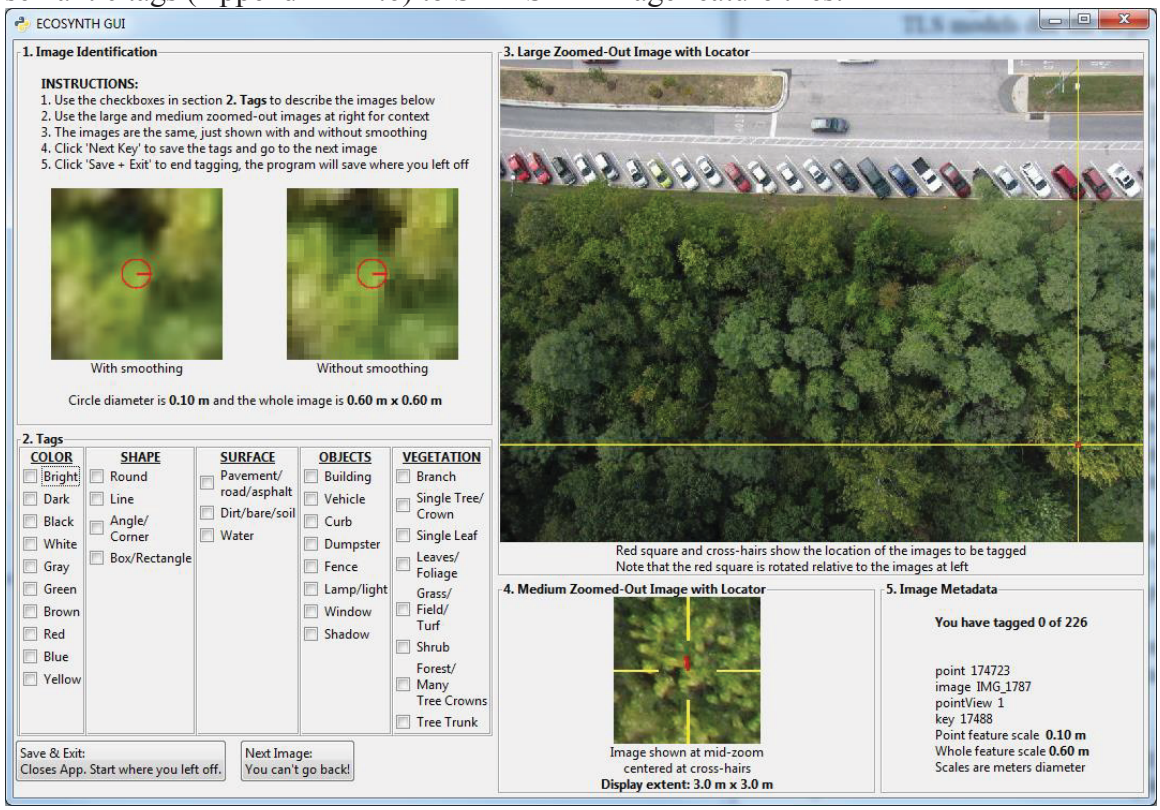
    c. 'Point Based Glueing' was first used to roughly align the point clouds together. Then, the 'Manual Rough Glue' function was used to manually rotate and

translate the leaf-on model along all three axes to be approximately visually aligned with the leaf-off.

d. With the models roughly visually aligned, the ICP 'Process' was run until there was no discernible change in the alignment of the models based on visual inspection. The resulting rotation and translation matrix for the leaf-on TLS point cloud of the tree trunk was then exported as a plain-text alignment file (ALN) and applied to the entire leaf-on model by changing the referenced file name for the rotation and translation matrix in the ALN file , re-loading the ALN file as a project, which automatically loads and applies the matrix to the full leaf-on TLS point cloud, and then following (4h) – (4j) above to 'Freeze' and export the aligned point cloud.

A2.5

Example screenshot of custom graphical user interface developed to help users provide semantic tags (Appendix A 2.6) to SFM SIFT image feature tiles.



A2.6

Set of common tags spanning five descriptive categories (color, shape, surface, vegetation objects, and other objects provided to users for tagging image feature tiles in custom graphical user interface (Appendix A 2.5).

COLOR	SHAPE	SURFACE	OBJECT	VEG OBJECTS
Bright	Round	Pavement / road / asphalt	Building	Branch
Dark	Line	Dirt / bare / soil	Vehicle	Single Tree / Crown
Black	Angle/Corner	Water	Curb	Single Leaf
Gray	Box/Rectangle		Dumpster	Leaves/Foliage
Green			Fence	Grass / field/ turf
Brown			Lamp/light	Shrub
Red			Window	Forest/ Many/ Tree Crowns
White			Shadow	Tree Trunk
Blue				
Yellow				

A2.7

Table showing breakdown of classification of points per target and per replicated for SFM and TLS single tree scans under leaf-on and leaf-off conditions. Values show proportion of points located within the target area for each replicate that fell within the red hue cutoff region (330 – 20). Blank cells ('—') indicate that no points were observed at the target area. Cells with value of 0.00 had > 1 point within the target area, but no points were within the red cutoff region. Averages, for which classification accuracy is calculated on, consider blank values ('—') as 0. Values of 'Average Red' indicate the average proportion of points located at a target that were red. Values greater than 0.5 are highlighted in bold face and indicate that a target was 'accurately classified'.

SFM Leaf-on point clouds

SFM Leaf-on	Target										
SFM replicate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0.33	--	--	--	--	0.00	--	--	0.00	--	1.00
2	--	0.67	0.75	--	--	--	--	--	0.18	--	0.50
3	--	0.59	--	--	0.33	--	--	--	0.23	--	0.17
4	0.00	0.50	1.00	1.00	1.00	0.44	--	--	0.40	--	1.00
5	0.00	0.86	0.86	--	0.60	--	0.00	--	0.25	--	0.75
6	0.13	0.67	--	0.00	0.40	0.50	0.00	--	0.10	--	1.00
7	0.00	0.83	0.33	0.00	--	--	--	--	0.25	--	0.29
8	0.25	1.00	1.00	0.00	1.00	0.00	--	--	--	--	0.50
9	0.00	0.60	1.00	0.00	--	0.50	0.00	--	0.20	--	0.00
10	0.30	0.71	0.33	0.00	0.00	0.67	--	--	0.40	--	1.00
Average Red	0.08	**0.64**	**0.55**	0.11	0.37	0.16	0.00	0.00	0.18	0.00	**0.58**
Had points (0.59)	0.80	0.90	0.70	0.60	0.60	0.60	0.40	0.00	0.90	0.00	1.00

TLS leaf-on sub-sampled point clouds

TLS Leaf-on	Target										
TLS replicate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0.25	0.09	0.89	1.00	0.33	1.00	0.36	1.00	0.00	0.27	0.14
2	0.00	0.13	0.36	1.00	0.48	0.67	0.69	--	0.00	0.36	0.40
3	0.00	0.23	0.55	1.00	0.63	1.00	0.50	1.00	0.00	0.35	0.00
4	0.00	0.11	0.56	1.00	0.36	0.86	0.50	--	0.00	0.35	0.13
5	0.40	0.29	0.50	1.00	0.63	0.60	1.00	--	0.00	0.37	0.13
6	0.17	0.00	0.88	0.67	0.62	0.75	0.70	--	0.00	0.27	0.17
7	0.38	0.00	0.83	1.00	0.58	0.71	0.55	1.00	0.00	0.46	0.25
8	0.50	0.00	0.53	1.00	0.39	0.67	0.33	0.00	0.00	0.45	0.17
9	0.14	0.25	0.88	1.00	0.64	0.75	0.65	--	0.00	0.29	0.33
10	0.29	0.00	0.80	1.00	0.61	0.70	0.65	1.00	0.00	0.46	0.20
Average Red	0.20	0.12	**0.66**	**0.96**	**0.52**	**0.78**	**0.59**	0.00	0.00	0.00	0.19
had points (0.95)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.50	1.00	1.00	1.00

A2.7 (continued)

SFM Leaf-off point clouds

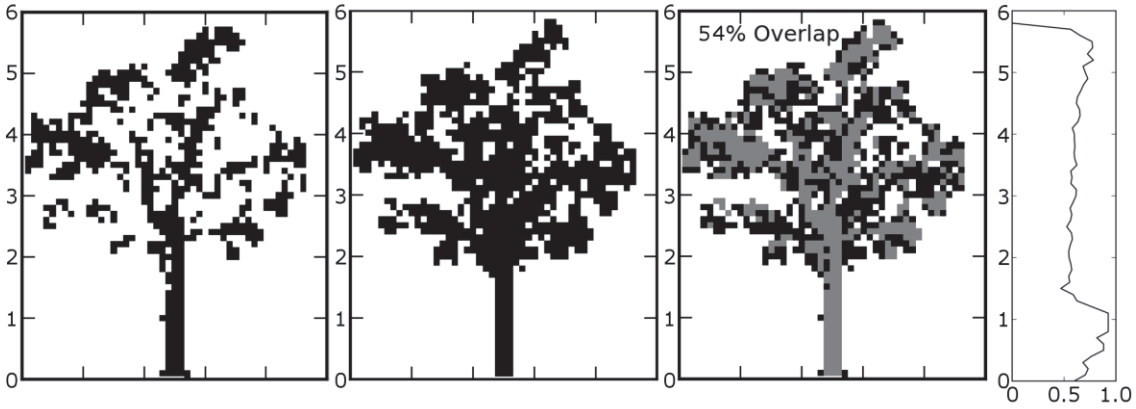
SFM Leaf-off	Target													
SFM replicate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	0.65	0.83	1.00	0.77	1.00	0.67	0.79	0.90	0.39	0.64	0.50	0.71	0.38	0.79
2	0.84	0.93	0.89	0.90	1.00	0.67	0.75	0.80	0.44	0.64	0.55	0.60	0.47	0.76
3	0.64	0.85	0.94	0.68	1.00	0.25	0.67	0.86	0.33	0.32	0.60	0.63	0.50	0.86
4	0.77	0.79	0.95	0.84	1.00	0.62	0.83	0.91	0.46	0.78	0.53	0.61	0.38	0.83
5	0.88	0.78	0.90	0.71	1.00	0.73	0.67	0.90	0.46	0.73	0.72	0.60	0.57	0.67
6	0.67	0.78	0.70	0.61	1.00	0.58	0.88	0.86	0.46	0.52	0.71	0.60	0.65	0.75
7	0.76	0.93	0.82	0.73	1.00	0.71	1.00	0.92	0.36	0.74	0.55	0.70	0.45	0.80
8	0.70	0.77	0.86	0.69	1.00	0.59	0.80	0.90	0.36	0.86	0.48	0.66	0.51	0.70
9	0.79	0.75	0.80	0.84	1.00	0.73	0.83	0.85	0.36	0.67	0.72	0.61	0.63	0.52
10	0.66	0.65	0.93	0.82	1.00	0.89	0.80	1.00	0.30	0.68	0.56	0.50	0.54	0.83
Average Red	**0.74**	**0.81**	**0.88**	**0.76**	**1.00**	**0.64**	**0.80**	**0.89**	0.39	**0.66**	**0.59**	**0.62**	**0.51**	**0.75**
Had points (1.0)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

TLS Leaf-off point clouds

TLS Leaf-off	Target													
TLS replicate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	0.51	0.11	0.59	0.43	1.00	0.26	0.57	0.16	0.06	0.32	0.16	0.31	0.40	0.51
2	0.61	0.00	0.52	0.45	0.84	0.40	0.67	0.14	0.04	0.29	0.13	0.28	0.34	0.53
3	0.42	0.11	0.53	0.35	0.82	0.29	0.42	0.11	0.09	0.25	0.29	0.24	0.43	0.51
4	0.51	0.07	0.25	0.38	0.90	0.43	0.65	0.12	0.10	0.18	0.17	0.32	0.39	0.47
5	0.59	0.18	0.33	0.44	0.82	0.26	0.61	0.08	0.09	0.14	0.27	0.20	0.47	0.49
6	0.54	0.11	0.45	0.38	0.87	0.30	0.70	0.21	0.13	0.17	0.37	0.27	0.43	0.50
7	0.60	0.08	0.40	0.40	0.88	0.25	0.53	0.08	0.05	0.28	0.28	0.31	0.39	0.57
8	0.48	0.07	0.41	0.40	0.95	0.25	0.56	0.11	0.12	0.21	0.29	0.30	0.49	0.41
9	0.56	0.11	0.35	0.37	0.76	0.43	0.62	0.16	0.00	0.35	0.17	0.27	0.46	0.51
10	0.48	0.12	0.43	0.49	0.85	0.28	0.68	0.20	0.04	0.14	0.26	0.26	0.51	0.42
Average Red	**0.53**	0.10	0.43	0.41	**0.87**	0.32	**0.60**	0.14	0.07	0.23	0.24	0.28	0.43	**0.50**
Had points (1.0)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

A2.8

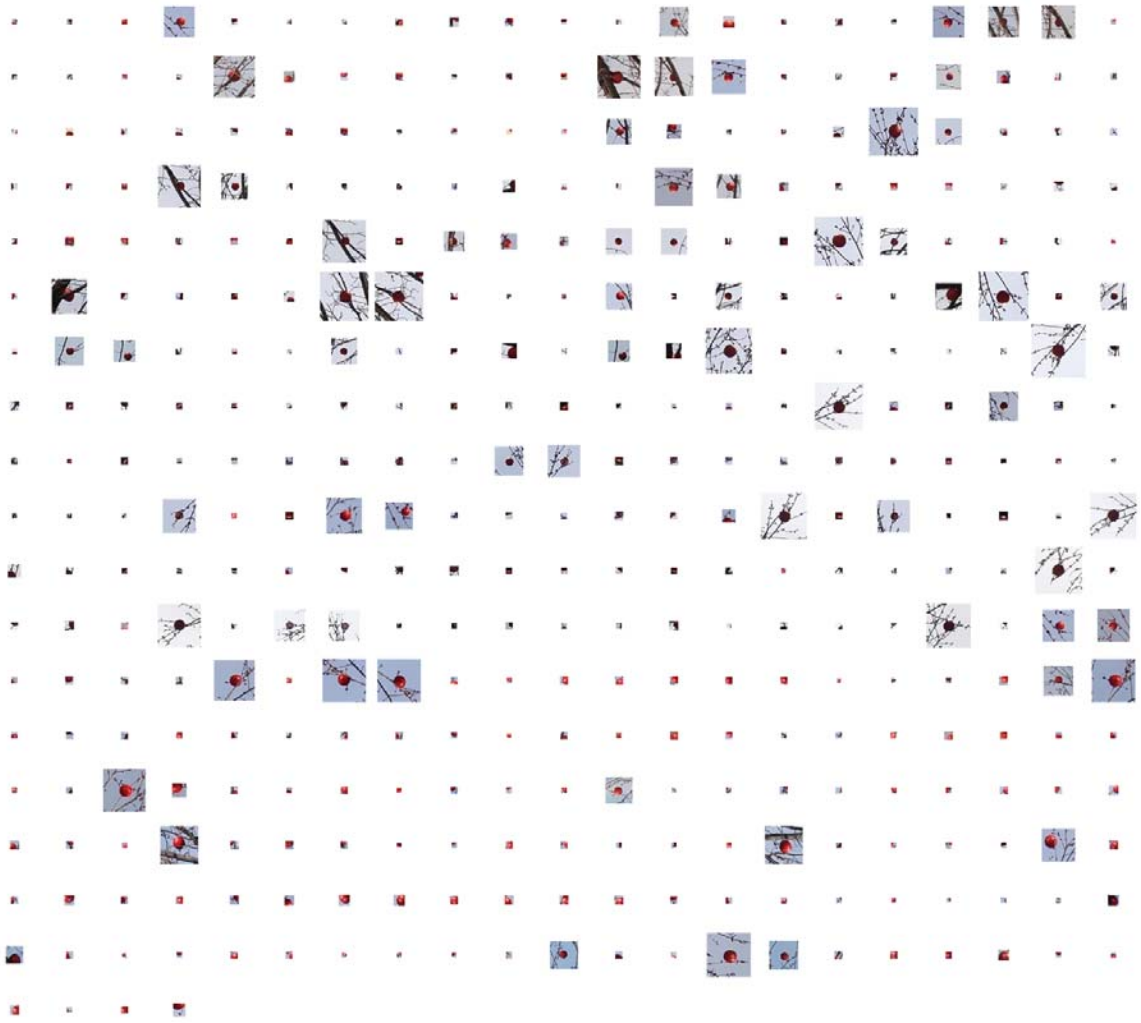
Tomographic slices of a 0.1 m (10 cm) section of combined leaf-on and leaf-off SFM and TLS point clouds centered at roughly the center of the tree stem. Pixels represent a 0.1 m x 0.1 m x 0.1 m cube (0.001 m^3; 1,000 cm^3) within which at least 1 point cloud point was located. All slices are co-registered to the same coordinate system (scale, rotation and translation) such that pixels are the same in each model.



A2.9

Samples of SFM image feature tiles from leaf-on and leaf-off single tree point cloud scans highlighting differences in the background around each feature that may contribute to differences in point vs. image feature tile RGB channel correlation and ratio. All SFM image feature tiles associated located within target areas for a single leaf-on replicate shown at actual relative scale, n = 55.

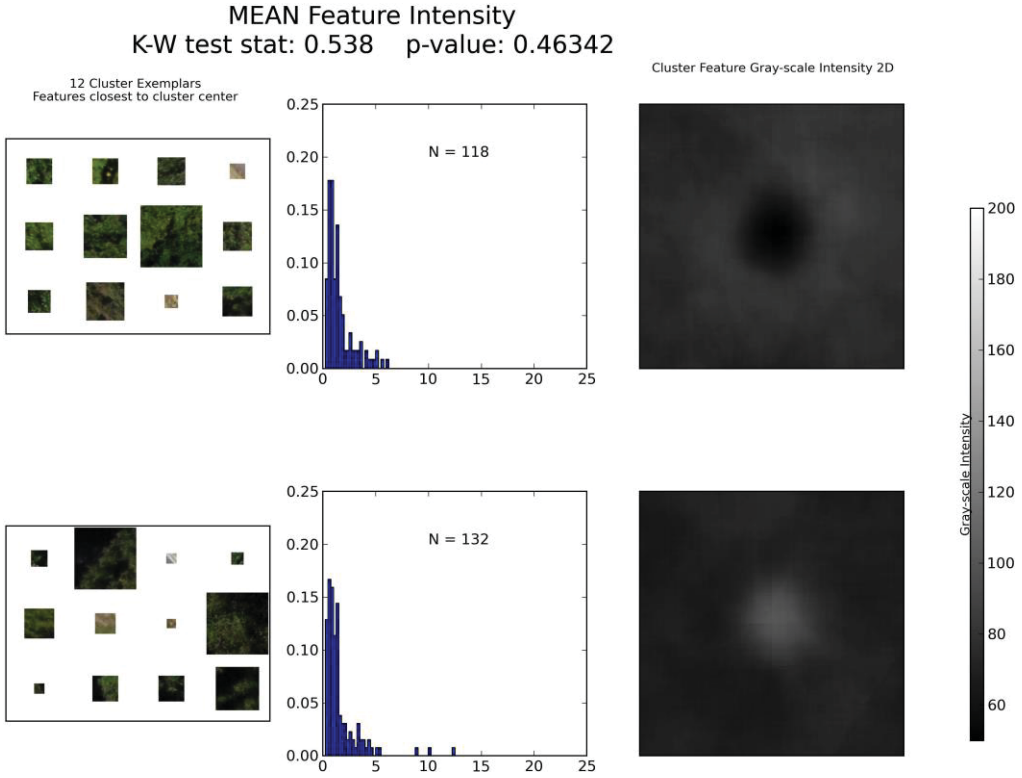
A2.9 (continued)
All SFM image feature tiles associated located within target areas for a single leaf-off replicate shown at actual relative scale, n = 382.



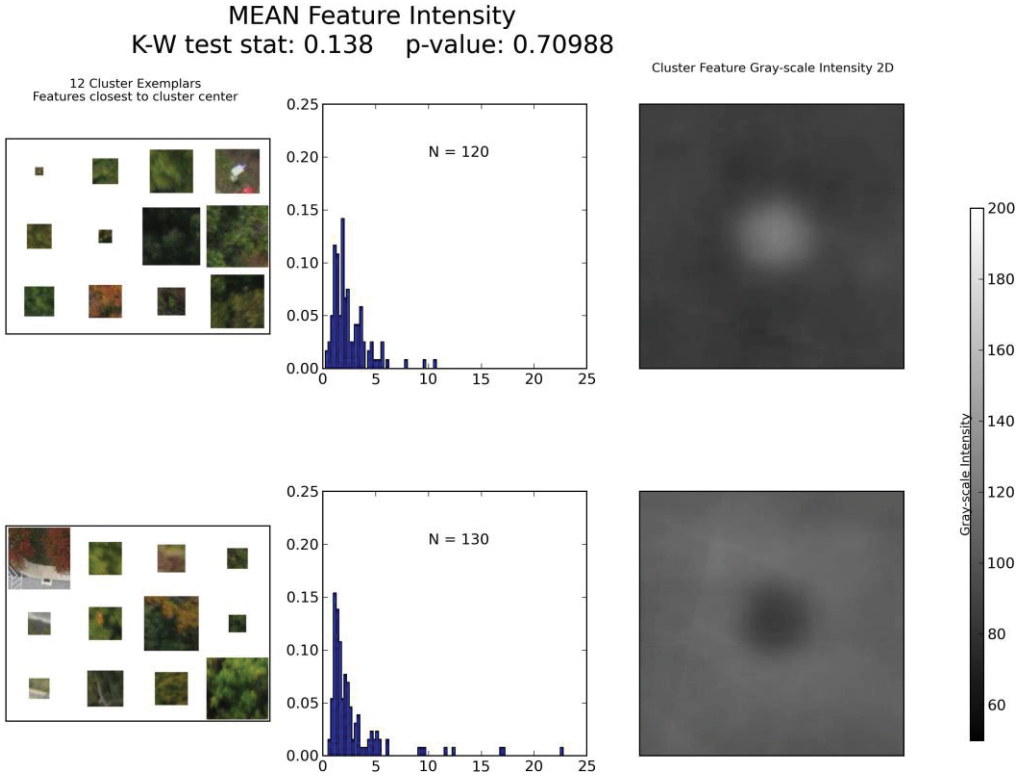
A2.10

Clustering results on SIFT 128D numeric descriptors for all datasets (Appendix A 2.2) and on a sample of points from leaf-on and leaf-off single tree datasets.

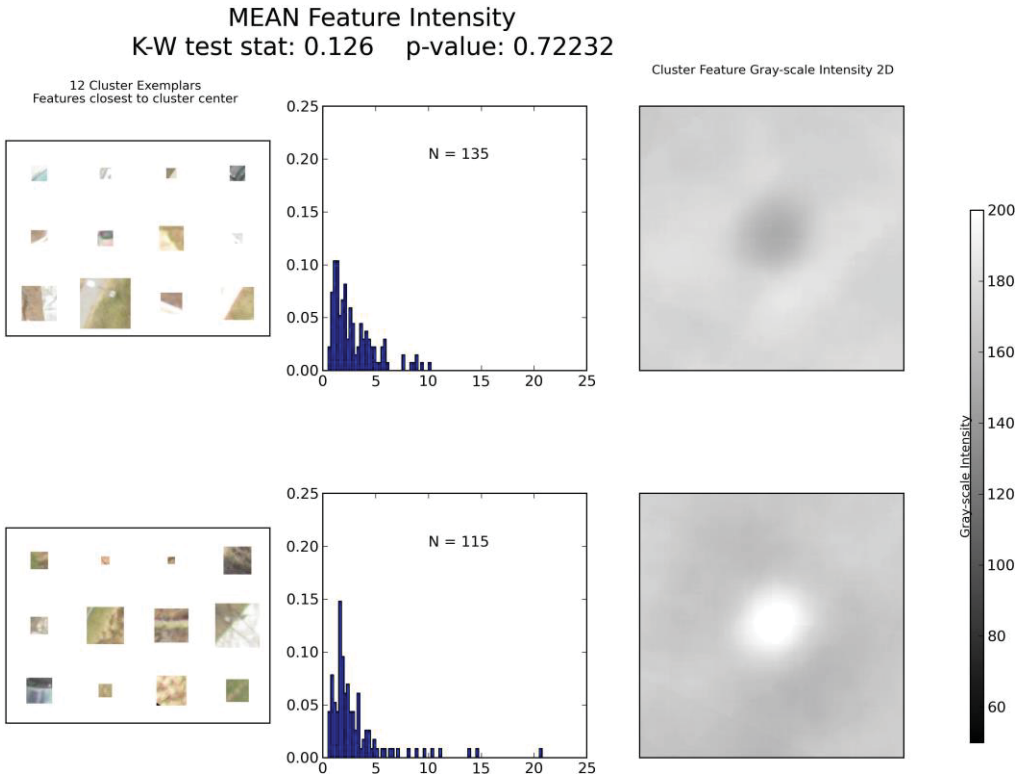
Knoll 10/08/10



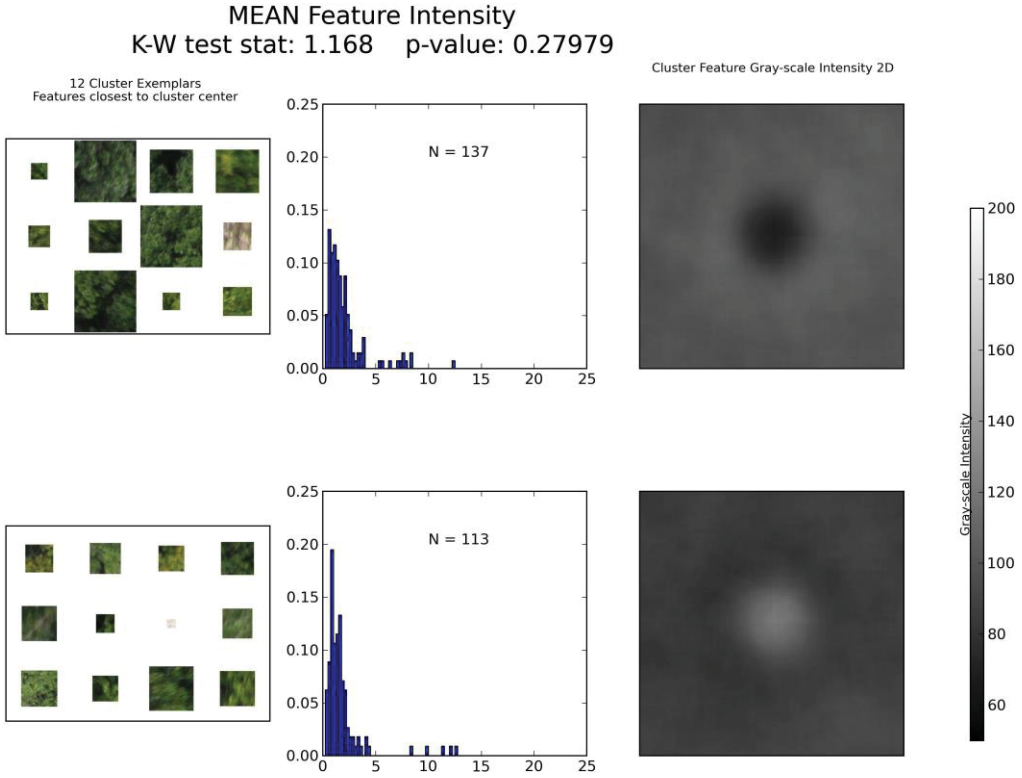
Knoll 10/24/10



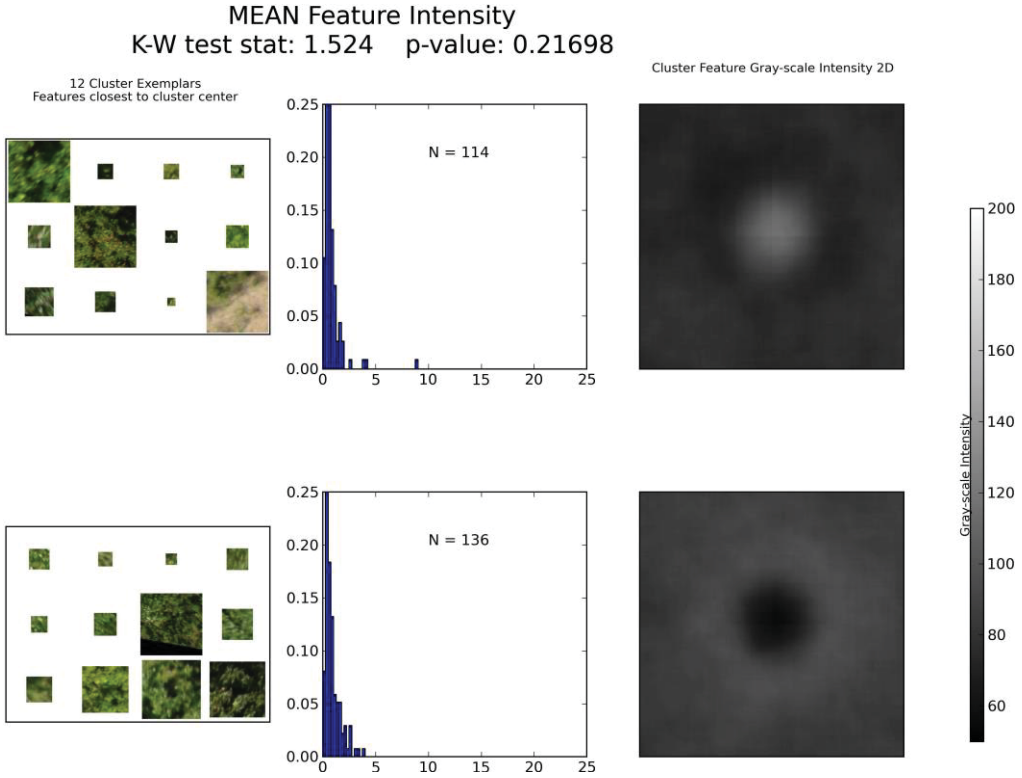
Knoll 3/11/11



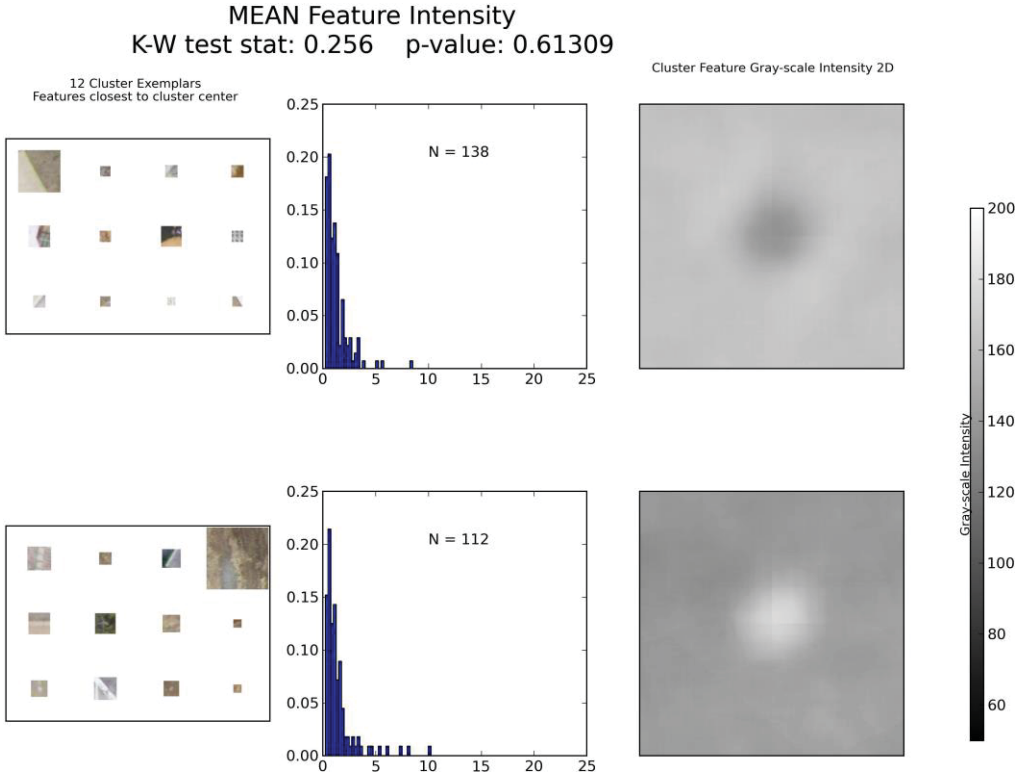
Herbert Run 10/06/10



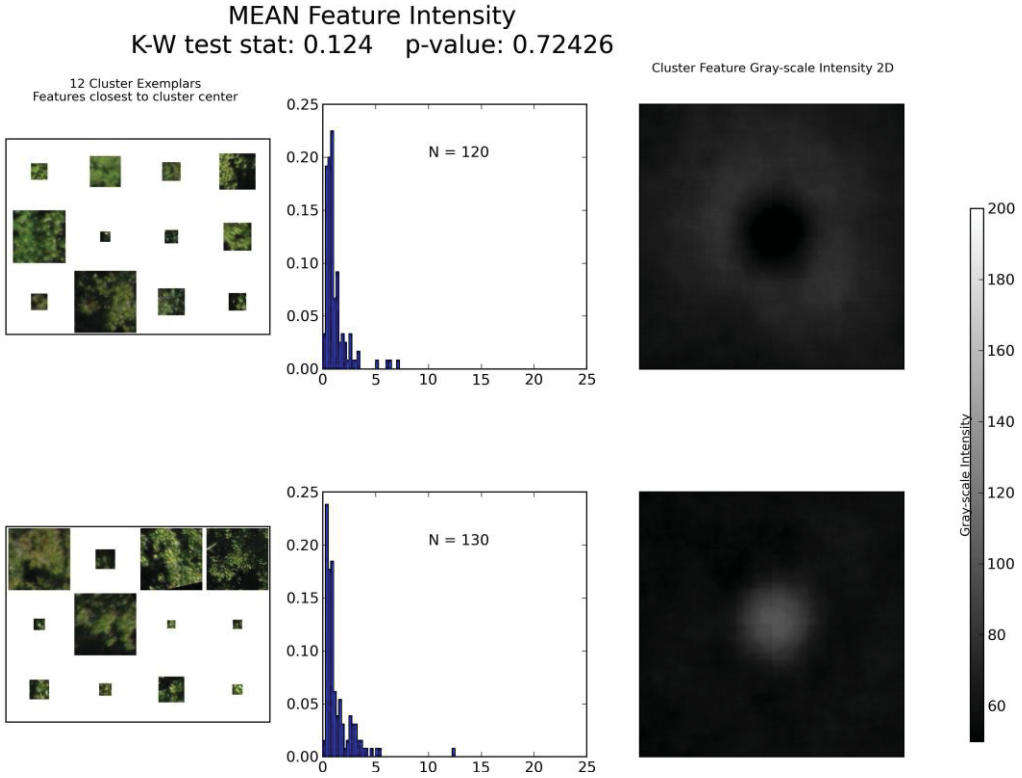
Herbert Run 8/24/12



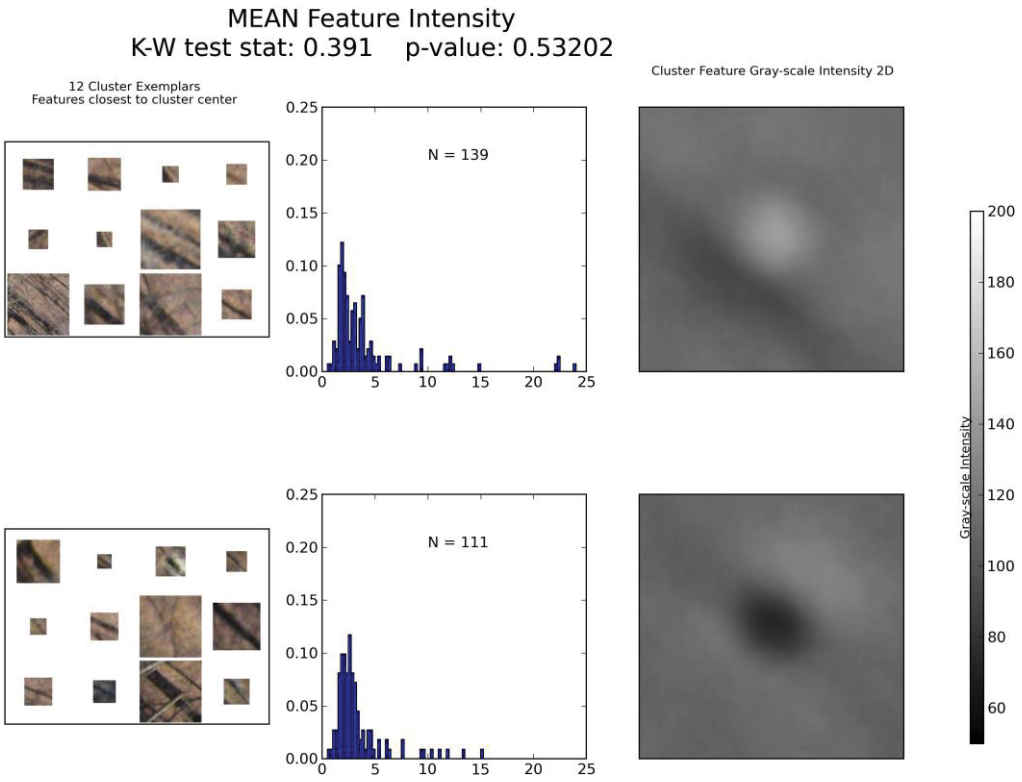
Herbert Run 03/05/11



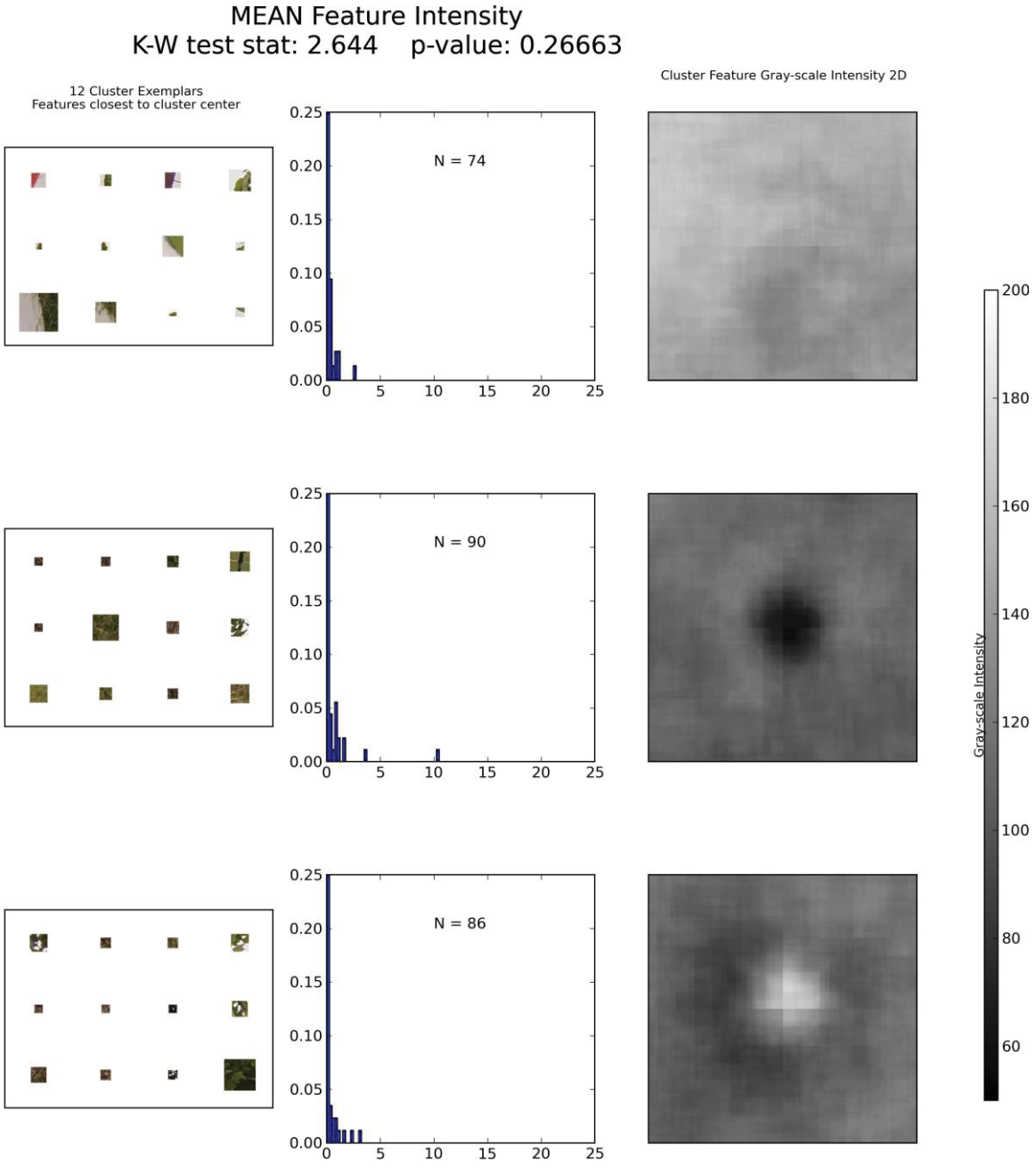
SERC 10/08/11



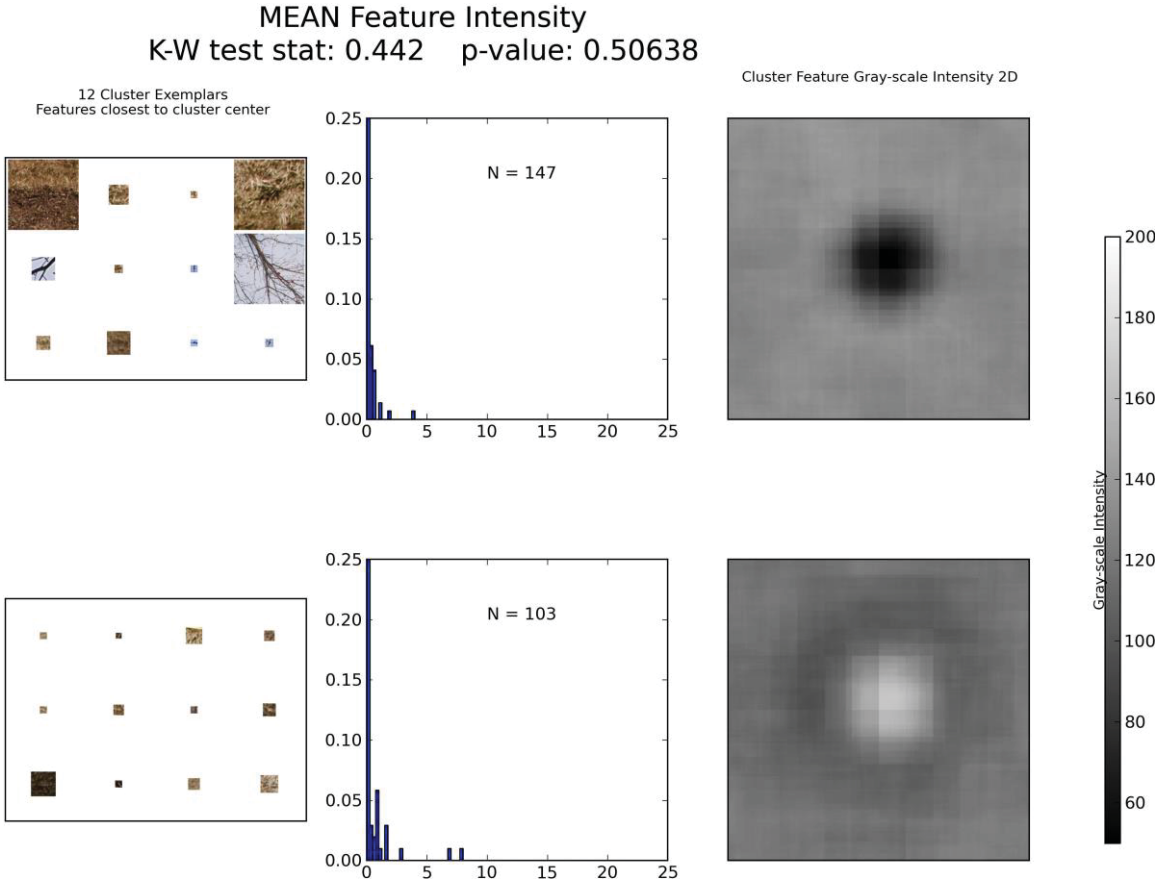
SERC 02/26/12


MEAN Feature Intensity
K-W test stat: 0.391 p-value: 0.53202
12 Cluster Exemplars
Features closest to cluster center
N = 139
N = 111
Cluster Feature Gray-scale Intensity 2D
Gray-scale Intensity
200
180
160
140
120
100
80
60
0.25
0.20
0.15
0.10
0.05
0.00
0
5
10
15
20
25

Single Tree Leaf-on 8/20/12

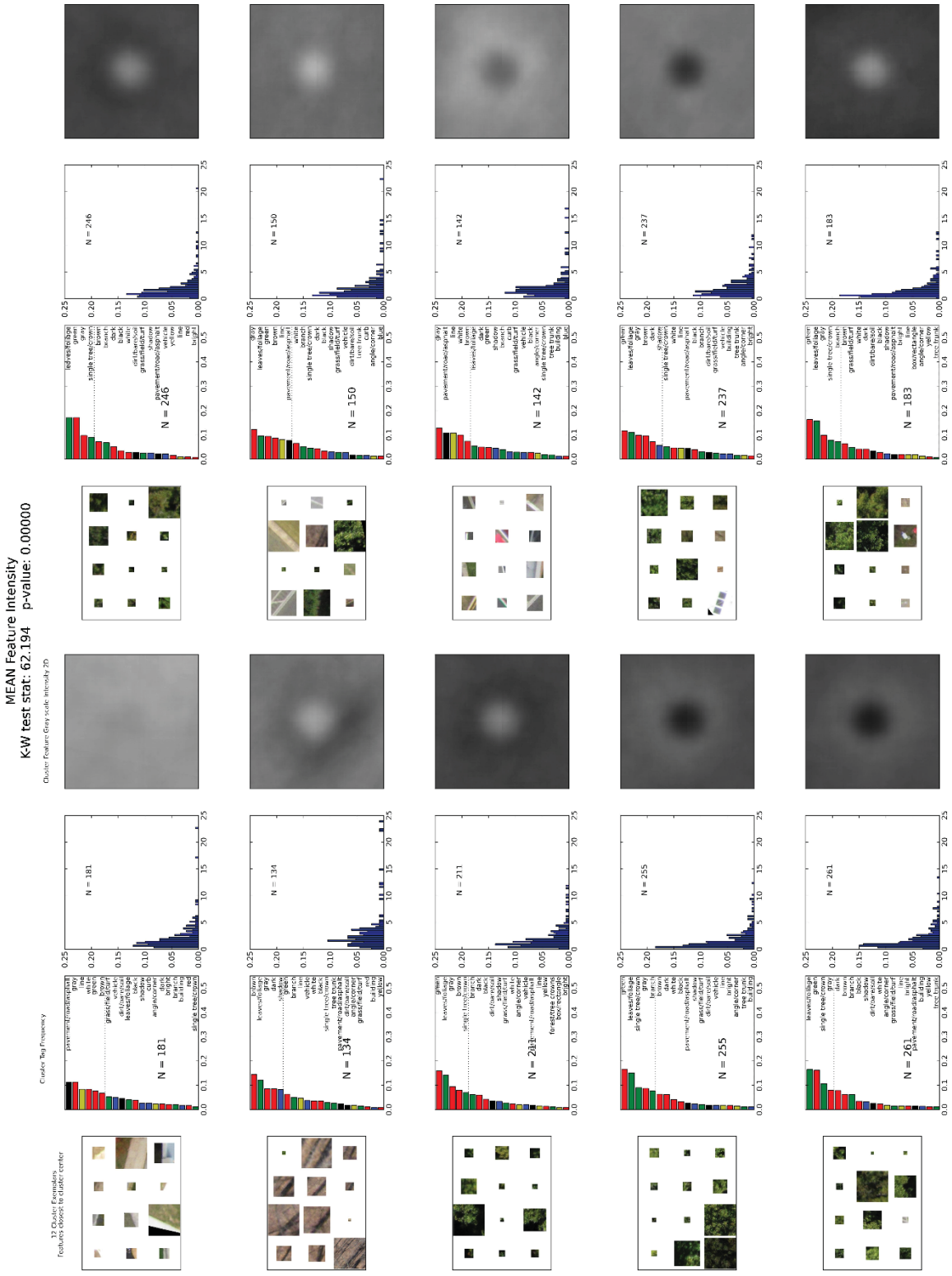


Single Tree Leaf-off 3/5/2013



A2.11

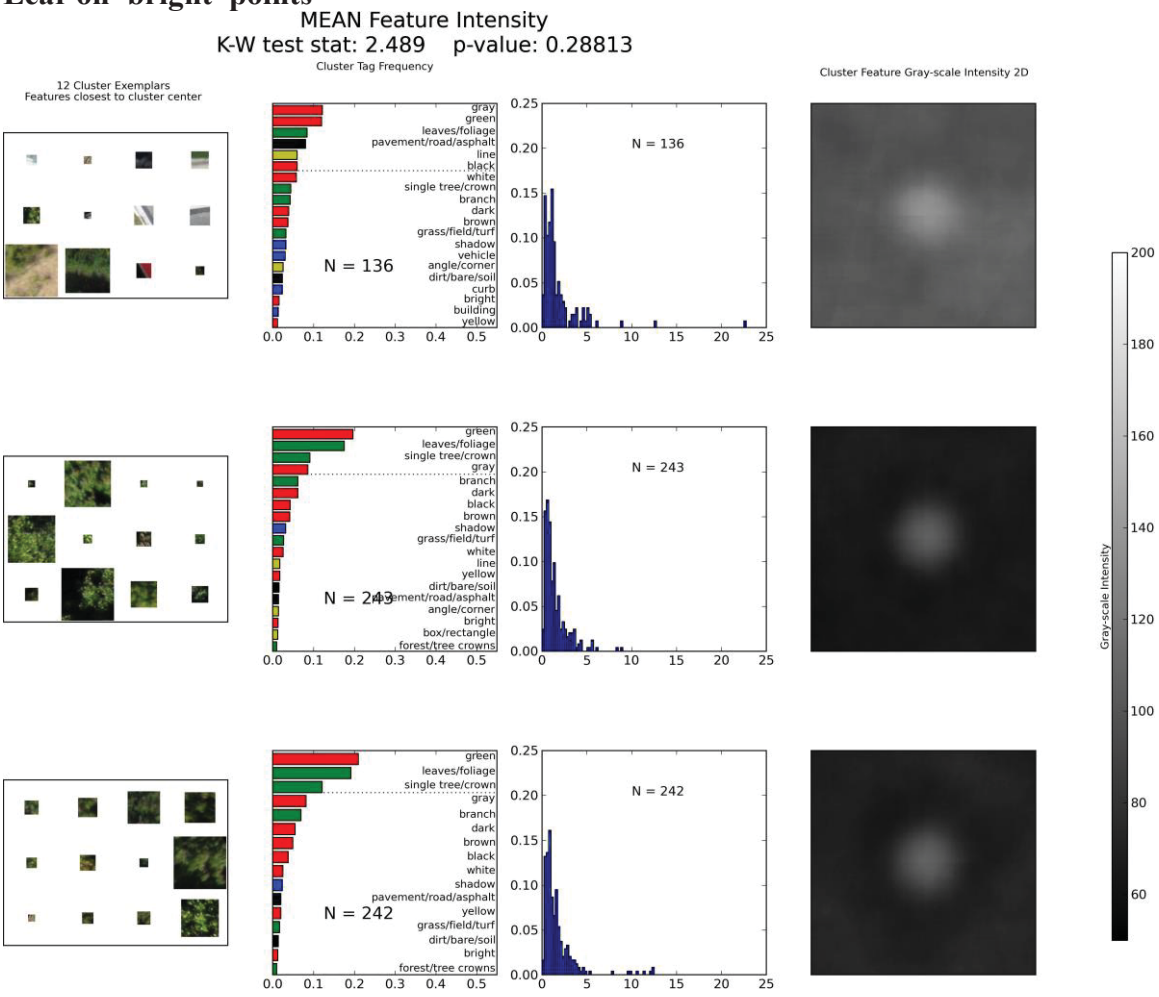
Clustering results on all SIFT 128D descriptors for all datasets at k=10 clusters.



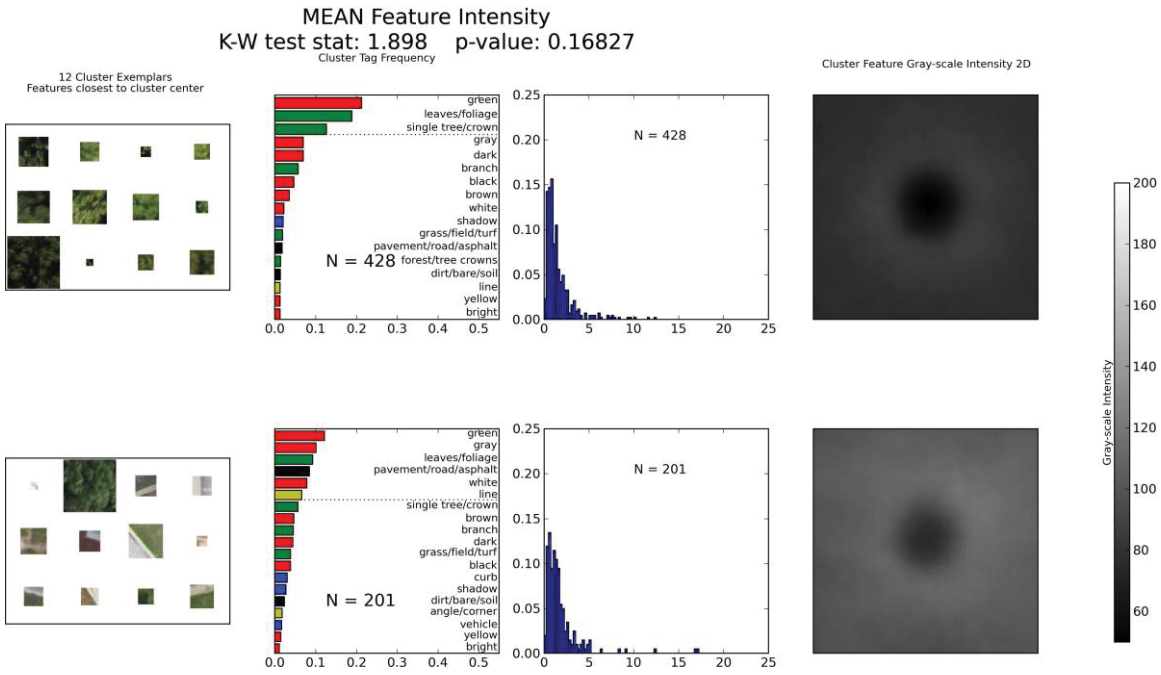
A2.12

Clustering results on all SIFT 128D numeric descriptors broken up by bright or dark spot and by leaf-on or leaf-off aerial datasets.

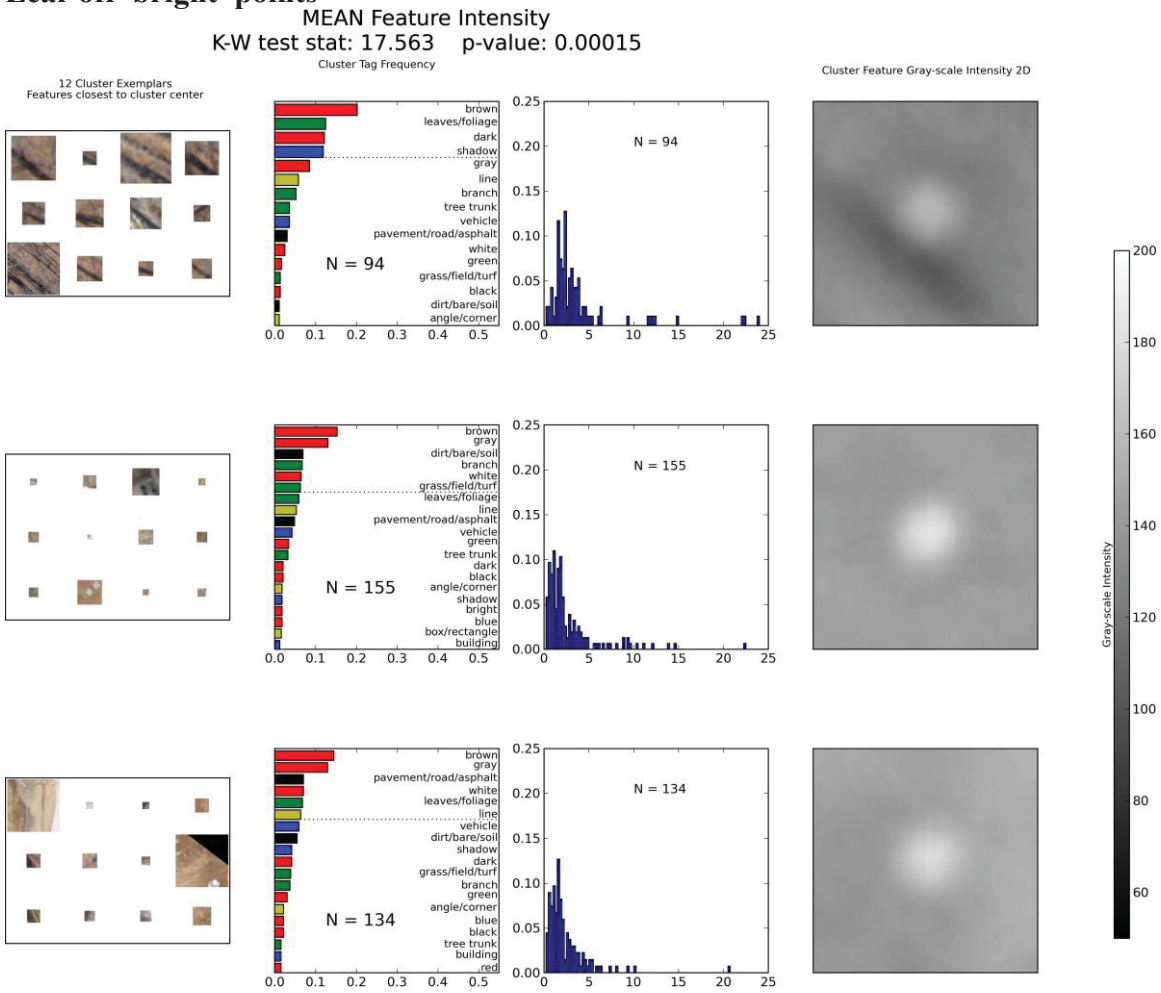
Leaf-on 'bright' points



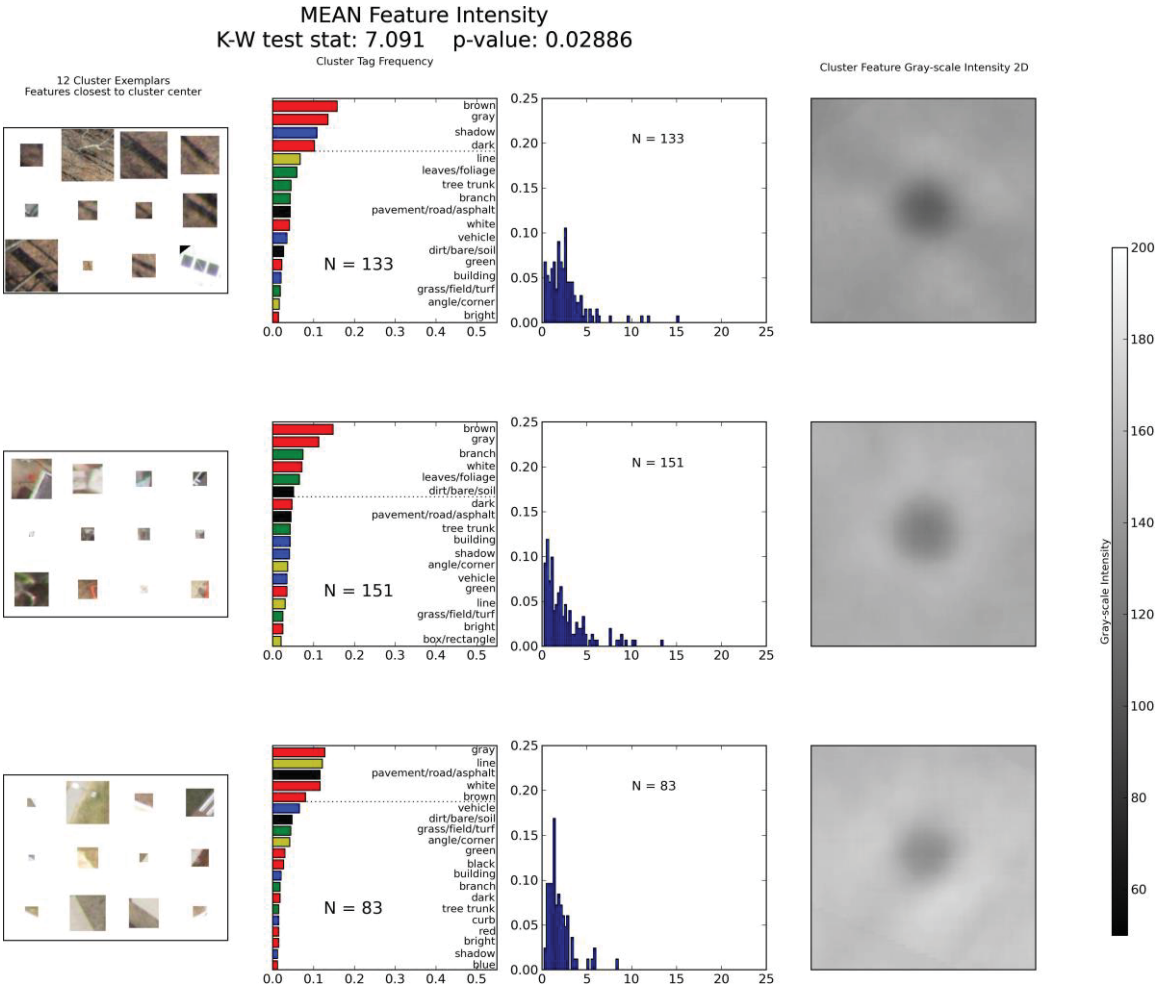
Leaf-on 'dark' points



Leaf-off 'bright' points



Leaf-off 'dark' points

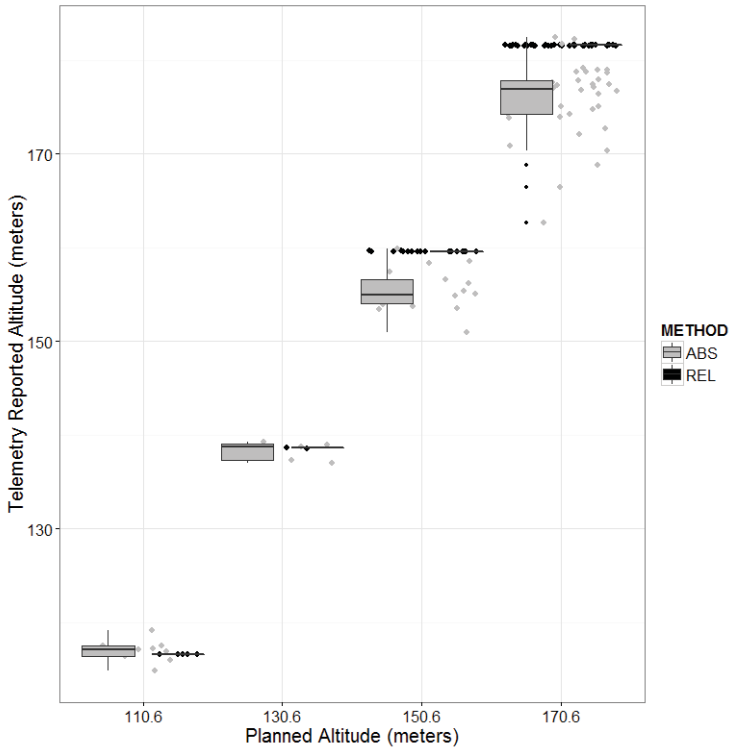


Appendix 3: Supplemental material for Chapter 4

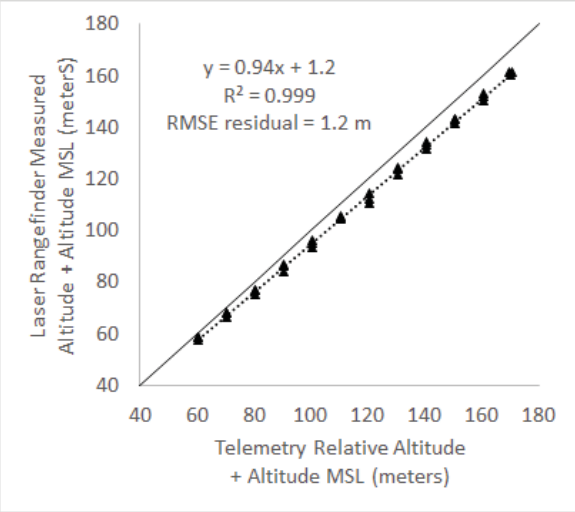
A3.1

Calibration of Arducopter telemetry altitude values. (a) box plots of planned versus actual altitude recordings from absolute and relative elevation values, (b) calibration curve of relative heights to actual heights, (c) raw measurements used for calibration.

A3.1a Boxplots of the relationship between planned altitude (meters above mean sea level, MSL) and mean values of 'absolute' (gray) and 'relative' (black) altitude from Arducopter telemetry files from n = 72 flight replicates. Mean altitude values are shown as points with horizontal 'jitter' to highlight variability in actual flight altitude values for a given planned altitude.



A3.1b Calibration model used to correct relative altitude values to laser rangefinder measured altitude for georeferencing of point clouds.



A3.1c Raw calibration values

		Telemetry Altitude	
Replicate Flight 1	**Laser measured height + MSL**	**Relative Height + MSL**	**Absolute Height + MSL**
1	57.6	60.6	67.8
2	66.6	70.6	76.2
3	75.3	80.5	85.5
4	84.3	90.5	94.8
5	93.4	100.5	103.8
6	104.6	110.5	114
7	112.6	120.5	122.9
8	121.7	130.5	132.4
9	131.5	140.5	141.7
10	141.5	150.5	151.8
11	150.3	160.6	160.6
12	160.3	170.2	169.9
Flight 2			
1	59.1	60.6	67.1
2	68.0	70.5	76.4
3	77.3	80.5	85.2
4	86.7	90.5	94.8
5	96.5	100.6	105.7
6	105.5	110.7	116.1
7	110.5	120.6	126
8	124.3	130.6	135.3
9	134.3	140.5	145.2
10	143.3	150.5	154.1
11	153.3	160.6	163.6
12	161.3	170.5	172.7
Flight 3			
1	59.1	60.6	71
2	68.5	70.6	79.5
3	77.2	80.5	88
4	86.8	90.5	97
5	95.3	100.5	106.2
6	105.3	110.6	116.2
7	114.7	120.5	125.3
8	124.0	130.6	135
9	133.3	140.6	144.7
10	143.3	150.6	153.8
11	152.3	160.6	162.7
12	161.3	169.8	170.9
	RMSE m	6.6	10.5

Glossary

BOF – Bag of features: A computer vision concept used for classifying images based on comparing the frequency / occurrence of a collection of image features within a specific grouping (a bag) to a reference database of features that includes information on feature semantic content.

CHM – Canopy height model: A digital dataset (raster / grid or vector / points) that depicts the height above the ground of forest canopy. Calculated by subtracting a DTM from a DSM on a per-pixel / per-point basis.

Computer Vision: Field of computer science that aims to recreate human vision tasks (classification, recognition, motion tracking, 3D/depth) by analysis of digital images.

DSM – Digital surface model: A digital dataset (raster/grid or vector / points) that depicts the elevation values of the surface relative to the datum, e.g., elevation above mean sea level. When DSM values are differenced to DTM values, a CHM can be produced.

DTM – Digital terrain model: A digital dataset (raster/grid or vector / points) that depicts that elevation values of the ground surface relative to the datum after a filtering algorithm has been applied to remove objects above the surface (trees, shrubs, buildings). Subtracting a raster DTM from a DSM produces a CHM.

Ecosynth: The combination of hardware (cameras, UAS) and software (SFM algorithms, processing code) that forms a personal remote sensing system for producing 3D-multispectal fusion datasets of vegetation at low-cost and on-demand.

Feature: A data structure in computer vision that refers to an entity that has been identified in a 2D image or 3D model by an algorithm (feature detector) for which descriptive information is retained (feature descriptor) to locate or compare that entity to others in other images or datasets.

Feature Descriptor: A numerical vector (string of numbers) that are created by an algorithm (feature detector) to provide a description of an image feature.

Feature Detector: A computer vision algorithm that is designed to automatically identify features within 2D images or 3D models based on locating particular patterns (edges, corners, lines, points, etc.) and then generating a descriptor around the location where the pattern was observed.

Remote Sensing Fusion: Typically, the combination or overlay of remote sensing datasets from different sensors that produce different kinds of measurements (3D and optical image), more generally, the creation of a digital dataset that represents the combined observation of the 3D and spectral properties of the land surface.

Scale-space: A transformation of a signal (1D, 2D, 3D) that depicts that signal at multiple scales, represented by multiple levels of smoothing. The scale-space representation of an image (2D) depicts the image as it would appear at increasingly coarse resolutions or viewed from greater distance. The scale-space representation of an image is a valuable tool for locating features using a feature detector.

SFM – Structure from motion: A concept in computer vision (and associated algorithms) that seeks to automatically determine the location from where images were taken in relationship to each other along with the 3D geometry or structure of the objects observed within images. SFM uses information about the correspondence or matches of features across images as input to a bundle adjustment algorithm, which seeks to determine the optimal set of parameters that minimizes the reprojection error of points within and across images in 2D and 3D space.

SIFT – Scale invariant feature transform: A computer vision feature detector algorithm that uses scale-space to identify features and produce feature descriptors in a gray-scale version of an image that are invariant to changes in image scale and also illumination and rotation. SIFT feature descriptors are represented by an XY location within an image, a measure of the scale or size of the feature, the orientation of the feature, and a 128 dimensional numerical vector that describes the pattern of image gradients around the feature location in a gray-scale version of the image.

UAS – Unmanned aerial system: The combination of hardware (aircraft, GPS, flight controller) and software (flight control program, mission planning software) that combined enable autonomous flight of a remote controlled aircraft.