Remote Sensing of Environment 136 (2013) 259–276

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision ☆

Jonathan P. Dandois, Erle C. Ellis *

Department of Geography and Environmental Systems, University of Maryland Baltimore County, 211 Sondheim Hall, 1000 Hilltop Circle, Baltimore, MD, 21250, USA

ARTICLE INFO

Article history:
Received 26 June 2012
Received in revised form 20 March 2013
Accepted 14 April 2013
Available online 5 June 2013

Keywords:
Canopy structure
Canopy phenology
Forest structure
Forestry
Forest ecology
Computer vision
Unmanned aerial systems
UAS
Unmanned aerial vehicle
UAV
Forest biomass
Forest carbon
Phenology
Forest monitoring
Tree height
Canopy height

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 (<130 m) aerial photographs acquired using off-the-shelf digital cameras mounted on an inexpensive (<USD$4000), lightweight (<2 kg), hobbyist-grade unmanned aerial system (UAS). Ecosynth 3D point clouds with densities of 30–67 points m^{-2} were produced using commercial computer vision software from digital photographs acquired repeatedly by UAS over three 6.25 ha (250 m × 250 m) Temperate Deciduous forest sites in Maryland USA. Ecosynth point clouds were georeferenced with a precision of 1.2–4.1 m horizontal radial root mean square error (RMSE) and 0.4–1.2 m vertical RMSE. Understory digital terrain models (DTMs) and canopy height models (CHMs) were generated from leaf-on and leaf-off point clouds using procedures commonly applied to LIDAR point clouds. At two sites, Ecosynth CHMs were strong predictors of field-measured tree heights (R^2 0.63 to 0.84) and were highly correlated with a LIDAR CHM (R 0.87) acquired 4 days earlier, though Ecosynth-based estimates of aboveground biomass and carbon densities included significant errors (31–36% of field-based estimates). Repeated scanning of a 50 m × 50 m forested area at six different times across a 16 month period revealed ecologically significant dynamics in canopy color at different heights and a structural shift upward in canopy density, as demonstrated by changes in vertical height profiles of point density and relative RGB brightness. Changes in canopy relative greenness were highly correlated ($R^2 = 0.87$) 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 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. 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; Goetz & Dubayah, 2011; Houghton et al., 2009), fire spread and risk modeling (Andersen et al., 2005; Skowronski et al., 2011), commercial and scientific forestry (Næsset & Gobakken., 2008), ecosystem modeling (Antonarakis et al., 2011; Thomas et al., 2008; Zhao & Popescu, 2009), quantitative assessments of habitat suitability and biodiversity (Jung et al., 2012; Vierling et al., 2008) 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 & Martin, 2009). Remote sensing systems with high temporal resolutions are driving similar advances in understanding ecosystem dynamics of forests locally (Richardson et al., 2009) and globally (Zhang & Goldberg, 2011), including the

* Corresponding author. Tel.: +1 410 455 3078; fax: +1 410 455 1056.
E-mail address: ece@umbc.edu (E.C. Ellis).

0034-4257/$ – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.04.005

response of terrestrial ecosystems to changes in climate and land use (Frolking et al., 2009; Morisette et al., 2008; 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 & 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 & Drake, 2000; Popescu et al., 2003). Canopy heights and other metrics of vertical structure are useful for estimating aboveground biomass and carbon density (Goetz & Dubayah, 2011; Lefsky et al., 2002), 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 (Falkowski 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 $0.05–$0.20 per hectare (Asner, 2009). However, typical commercial aerial LIDAR acquisitions often cost a minimum of $20,000 per flight regardless of study area size (Erdody & 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; Dubayah et al., 2010; Frolking et al., 2009).

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 (Anderson et al., 2008; Ke et al., 2010; Turner et al., 2003; Vitousek et al., 2009). 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 & 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. 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 will be demonstrated by comparing vertical profiles of vegetation structure (Parker & Russ, 2004) and RGB relative brightness (Mizunuma et al., 2013; Richardson et al., 2009) 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 & Ellis, 2010; Dey et al., 2012; Rosnell & 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 (Hirschmugl et al., 2007; Ofner et al., 2006), 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