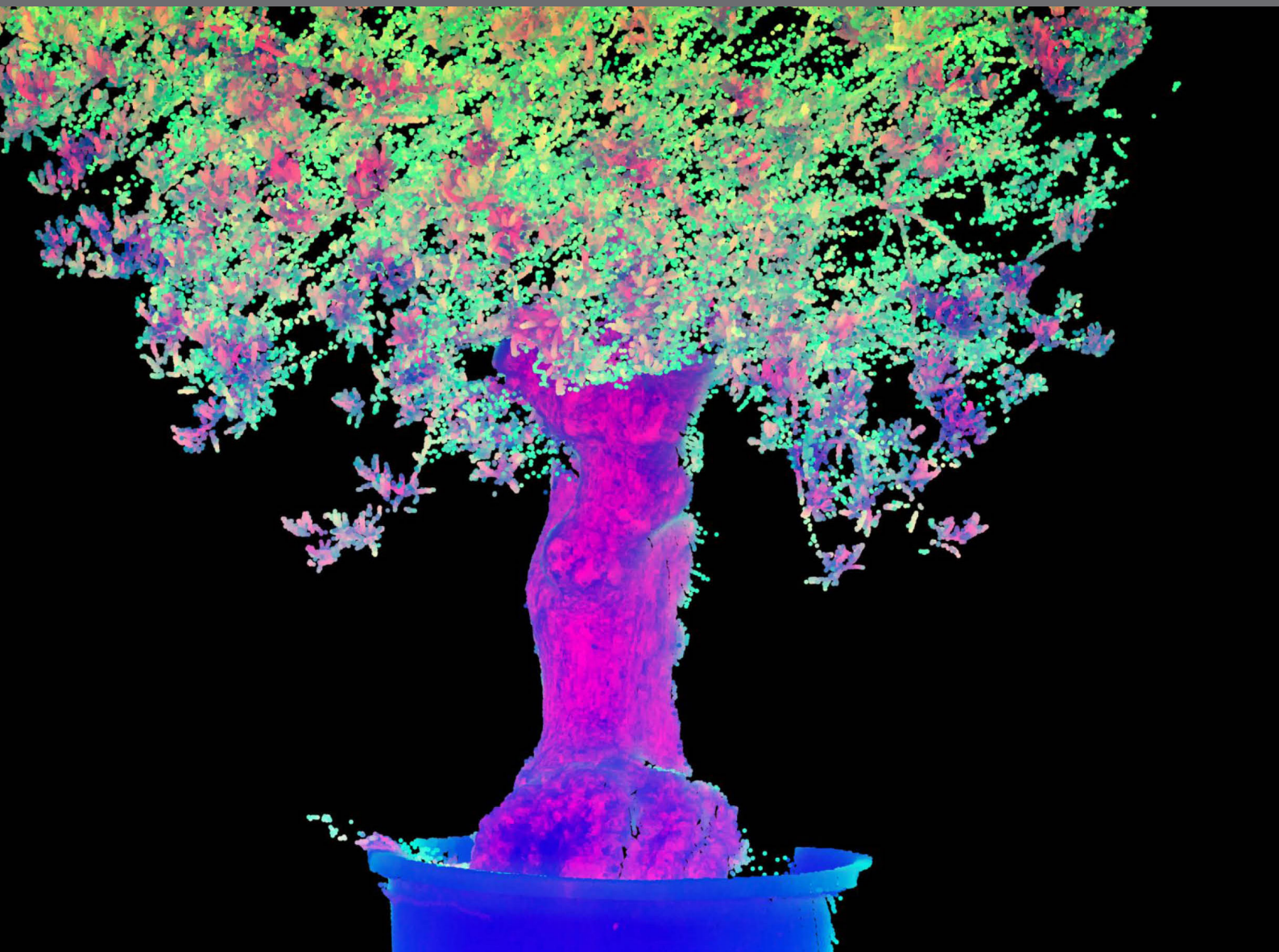


OPTICAL APPROACHES TO CAPTURE PLANT DYNAMICS IN TIME, SPACE, AND ACROSS SCALES

EDITED BY: Eetu Puttonen, Norbert Pfeifer, Alexander Bucksch and
András Zlinszky

PUBLISHED IN: Frontiers in Plant Science





frontiers

Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-543-0

DOI 10.3389/978-2-88945-543-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

OPTICAL APPROACHES TO CAPTURE PLANT DYNAMICS IN TIME, SPACE, AND ACROSS SCALES

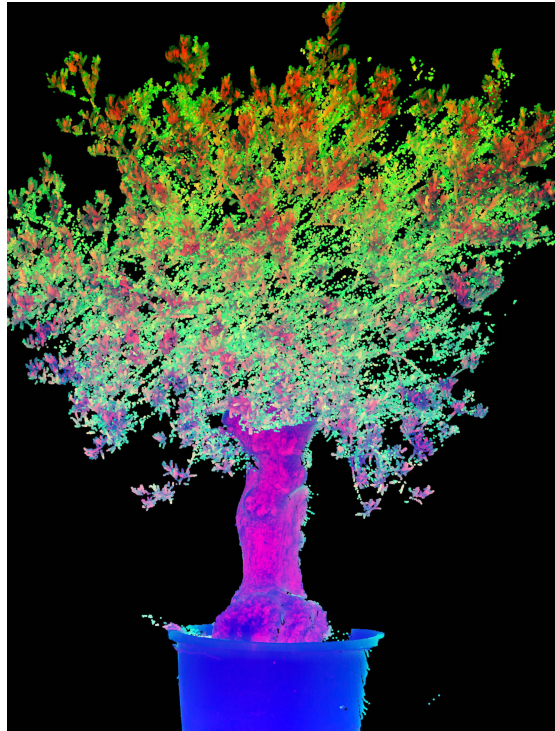
Topic Editors:

Eetu Puttonen, National Land Survey of Finland, Finland

Norbert Pfeifer, Technische Universität Wien, Austria

Alexander Bucksch, University of Georgia, United States

András Zlinszky, Centre for Ecological Research, Hungarian Academy of Sciences, Hungary



Side view of the 3D model of an olive tree produced by terrestrial laser scanning for the purpose of following overnight movement. The colors were produced by RGB blending, with the number of points in a search neighbourhood scaled to the red channel, point amplitude negatively scaled to the green channel, and height above ground negatively scaled to the blue channel. The circle on the bottom right is an artificial target sphere used for co-registration of the individual scans.

Image: Dr. Zlinszky.

Quantifying temporal changes in plant geometry as a result of genetic, developmental, or environmental causes is essential to improve our understanding of the structure and function relationships in plants. Over the last decades, optical imaging and remote sensing developed fundamental working tools to monitor and quantify our environment and plants in particular. Increased efficiency of methods lowered the barrier to compare, integrate, and interpret the optically obtained plant data across larger spatial scales and across scales of biological organization. In particular, acquisition speed at high resolutions reached levels that allow capturing the temporal dynamics in plants in three dimensions along with multi-spectral information beyond human visual senses. These advanced imaging capabilities have proven to be essential to detect and focus on analyzing temporal dynamics of plant geometries.

The focus of this Research Topic is on optical techniques developed to study geometrical changes at the plant level detected within the wavelength spectrum between near-UV to near infrared. Such techniques typically involve photogrammetric, LiDAR, or imaging spectroscopy approaches but are not exclusively restricted to these. Instruments operating within this range of wavelengths allow capturing a wide range of temporal scales ranging from sub-second to seasonal changes that result from plant development, environmental effects like wind and heat, or genetically controlled adaption to environmental conditions. The Research Topic covered a plethora of methodological approaches as suggestions for best practices in the light of a particular research question and to a wider view to different research disciplines and how they utilize their state-of-the-art techniques in demonstrating potential use cases across different scales.

Citation: Puttonen, E., Pfeifer, N., Bucksch, A., Zlinszky, A., eds. (2018). Optical Approaches to Capture Plant Dynamics in Time, Space, and Across Scales. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-543-0

Table of Contents

- 05 Editorial: Optical Approaches to Capture Plant Dynamics in Time, Space, and Across Scales**
Eetu Puttonen, Alexander Bucksch, András Zlinszky and Norbert Pfeifer
- 08 Recruiting Conventional Tree Architecture Models Into State-of-the-Art LiDAR Mapping for Investigating Tree Growth Habits in Structure**
Yi Lin, Miao Jiang, Petri Pellikka and Janne Heiskanen
- 21 Not All Trees Sleep the Same—High Temporal Resolution Terrestrial Laser Scanning Shows Differences in Nocturnal Plant Movement**
András Zlinszky, Bence Molnár and Anders S. Barfod
- 46 Aerial Images and Convolutional Neural Network for Cotton Bloom Detection**
Rui Xu, Changying Li, Andrew H. Paterson, Yu Jiang, Shangpeng Sun and Jon S. Robertson
- 63 Fast High Resolution Volume Carving for 3D Plant Shoot Reconstruction**
Hanno Scharr, Christoph Brieze, Patrick Embgenbroich, Andreas Fischbach, Fabio Fiorani and Mark Müller-Linow
- 75 Chlorophyll Fluorescence Imaging Uncovers Photosynthetic Fingerprint of Citrus Huanglongbing**
Haiyan Cen, Haiyong Weng, Jieni Yao, Mubin He, Jingwen Lv, Shijia Hua, Hongye Li and Yong He
- 86 Can Leaf Water Content Be Estimated Using Multispectral Terrestrial Laser Scanning? A Case Study With Norway Spruce Seedlings**
Samuli Junttila, Junko Sugano, Mikko Vastaranta, Riikka Linnakoski, Harri Kaartinen, Antero Kukko, Markus Holopainen, Hannu Hyypä and Juha Hyypä
- 100 Quantitative Analysis of Cotton Canopy Size in Field Conditions Using a Consumer-Grade RGB-D Camera**
Yu Jiang, Changying L, Andrew H. Paterson, Shangpeng Sun, Rui Xu and Jon Robertson
- 120 Leaf Movements of Indoor Plants Monitored by Terrestrial LiDAR**
Mónica Herrero-Huerta, Roderik Lindenbergh and Wolfgang Gard
- 129 A Low-Cost Imaging Method for the Temporal and Spatial Colorimetric Detection of Free Amines on Maize Root Surfaces**
Truc H. Doan, Tu A. Doan, Michael J. Kangas, AdreAnna E. Ernest, Danny Tran, Christina L. Wilson, Andrea E. Holmes, Erin L. Doyle and Tessa L. Durham Brooks
- 142 In-field High Throughput Phenotyping and Cotton Plant Growth Analysis Using LiDAR**
Shangpeng Sun, Changying Li, Andrew H. Paterson, Yu Jiang, Rui Xu, Jon S. Robertson, John L. Snider and Peng W. Chee
- 159 Considerations for Achieving Cross-Platform Point Cloud Data Fusion Across Different Dryland Ecosystem Structural States**
Tyson L. Swetnam, Jeffrey K. Gillan, Temuulen T. Sankey, Mitchel P. McClaran, Mary H. Nichols, Philip Heilman and Jason McVay



Editorial: Optical Approaches to Capture Plant Dynamics in Time, Space, and Across Scales

Eetu Puttonen^{1,2*}, Alexander Bucksch^{3,4,5}, András Zlinszky⁶ and Norbert Pfeifer⁷

¹ Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, Kirkkonummi, Finland, ² Center of Excellence in Laser Scanning Research, Masala, Finland, ³ Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, United States, ⁴ Institute of Bioinformatics, University of Georgia, Athens, GA, United States, ⁵ Department of Plant Biology, University of Georgia, Athens, GA, United States, ⁶ Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences, Tihany, Hungary, ⁷ Department of Geodesy and Geoinformation, Technische Universität Wien, Vienna, Austria

Keywords: plant growth, plant development, geometric changes in plants, seasonal changes in vegetation structure, change detection

Editorial on the Research Topic

Optical Approaches to Capture Plant Dynamics in Time, Space, and Across Scales

The quest to decipher the phenotype to genotype relationships involves the quantification of plant adaption to the environment at various scales to solve some of the world's most pressing problems (Bucksch et al., 2017). The role of phenotype to genotype relationships within initiatives to increase crop yields for food, fiber, and fuel and to improve prediction of future environmental conditions (IPCC, 2014) is central to the function and well-being of societies world-wide. Quantifying plant morphology over time captures the dynamic structure and function relationships of how plants interact with and respond to environmental stimuli (Balduzzi et al., 2017). A deeper understanding of plant adaption can be achieved if technologies to monitor plants across spatial and biological scales are further developed. Spanning biological scales from the community over the organismal down to the molecular scale is inherently coupled to spatial scales. A wide variety of technological concepts are facilitating the revival of the science of plant morphology and anatomy (Ledford, 2018).

Over the last decades, optical imaging and remote sensing developed the fundamental working tools to monitor and quantify our environment and plants in particular. Satellite imaging increased its spatial, temporal, and spectral resolution to levels where individual trees in forests can be partly identified. However, it remains a challenge to develop pipelines that quantify the traits of plants and their processes from the level of plant populations down to the detail obtained with microscopy. The reason for the pipeline challenge is the heterogeneity of the obtained data ranging from unorganized point clouds over surface models to rasterized imaging and hyperspectral data. For example, airborne methods increased their importance in ecology by measuring aggregate tree traits such as crown width or stem diameter to study community composition (Wieser et al., 2017). However, combining airborne and high-resolution terrestrial data is non-trivial despite improved automatization and higher resolutions of measured plant traits. Unmanned aerial vehicles (UAV) equipped with cameras and multispectral sensors allow the fusion of temporal, spatial, and spectral data to record plant dynamics on ecosystem level at resolutions that also allow the quantification of individual plant morphology.

OPEN ACCESS

Edited and reviewed by:

Yann Guédon,
Centre de Coopération Internationale
en Recherche Agronomique pour le
Développement (CIRAD), France

*Correspondence:

Eetu Puttonen
eetu.puttonen@nls.fi

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 28 March 2018

Accepted: 24 May 2018

Published: 11 June 2018

Citation:

Puttonen E, Bucksch A, Zlinszky A
and Pfeifer N (2018) Editorial: Optical
Approaches to Capture Plant
Dynamics in Time, Space, and Across
Scales. *Front. Plant Sci.* 9:791.
doi: 10.3389/fpls.2018.00791

In particular, the advances in algorithms to handle heterogeneous data of various sensors revealed details of plant adaptation to discover the genes controlling adaptation mechanisms. The wealth of options presents a new challenge in testing and selecting an appropriate approach that scales well with the research questions. Making modern technology applicable to the hypothesis driven science process requires more than learning a few techniques. In the light of a particular hypothesis it requires also resources to acquire, build, and refactor equipment and software. Here, a wider view to different research disciplines and how they utilize their state-of-the-art techniques can be most beneficial in demonstrating potential use cases. Ideally, this will lead to joint use of best practices in technology innovation and discovery in the plant sciences.

The Topic received a diverse set of 11 original research manuscripts that cover a wide array of applications across different scales. On the spatial scale, Scharr et al. present a fast imaging approach to perform volumetric 3D carving of plant roots on submillimeter level. Sun et al. apply LiDAR for high-throughput phenotyping on cotton plants while Lin et al. use LiDAR mapping to investigate how Hallé's models of tree architecture reflect in real data. Xu et al. apply Structure-from-Motion to images collected with unmanned aerial vehicle in cotton fields to detect the exact flowering time and the precise locations of cotton balls using convolutional neural networks. Swetnam et al. link individual plant morphology to plant populations by fusing point clouds collected with ground-based, UAV mounted, and airborne systems to study different structural states of a dryland ecosystem.

Temporal dynamics were quantified by Zlinszky et al. who measured variation in nocturnal branch movements between different species with short interval LiDAR. Similarly, Herrero-Huerta et al. studied the movement of leaves in *Calathea roseopicta* under different lighting conditions with a LiDAR system. The outcome of these investigations challenges our view of plants as passive, static organisms, and shows that they are capable of short-term changes in shape at the whole plant level.

While not completely non-invasive, Doan et al. published a first low-cost imaging system to detect formation of free amino acid groups on maize roots to elucidate exudate responses. In doing so, a special, ninhydrin-injected paper was monitored over a 3-week growth period. Jiang et al. reported an application using low-cost consumer RGB-D cameras to monitor and quantify the growth of cotton canopies over a 3-month period.

In addition, two manuscripts harnessed spectral information to study dynamic responses in plant leaves outside the visible region. Cen et al. detected photosynthetic fingerprints of citrus Huanglongbing (so-called "greening") disease with fluorescence

spectroscopy by combining imaging of specific fluorescence properties with advanced image analysis. Junttila et al. estimated leaf equivalent water thickness using combinations of different LiDAR wavelengths. Their results showed significant correlation with the standard laboratory reference measurements. These findings demonstrate the significant potential in expanding multi-wavelength laser scanning measurements to whole forests to accurately monitor ecophysiological parameters.

Our Research Topic proves that optical technologies have reached a level of precision, reliability and detail that is sufficient to study minute physiological processes in plants. In addition to documenting plant status at various levels, imaging, and scanning methods allow following processes of adaptation, movement or disease in real time as they happen. In addition to serving as early detection systems (Cen et al.; Junttila et al.) for management purposes, these methods allow new insights into plant function. One particular example for this is the observed short-term periodicity in plant movement and trunk diameter, which suggests a new approach to water transport in trees (Zlinszky and Barford, 2018).

The Research Topic covered a plethora of methodological approaches as suggestions for best practices in the light of a particular research question. As such, the Research Topic collected papers that demonstrate how technology development and scientific discovery in the plant sciences can accelerate each other toward deciphering the adaption of plants as part of the quest to reveal the phenotype to genotype map. The future challenge will be to extend the multi-scale measurement of static traits to multi-scale measurement of chemical processes within and between plants. For example, how do diseases or ripening processes affect the community structure or how does microclimate fluctuation of single plants affect overall edible yield? How do shape changes in canopies relate to the hydraulic properties at tissue level? We believe that the optical methods in this Research Topic will become openly available such that all researchers across the plant sciences will be capable to perform research at this exciting frontier.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

AZ received funding under the grant No. OTKA PD 115833.

REFERENCES

Bucksch, A., Atta-Boateng, A., Azihou, A. F., Battogtokh, D., Baumgartner, A., Binder, B. M., et al. (2017). Morphological plant modeling: unleashing geometric and topologic potential within the plant sciences. *Front. Plant Sci.* 8:900. doi: 10.3389/fpls.2017.00900

Balduzzi, M., Binder, B. M., Bucksch, A., Chang, C., Hong, L., Iyer-Pascuzzi, A. S. et al. (2017). Reshaping plant biology: qualitative and quantitative descriptors for plant morphology. *Front. Plant Sci.* 8:117. doi: 10.3389/fpls.2017.00117

IPCC (2014). "Climate Change 2014: synthesis report," in *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team]*, eds

- R. K. Pachauri and L. A. Meyer (Geneva: IPCC), 151. Available online at: <https://www.ipcc.ch/index.htm>
- Ledford, H. (2018). The lost art of looking at plants. *Nature* 553, 396–398. doi: 10.1038/d41586-018-01075-5
- Wieser, M., Mandlbauer, G., Hollaus, M., Otepka, J., Glira, P., and Pfeifer, N. (2017). A case study of UAS borne laser scanning for measurement of tree stem diameter. *Remote Sens.* 9:1154. doi: 10.3390/rs9111154
- Zlinszky, A., and Barford, A. (2018). Short interval overnight laser scanning suggest sub-circadian periodicity of tree turgor. *Plant Signal. Behav.* 13:2. doi: 10.1080/15592324.2018.1439655

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Puttonen, Bucksch, Zlinszky and Pfeifer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Recruiting Conventional Tree Architecture Models into State-of-the-Art LiDAR Mapping for Investigating Tree Growth Habits in Structure

Yi Lin^{1*}, Miao Jiang², Petri Pellikka³ and Janne Heiskanen³

¹ Beijing Key Lab of Spatial Information Integration and Its Applications, School of Earth and Space Sciences, Institute of Remote Sensing and GIS, Peking University, Beijing, China, ² Institute of Mineral Resources Research, China Metallurgical Geology Bureau, Beijing, China, ³ Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

OPEN ACCESS

Edited by:

András Zlinszky,
Institute of Ecology Research Center,
Hungarian Academy of Sciences,
Hungary

Reviewed by:

Christophe Pradal,
Agricultural Research Center for
Development, France
Evelyn Costes,
INRA UMR Amélioration génétique et
adaptation des plantes
méditerranéennes et tropicales,
France

*Correspondence:

Yi Lin
yi.lin@pku.edu.cn

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 30 September 2017

Accepted: 05 February 2018

Published: 20 February 2018

Citation:

Lin Y, Jiang M, Pellikka P and
Heiskanen J (2018) Recruiting
Conventional Tree Architecture Models
into State-of-the-Art LiDAR Mapping
for Investigating Tree Growth Habits in
Structure. *Front. Plant Sci.* 9:220.
doi: 10.3389/fpls.2018.00220

Mensuration of tree growth habits is of considerable importance for understanding forest ecosystem processes and forest biophysical responses to climate changes. However, the complexity of tree crown morphology that is typically formed after many years of growth tends to render it a non-trivial task, even for the state-of-the-art 3D forest mapping technology—light detection and ranging (LiDAR). Fortunately, botanists have deduced the large structural diversity of tree forms into only a limited number of tree architecture models, which can present *a-priori* knowledge about tree structure, growth, and other attributes for different species. This study attempted to recruit Hallé architecture models (HAMs) into LiDAR mapping to investigate tree growth habits in structure. First, following the HAM-characterized tree structure organization rules, we run the kernel procedure of tree species classification based on the LiDAR-collected point clouds using a support vector machine classifier in the leave-one-out-for-cross-validation mode. Then, the HAM corresponding to each of the classified tree species was identified based on expert knowledge, assisted by the comparison of the LiDAR-derived feature parameters. Next, the tree growth habits in structure for each of the tree species were derived from the determined HAM. In the case of four tree species growing in the boreal environment, the tests indicated that the classification accuracy reached 85.0%, and their growth habits could be derived by qualitative and quantitative means. Overall, the strategy of recruiting conventional HAMs into LiDAR mapping for investigating tree growth habits in structure was validated, thereby paving a new way for efficiently reflecting tree growth habits and projecting forest structure dynamics.

Keywords: tree growth habits, Hallé architecture model (HAM), light detection and ranging (LiDAR), tree species classification, morphological features

INTRODUCTION

Trees play a fundamental role in maintaining forest ecosystems by adapting their biophysical, biochemical or physiological characteristics to the integrity of biotic and abiotic environmental influences (Saxe et al., 2001). The adaptation effects apparently show in their structures and growth habits (Creber and Chaloner, 1984). For example, the contemporary structures of trees mirror

the inherent rules of their various organic components being organized; by following these rules, people estimated the function capacities of trees such as crown photosynthesis efficacy (Lovell et al., 2006). As regards to the dynamical growth habits of trees such as their continuous or rhythmic growth modes, tree habit modeling facilitates decisions-making in advance, which is favorable for sustainable forest managements (Pretzsch et al., 2006). Hence, learning tree growth habits in structure has considerable implications for understanding from how single trees cope with diversified environmental stresses to how integral forests function in multiplex terrestrial processes (Pinard et al., 1999). This task is especially interesting for the trees growing in the boreal environment, since characterizing tree growth habits in structure in such a condition that is enduring quicker warming than global warming (Shepherd, 2016) facilitates projecting how trees will respond to the climatic scenarios in the future.

However, mensuration of tree growth habits in structure across large areas is a non-trivial task. This is due to the difficulty of measuring and modeling tree crown structures that comprise leaves and branches of varying shape forms, growth modes, and organization layouts at the fine scales. In addition, the conventional approaches, e.g., field investigation and estimation of tree growths by comparing multi-temporal canopy surface models (Yu et al., 2006), are far from enough for accurately mapping crown structure and dynamics. Therefore, new techniques with higher potentials of characterizing tree growth habits in structure are widely and urgently demanded.

Remote sensing (RS) supplies a more efficient solution plan to accomplish this task (Heinzel and Koch, 2012). The state-of-the-art RS technology—static terrestrial laser scanning (TLS), with the merits of 3D high-density sampling of objects, seems to be the best choice to date. Based on TLS data, an LiDAR data to tree *Architecture (L-Architect)* architecture model has been designed to provide a feasible framework for quantifying the spatial distributions of tree components and explicitly describing 3D tree architectures (Côté et al., 2011). From TLS-recorded 3D points, quantitative retrievals of crown structure and foliage assemblage situations have also been tried (Yang et al., 2013). In a broader sense, TLS has also been enthusiastically attempted for deriving various structural parameters of trees, including leaf area density (Hosoi and Omasa, 2007), tree height and crown diameter (Moorthy et al., 2011; Lin et al., 2012), and tree biomass (Calders et al., 2015). However, although TLS has proved to be able to characterize tree structures well, it is difficult to derive tree growth habits directly from TLS point clouds due to the complex structures and organizations of leaves and branches. Eitel et al. (2013) placed a TLS system at a fixed location for monitoring forest plot dynamics, but the time span was too short to infer the growth habits in structure. Worse still, TLS is principally restricted by its limited coverage; consequently, TLS is more assumed to supply the reference samples for calibrating large-area forest canopy structure measurements (Côté et al., 2012).

As another technical branch of light detection and ranging (LiDAR), airborne laser scanning (ALS) that is applicable for large-scale forest mapping can solve this shortage to some extent. In retrospect, the usages of ALS for forest inventory had a longer

history than TLS (Yang et al., 2013), while this history showed that applying ALS for forest inventories still needs to consider tree architectures when mapping accuracies are concerned. This is evidenced by the fact that numerous ALS-based forest researches were conducted based on presumptions of diversified tree crown archetypes, as reviewed by Calders et al. (2013). For example, Koetz et al. (2007) used a hemi-ellipsoidal tree crown archetype, and Ferraz et al. (2012) assumed an ellipsoidal one. The results indicated that geometric archetypes of tree crowns did help in forest investigations, yet the previous selections of archetypes were mostly based on the external shape outlines of tree crowns represented by 3D ALS points (Koetz et al., 2007; Ferraz et al., 2012; Calders et al., 2013). That is, almost neither geometries nor topologies of crown-internal structures such as the layouts of leaves and branches were examined, and most of the previous studies have not considered tree architectures in a real sense. This is caused by that the relatively low sampling density of airborne LiDAR is the most kernel factor restricting its performance on detailed representations of crown architectures. This adverse condition also restricts the solutions of fine-scale plant architecture modeling (Davidson et al., 2008) from playing their full roles. Although the latest ALS systems are continuously improved with their sampling densities increased, they are far from enough for characterizing crown-internal structural details (Ko et al., 2013).

Fortunately, botanists have generalized tree forms of a gigantic diversity into a limited number of tree architecture models, which facilitate sketchily giving *a-priori* knowledge of tree structure and development. Hallé et al. (1978) proposed a scheme to deduce the branch and stem architectures of tropical trees into 23 tree architecture models (hereafter called Hallé architecture models, HAMs). These HAMs have also been applied onto multiple temperate tree species (Fisher, 1984; Costes et al., 2006). As a representative set of tree architecture models, their proposal was rooted in that there are inherent modes of organizing stems and branches for different tree species, and the associated genetically-determined rules don't vary along with trees growing (Negrón et al., 2013). Tree architecture models have also been used to determine the branching forms considered in plant constructions (Barthélémy and Caraglio, 2007), e.g., based on the hidden semi-Markov models (Guédon et al., 2001) or the stochastic models (Costes and Guédon, 2002; Renton et al., 2006). HAMs have also shown their usefulness in improving the understanding of tree, canopy, and even forest properties. For example, Mutke et al. (2005) applied a Rauh HAM for characterizing the grafted stone pines (*Pinus pinea* L.) to figure out their shoot growths and bud differentiations. Feng et al. (2012) synthesized a Rauh HAM and a forest dynamics model to simulate the development of a forest stand, with the stand-level ecological/silvicultural model and the tree-level architectural representations integrated to provide the details of the individual trees at the stand level.

With the HAMs used as a linking bridge, ALS can supply more information on the structural, morphological, and biophysiological properties of trees. Theoretically, although ALS cannot make the same performance on tree structure reconstruction as TLS does, it can reveal the sketchy features such as the spatial pattern of laser points, the penetrations for

different crown layers and crown integral symmetry. These three case features correspond to the external shapes of trees and crown-internal structures (i.e., leaf distributions and branching modes), which are of the potentials for deriving the appropriate HAMs. Actually, in the previous ALS-based endeavors, it has already been realized that grasping more detailed structural features of crowns can better the depth of tree understanding (Hyde et al., 2005; Wang and Glenn, 2008; Ferraz et al., 2012). With the promotion of ALS performance, particularly on the sampling density, the trend of making use of crown-internal structure is getting more and more distinct. For instance, tree genera classification has been explored with the geometric traits of branches measured with high-density ALS (Ko et al., 2013). The HAMs proposed by Hallé et al. (1978) were defined in a descriptive (qualitative) way; hence, no explicit attempts, so far, have emerged to connect these theoretical models to real ALS point clouds to derive more about tree growth habits in structure. Consequently, a question can be asked—Investigating tree growth habits in structure: using traditional HAMs, state-of-the-art ALS, or both?

To fill this gap, this study attempted to recruit the conventional tree architecture models into the state-of-the-art ALS based investigation of tree-level properties, here tree growth habits in structure. As suggested by Hallé et al. (1978), HAMs depend on tree species. Accordingly, the used strategies were to classify tree species based on the ALS-collected data, identify the HAM for each of the targeted tree species, and derive the properties of the targeted trees by referring to their related HAMs. To validate this solution plan, this work needed to implement the following three tasks: (1) converting the descriptive HAM-indicated structural traits into a set of quantitative parameters by following the characteristics in the ALS-based tree representation; (2) classifying tree species based on the proposed crown- and tree-associated feature parameters; and (3) identifying the HAM for each of the tree species to supply the information about tree growth habits in structure.

MATERIALS AND METHODS

Study Area and Data Collection

The study area was located on the Seurasaari Island, Helsinki, Southern Finland (N60°10'52", E24°53'4"). The island is a wooded park and has rocks, hills, wetlands, and herb-rich forests, covering approximately a total of 46 hectares. The tree species include Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*), European rowan (*Sorbus aucuparia*), European aspen (*Populus tremula*), alder (*Alnus* sp.), birch (*Betula* sp.), English oak (*Quercus robur*), Norway maple (*Acer platanoides*), and small-leaved lime (*Tilia cordata*), in a descending order of abundance. The studied trees lie in the middle of the forest with varying understory vegetation.

The airborne LiDAR data were collected by using an Optech ALT3100 laser scanner (Optech Inc., Ontario, Canada) in September 2010. The flight altitude was set at ~400 m. One laser pulse triggered 1, 2, 3, or 4 returns. The density of the collected point clouds is ~10 points per m², which can be classified into an intermediate point density compared with the very-high

density reported by Ko et al. (2013). To acquire the reference data, the TLS data were collected with a Leica HDS6100 system (Leica Geosystem AG, Heerbrugg, Switzerland) in September 2010. HDS6100 is a 690 nm phase-based continuous-wave laser scanner with a 360° × 310° Field-of-view; its data acquisition rate is 508,000 points per second. The distance measurement accuracy is ±2 mm at a range distance of 25 m. The point spacing is 6.3 mm at 10 m. The TLS measurements were finished in a multi-scan mode (Liang and Hyypä, 2013), which can result in good point coverage and further supply tree structure representations as the reference data. More detailed specifications of the test settings were described by Holopainen et al. (2013). The field inspections were conducted in January 2012, and the species of the target trees were visually identified.

Data Preparation

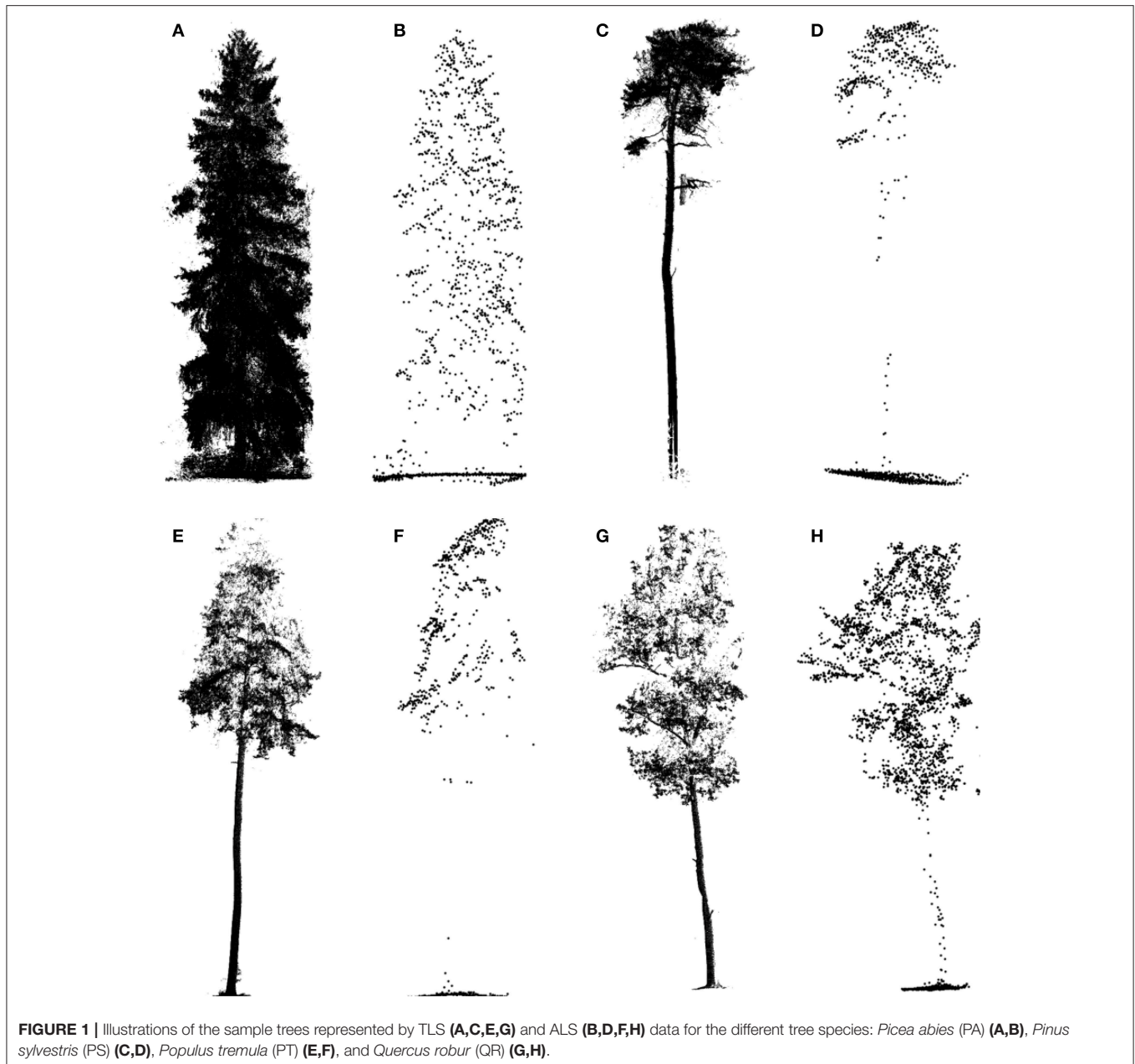
Based on the pre-set reference markings, the ALS and TLS data were registered by using the Cyclone software (Leica Geosystem AG, Heerbrugg, Switzerland). The procedures of tree isolation and stem segmentation were interactively accomplished by using the Terrascan software (Terrasolid Oy, Helsinki, Finland), and the detailed operations of these isolation and segmentation were delineated by Holopainen et al. (2013). The stem axes of the isolated trees were determined by principal component analyses (PCA) of the extracted laser points on stems (Lehtomäki et al., 2010). The results can help calculate tree feature parameters in a uniform manner.

After data preprocessing, a total of 40 trees with their structural representations simultaneously by the ALS and TLS data were chosen for the tests. The trees belong to four typical boreal tree species, including nine Norway spruces (*P. abies*, PA), 14 Scots pines (*P. sylvestris*, PS), 7 European aspens (*P. tremula*, PT), and 10 English oaks (*Q. robur*, QR). The specimens represented by the ALS and TLS data are illustrated in **Figure 1**. Their descriptive statistics in terms of tree height and crown length (i.e., crown thickness Hauglin et al., 2014) are listed in **Table 1**.

Schematic Program

The schematic program of combining the HAMs and ALS was designed (**Figure 2**). The dashed arrow line marks the status that it is hard to directly derive tree growth habits in structure from ALS data. After all, different to TLS that is even available for tree architecture modeling at the branch scales (Côté et al., 2011), most contemporary ALS systems show low point densities. The HAMs are used to work as a bridging link to fill this technical gap. Specifically, the HAM-delineated tree structural organization rules are used to guide tree species classification; then, the HAM that can best model each tree species is identified; finally, the structural growth habits for each of the tree species of interest are derived from its associated HAM. In this schematic framework, tree species classification serves as a kernel procedure; hence, seeking how to add the accuracy of this procedure is very important.

For LiDAR-based tree species classification, a large number of approaches has been proposed, even aiming at the scenario

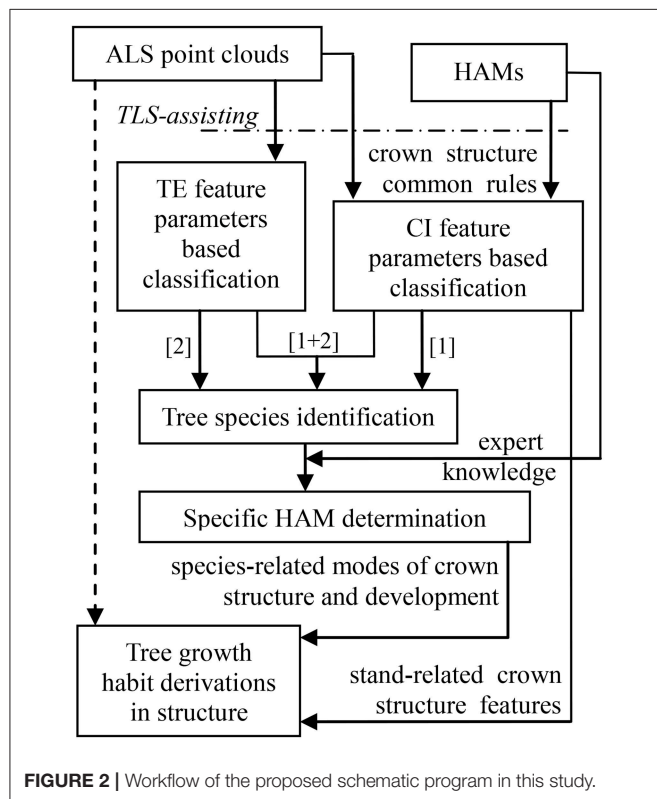


that merely limited-size sample data are collected (Puttonen et al., 2010). Their specific methodologies have proved to work well on the corresponding ALS data (Dalponte et al., 2008; Heinzl and Koch, 2012; Ko et al., 2013) and TLS data (Lin and Herold, 2016) that are of varying conditions. However, almost all of these previous studies suggested that both distinctive feature variables and high-performance classifiers are still needed. This schematic program was aimed at these two aspects, i.e., comprehensively proposing feature parameters and classifiers. For example, the parameters capable of reflecting tree external (TE) structures such as tree height can be intuitively extracted, but the feature parameters reflecting crown internal (CI) structures such as tree symmetry need to be proposed

by referring to the structural characteristics presented by the associated HAMs. The classifications based on these two kinds of parameters correspond to the routines [1] and [2] in Figure 2, respectively. Tree species classification can be deployed on the two sets of feature parameters separately or together ([1]+[2]) to enhance the overall accuracy. The following sub-Section Tree Species Classification specifically describes how to derive feature parameters from the HAMs and the selected classifier appropriate for the associated classification scenario in the present study. Next, the sub-section HAM Identification and Information Derivation delineates how to determine the related HAMs, via referring to the expert knowledge supplied by the potential HAMs, and derive tree growth habits in structure.

TABLE 1 | Descriptive statistics for the sample trees.

Tree species	Number	Tree height (m)			Crown length (m)		
		Max	Min	Mean	Max	Min	Mean
<i>Picea abies</i>	9	28.38	17.60	23.68	27.38	14.02	21.82
<i>Pinus sylvestris</i>	14	23.73	16.66	21.06	18.83	5.67	11.17
<i>Populus tremula</i>	7	25.98	20.39	23.63	20.11	11.64	16.37
<i>Quercus robur</i>	10	25.94	15.17	20.21	18.94	13.81	16.69

**FIGURE 2** | Workflow of the proposed schematic program in this study.

Tree Species Classification

TLS-Assisted Analysis of HAM Signatures

For the four tree species, their HAMs were visually interpreted in advance, via referring to Prusinkiewicz and Remphrey (2000), in order to obtain the ground-truth data that were used to test the performance of the proposed schematic program. The manually-identified HAMs are demonstrated in **Figure 3**. The Massart model shows plagiotropic structure, with the main branches being in whorls and rhythmic growths. The Rauh model shows different orthotropic structure, with its main branches being morphogenetically equivalent to its trunk, all in a rhythmic growth mode. The Roux model demonstrates plagiotropic, monopodial, and non-phyllomorphic structure and continuous growths. The Attim model exhibits orthotropic structure and more or less continuous growths, also with its main branches being morphogenetically equivalent to its trunk. More specific information about tree growth habits in structure can refer to the relevant literature (Hallé et al., 1978; Fisher, 1984; Costes

et al., 2006). Given that the focus of this study was on examining the proposed schematic program rather than on revealing more growth habits of the sample trees, the above-listed structures such as plagiotropic structure and continuous growth were referred to for proposing more appropriate feature parameters for tree species classification and deriving tree growth habits in structure.

Proposal and Derivation of Tree Structure Feature Parameters

The performance on exploiting tree structure growth habits is largely dependent on whether the assumed feature parameters are suitable for tree species classification. However, none of the previous studies have generalized each of the 23 HAMs into a set of distinguished feature parameters. To fill this gap, the first work was to propose feature parameters effective for operating the step of tree species classification, with the HAM-indicated descriptive signatures and the characteristics of ALS-based tree representation considered in a comprehensive way. That is, the derived tree structure feature parameters from the ALS-based tree representations need to characterize the marked differences of orthotropic vs. plagiotropic branching and rhythmic vs. continuous growth reflected by the four models as much as possible.

First, the CI structure feature parameters were proposed. For different branching patterns within crowns, the related ALS-collected point clouds demonstrate different spatial-distribution modes. For example, the orthotropic branches often correspond to a symmetric distribution of point clusters that represent branches, but the plagiotropic branches more often mean high laser penetration in trees. Such knowledge comprises the rules linking between the structural features and the HAMs. In accordance to this rule set, the HAM-associated CI structure parameters were extracted by applying the 8-segments framework developed by Lin and Hyypä (2016). Next, a total of 13 CI parameters were extracted, with their derivations from the specific characteristics of the HAMs and their quantitative definitions listed in **Table 2**. For instance, the fourth parameter (P4) is defined as the largest voxel density within 1 m super-voxels to reflect the HAM situation of leaf clustering. These parameters can allude to different branching and leaf distribution modes, which are characteristic for different HAMs.

Then, the TE structure feature parameters were also proposed. The procedures of deriving the HAM-related TE structural features followed the workflow proposed by Lin and Herold (2016). In addition to the common structural parameters (Lin and Herold, 2016), additional structural parameters were defined (**Table 3**), by seeking the corresponding more robust parameters or by combining two common parameters. For the first case, the diameter of the circle with area equal to the crown cover (D_{EA}) was defined, which can reduce the effects of crown shape variability. For the second case, the ratios between a couple of feature parameters and tree height were also examined, as shown by the parameter of $LcHt$ (**Table 3**). This operation can handle the influences of varying tree ages on tree species classification. The specific 10 TE parameters can also characterize HAMs, and their combinations are of the potential for reflecting the structural differences between the four tree species.

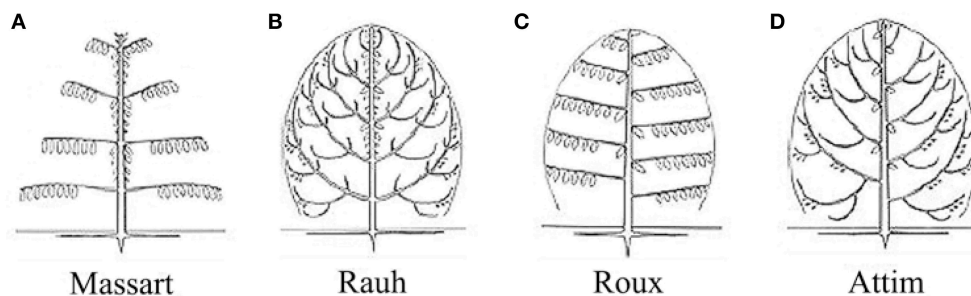


FIGURE 3 | Illustrations of the HAMs proposed by Hallé et al. (1978): **(A)** Massart, **(B)** Rauh, **(C)** Roux, and **(D)** Attim.

TABLE 2 | CI feature parameters derived from the ALS point clouds by referring to the HAMs.

Abbr	Definition	HAM-rule	Formula
P1	Ratio between the height of the equivalent centers for the voxels within each profile and crown length (average for 8 profiles)	Horizontal branch arrangement	$P1 = \frac{\sum_8 \left(\sum_{i=1}^n H_{G_i} / n \right)}{8 \cdot L_c}$
P2	Ratio between the radius of the equivalent centers for the voxels within 8 profiles and crown radius (average for 8 profiles)	Vertical branch arrangement	$P2 = \frac{\sum_8 \left(\sum_{i=1}^n R_{G_i} / n \right)}{8 \cdot R_c}$
P3	Ratio between the area of the below 1/3 laser points and crown base area	Old branches longer or not	$P3 = A_{\text{down } \frac{1}{3} \text{ pts}} / A_{\text{Base}}$
P4	Largest voxel density within 1 m super-voxels	Leaf clustering	$P4 = \max(N_{G_{1m}})$
P5	Standard deviation of P1 for 8 profiles	Consistency	$P5 = \text{std}(P1_i i = 1, \dots, 8)$
P6	Standard deviation of P2 for 8 profiles	Consistency	$P6 = \text{std}(P2_i i = 1, \dots, 8)$
P7	Ratio between the voxels of stem space and all voxels (space: 1/3 height, 1/2 crown diameter)	Branching density	$P7 = \frac{\sum N_{G_{\text{down } \frac{1}{3} \text{ crown \& within } \frac{1}{2} \text{ crown}}}}{\sum N_{G_{\text{crown}}}}$
P8	Ratio between the height of the equivalent centers for the whole crown and crown length	Main part of growth	$P8 = \frac{\sum_8 \left(\sum_{i=1}^n H_{G_i} / n \right)}{8 \cdot L_c}$
P9	Standard deviation of voxels for crown layers	Consistency	$P9 = \text{std}(N_{G_i \text{Layer}=1, \dots, V})$
P10	Ratio of the sum of the difference between adjacent crown layers and all voxels	Vertical consistency	$P10 = \frac{\sum_{i=1}^8 \text{Diff}(S_i, S_{i+1})}{\sum \text{Diff}(G_i, G_i)}$
P11	Ratio between the alpha volume (0.5 m) and the volume of the convex hull of crown	Branching dependence	$P11 = \frac{\text{convexhull}(P_{i t=0.5m})}{\text{convexhull}(P_{i t=\infty})}$
P12	Ratio between the similarity of two opposite and two adjacent profiles	Branching symmetry	$P12 = \frac{\sum_{i=1}^4 \text{Cor}(S_i, S_{i+4})}{\sum_{i=1}^8 \text{Cor}(S_i, S_{i+1})}$
P13	P5/P6	Consistency	$P13 = P5/P6$

Where L_c is crown length, S_i is the i th segment, N_{G_i} is the point number within the i th voxel, H_{G_i} is the height of the i th voxel, all of the voxels at the same height compose the related crown layer, R_{G_i} is the radius of the i th voxel, and n indicates the total number of the voxels for each tree.

Then, the performance about the derivation of the CI and TE feature parameters was assessed. Tree height (TH) and diameter at breast height (DBH) were deemed as the representative

variables. After correlation analysis of the estimated and reference TH sets, the performance of the derivations was quantified in terms of the coefficient of determination (R^2) and

TABLE 3 | TE feature parameters derived from the ALS point clouds by referring to the HAMs.

Abbr.	Definition	HAM-rule	Formula
Ht	Tree height (Ht)	Main part of growth	$H_t = \max(H_i) - \text{DTM}$
LcHt	Ratio between crown length (Lc) and tree height (Ht)	Consistency	$\text{LcHt} = L_c/H_t$
D _{EA}	Area-equivalent crown diameter (D _{EA})	Horizontal branch arrangement	$D_{EA} = 2 \cdot \text{sqrt}(A_{\text{Base}})$
Alpha	Alpha of Gaussian fitting in crown ellipsoid modeling	Branching symmetry	$\alpha \left \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right) \right $
LcD _{EA}	Ratio between crown length (Lc) and crown diameter	Branching symmetry	$\text{LcD}_{EA} = L_c/D_{EA}$
LlsLhls	Ratio between (tree height-LLS) and (tree height-HLS)	Main part of growth	$\text{LlsLhls} = (H_t - H_{\text{LLS}})/(H_t - H_{\text{HLS}})$
LsLcs	Ratio between LS and LCS	Main part of growth	$\text{LsLcs} = L_s/L_{cs}$
Gc	Mean height for all of the voxels	Vertical branch arrangement	$G_c = \sum_{i=1}^n H_{G_i}/n$
P _L	Laser penetration (P _L) into crown	Vertical branch arrangement	$P_L = \frac{1}{2} \cdot \left(\frac{\sum (n_{G_{YZ}} \times A_G)}{A_{CV}^{YZ}} + \frac{\sum (n_{G_{XZ}} \times A_G)}{A_{CV}^{XZ}} \right)$
LAI	Leaf area index (LAI) for all of the voxels	Leaf clustering	$\text{LAI} = \sum (n_{G_{XY}} \times A_G)/A_{CV}^{XY}$

Where L_c is crown length, A_{Base} indicates the crown base area, H_{HLS} is the height of the highest branch within crown lower surface (HLS), H_{LLS} is the height of the lowest branch within crown lower surface (LLS), L_s is the longest spread of crown cover (LS), L_{cs} is the longest cross-spread of crown cover (LCS), H_{G_i} is the height of the i th voxel, A_G is the side area of each voxel, $n_{G_{YZ}}$, $n_{G_{XZ}}$ and $n_{G_{XY}}$ each means the number of voxels in the XZ, YZ, and XY projection plane, A_{CV}^{XZ} , A_{CV}^{YZ} and A_{CV}^{XY} each means the area of the convex hull of the voxels in the XZ, YZ, and XY projection plane, respectively, and n indicates the total number of the voxels for each tree. For each tree, the variable of DTM (digital terrain model) relates to the height of the location relating to its root.

root mean squared error (RMSE), which are defined as

$$R^2 = 1 - \left(\sum (d_i - \hat{d}_i)^2 \right) / \left(\sum (d_i - \bar{d})^2 \right), \quad (1)$$

$$\text{RMSE} = \sqrt{\sum (d_i - d_i^R)^2 / n}, \quad (2)$$

where d_i is the estimated TH value for the i th tree; d_i^R is its reference TH value; their difference ($d_i - d_i^R$) is termed as TH deviation; \hat{d}_i is the TH value estimated by following the estimation-reference TH regression relationship revealed by correlation analysis; and $\bar{d} = \sum d_i / n$.

Classification Implementation

Before the classification, a non-parametric test was manipulated to evaluate the distinctions between the four tree species, in terms of all of the calculated CI and TE feature parameters. The significance of the differences between the four tree species was handled through conducting the multivariate analysis of variance (MANOVA) (<http://www.statsoft.com/Textbook/ANOVA-MANOVA#multivariate>). The specific operation was run in the Data Processing System software environment (Tang and Zhang, 2013), with the resulting p -values used to characterize the differences.

The classification of the four tree species was based on LIBSVM, i.e., a support vector machine (SVM) package (Melgani and Bruzzone, 2004; Pfeifer and Briese, 2007; Dalponte et al., 2008; Chang and Lin, 2014). The sample trees were classified by using the SVM classifier in a leave-one-out cross-validation

(LOOCV) way (Lin and Herold, 2016; Lin and Hyypä, 2016). Specifically, each tree was classified with the SVM classifier that was trained based on the remaining trees. The assessment of their accuracies was operated in terms of recall and precision, i.e., for any tree species, the ratio between the number of the correctly-classified trees and the number of its trees and the ratio between the number of the correctly-classified trees and the number of the classified-into-that-species trees, respectively.

To seek the optimal classification result, the LOOCV-SVM-based classification was deployed on all of the cases of combining the CI and TE feature parameters. The numbers of the cases that combine 1 to 13 CI feature parameters are 13, 78, 286, 715, 1287, 1716, 1716, 1287, 715, 286, 78, 13, and 1, respectively, and the numbers of the cases that combine 1 to 11 TE feature parameters are 11, 55, 165, 330, 462, 462, 330, 165, 55, 11, and 1, respectively. The resulting recalls and precisions were compared and the optimal accuracies were achieved. The CI and TE parameters corresponding to their optimal results (termed as CI_{Opt} and TE_{Opt} variables) were combined together as the input variables (equivalent to the case of [1]+[2] in the schematic program) to explore the best classification result. In addition to the recall and precision, Cohen's kappa coefficient (κ) (Cohen, 1960) was calculated to assess the performance of the classifications individually based on the CI, TE or the combination of the CI_{Opt} and TE_{Opt} feature parameters. κ measures the inter-rater agreements between categorical items. As a more robust measure than the simple percent agreement calculations like recall and precision, κ also considers the agreement occurring by chance.

HAM Identification and Information Derivation

HAM Determination and Qualitative Information Retrieval

After tree species classification, two best-performing parameters were extracted. Their related feature parameters for HAM identification included CI alone, TE alone, or the combination of CI and TE. With the assistance of expert knowledge about the HAMs in terms of these feature parameters, the distribution of CI and TE value-pairs (i.e., the combination of CI and TE) was assumed to accurately determine the HAM corresponding to each of the four tree species. Specifically, the scatterplots of those calculated value-pairs were divided into four local quadrants, in accordance to their relative ranges for the four HAMs, e.g., the Massart model relating to low P1 values and low P13 values, the Rauh model to high P1 and low P13, the Roux model to low P1 and high P13, and the Attim model to high P1 and high P13. Next, the HAM for each group of the classified trees was identified according to its related center of the value-pairs lying in which one of the local quadrants. By this means, the influences caused by some special specimens can be reduced. Based on the determined HAMs, the knowledge about tree growth habits in structure can be derived, and the specific operations are listed as follows.

Quantitative Information Derivation and Assessment

The feature parameters proposed for HAMs determination can also be assumed as the quantitative indicators of tree growth habits in structure. After the steps of tree species classification and HAMs identification, the derived feature parameters based on the classified specimens, in an isolated or integrated manner, were used to supply the information on the quantitative characteristics of tree growth habits.

The mis-classified or -identified specimens, no doubt, may introduce errors into the derived feature parameters; hence, the accuracies of the derived parameters need to be quantitatively assessed. Given that the objective of this study was to investigate tree growth habits in structure at forest plot or stand rather than individual tree levels, statistics based on Gaussian distribution were used to generate a limited number of indices, in favor of efficient accuracy assessment. For each of the prescribed feature parameter, its derived values for all of the specimens of each classified species were fitted to a Gaussian distribution function. The resulting expectation (μ) and standard deviation (σ) were compared to the derivations from the ground-reference data, and the numerical divergence was used to quantitatively evaluate the performance of the proposed method for retrievals of tree growth habits in structure.

RESULTS

ALS-Based Tree Characterization

Given that TH and DBH were selected as the representative variables to assess the derivations of the CI and TE feature parameters, their derivations from the ALS and TLS data were examined. The performance of TLS-based DBH derivations and ALS-based TH estimations are shown in **Figure 4**, with

high-value R^2 and low-value RMSEs but with linearly-varying estimation biases at different value ranges that are principally caused by the systematic errors of the scanning systems. The comparison between the TLS-derived and manually-measured DBHs can testify the capability of the used TLS for detailed tree representation, and the comparison between the ALS- and TLS-derived THs can verify the availability of the ALS for tree characterization. After all, TLS has been used for collecting the ground-truth data (Côté et al., 2012), and ALS data can derive other feature parameters like DBH via allometric relations (Hauglin et al., 2014; Calders et al., 2015). More assessments of the derivations of other structural feature parameters can refer to Holopainen et al. (2013). All of the results validated the premier of this study, i.e., the used ALS and TLS are applicable for characterizing the basic structures of the sample trees.

Tree Species Classification

We first derived the proposed CI feature parameters (Figure S1), based on which the MANOVA p -value was calculated to be 0.0001 (Table S1). This meant significant differences exist between the four tree species, but the CI-based classification results (Figure S1) told that using any one of the 13 CI parameters individually cannot fulfill distinguishing the four tree species. Then, classifications were conducted for all of the cases of combining 1–13 CI parameters (Figure S2), which relates to the operations of the routine [1] marked in **Figure 2**. The overall performance firstly went better and then leveled off. The highest classification accuracy is 82.5% ($\kappa = 0.794$), in the case of inputting the five CI parameters—P3, P4, P6, P7, and P8—as the dependent variables into the classification (Table S2).

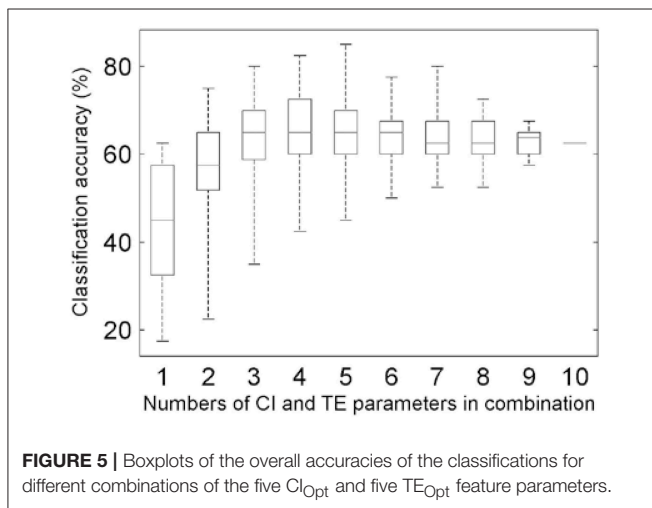
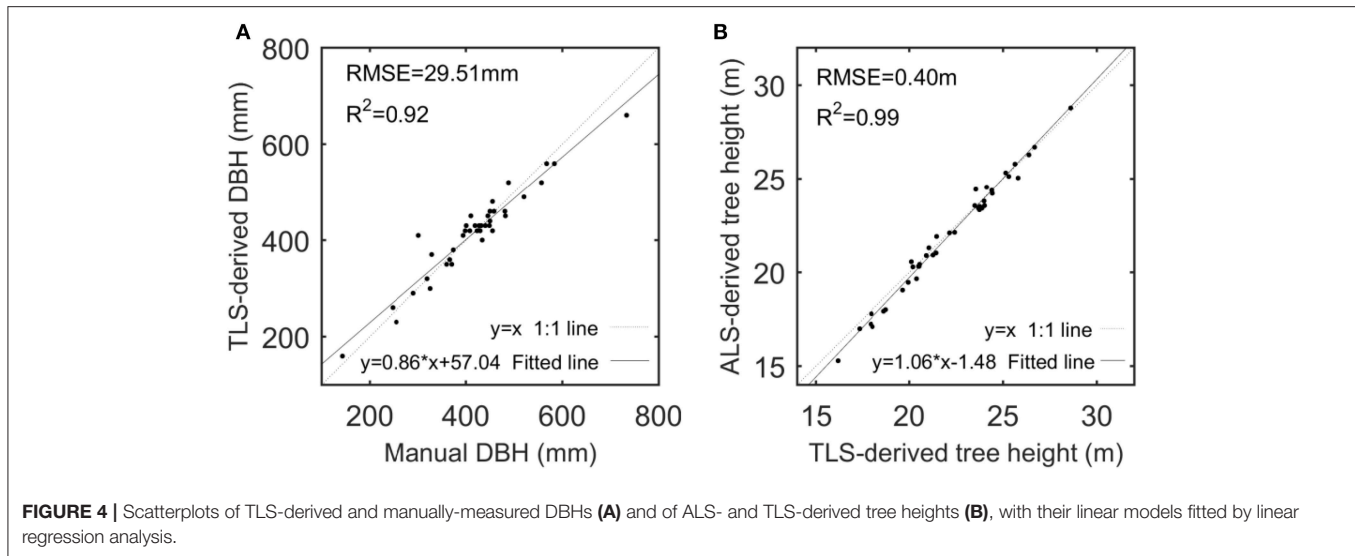
Then, the TE parameters were derived (Figure S3), and their related MANOVA p -value was 0.0001 (Table S3). The classifications for all of the cases of combining 1–10 TE structure parameters were carried out (Figure S4), which corresponds to the routine [2] in **Figure 2**. The highest classification accuracy is 85.0% (Table S4), in the case of combining the LcHt, LcDEA, Gc, DEA, and σ TE parameters ($\kappa = 0.764$).

The combinations of the above-sought CI_{Opt} and TE_{Opt} parameters were input to implement the classification (MANOVA p -value = 0.0001; Table S5). The accuracies for all of the cases of combining the optimal five CI parameters and the optimal five TE parameters were calculated (**Figure 5**), and this step relates to the routine [1]+[2] in **Figure 2**. The highest accuracy is 85.0%, which corresponds to the optimal case of TE-based classification ($\kappa = 0.794$; Table S6). Further, for the task of searching two feature parameters to determine the appropriate HAMs for the four tree species, the best case of combining the CI-typed parameters of P4 and P7 was found, exhibiting the highest classification accuracy of 75.0% ($\kappa = 0.658$).

HAM Determination and Information Retrieval

Qualitative Derivation

After tree species classifications, the HAM for each of the four tree species was identified based on the expert knowledge, assisted by referring to the distributions of the two P4 and



P7 variables in pairs (Figure 6). Specifically, for the four tree species, the centers of their P4–P7 pair distributions were derived, and a comparison of these centers to the opposites calculated by referring to the manually-trained crown-HAM links was operated to determine the HAM for each of the four tree species. The links include that the Massart model relates to low P4 and low P7 values (corresponding to the PA species), the Rauh model exhibits high P4 and low P7 (PS), the Roux model shows moderate P4 and low P7 (PT), and the Attim model presents low P4 and high P7 (QR). By this means, the *a-priori* knowledge about tree growth habits in structure was acquired for the four tree species (Table 4).

Quantitative Derivation

The information about tree growth habits in structure, further, was quantitatively derived, as illustrated by the boxplots of the derived parameters P4 and P7 (Figure 7). The biases

between the boxplots show the distinctions between the four tree species. For each tree species, the derived parameters are approximate to their values from the ground-truth data, as evidenced by their comparative statistics (Table 5). Specifically, their absolute deviations ($|\Delta\mu|$) of the expectations of the fitted Gaussian distribution before and after the proposed tree species classification and HAM identification are $<7\%$ of the actual values for any of the four tree species, and their absolute deviations ($|\Delta\sigma|$) of the standard deviations (σ) are all $<16\%$ of the reference values. These results suggested that based on the proposed method, the quantitative characteristics about tree growth habits in structure for the four tree species can be derived for the trees at plot or stand scales.

DISCUSSION AND CONCLUSION

The results of the case study based on the four tree species growing in the boreal environment validated the hypothesis that the recruitment of conventional descriptive tree architecture models such as the HAMs into the state-of-the-art ALS-based forest mapping facilitates more efficiently deriving the typically hard-to-discern tree growth habits in structure. The contributions of this study are 2-fold. With ALS data used as a bridge, the application range of the HAMs can be expanded to the boreal forests; meanwhile, with the HAMs used as a link, the application discipline of ALS can be extended to tree growth habits in structure. In other words, the answer to the question posed in the Introduction, no doubt, is “both.” Overall, this study developed a novel pan-solution strategy, i.e., by recruiting the previously-established tree architectural, physiological, and/or biophysicochemical models into the cutting-edge forest mapping technologies, for more efficiently understanding various forest growth habits, which facilitate investigating forest ecologies and projecting forest dynamics and processes under climate changes.

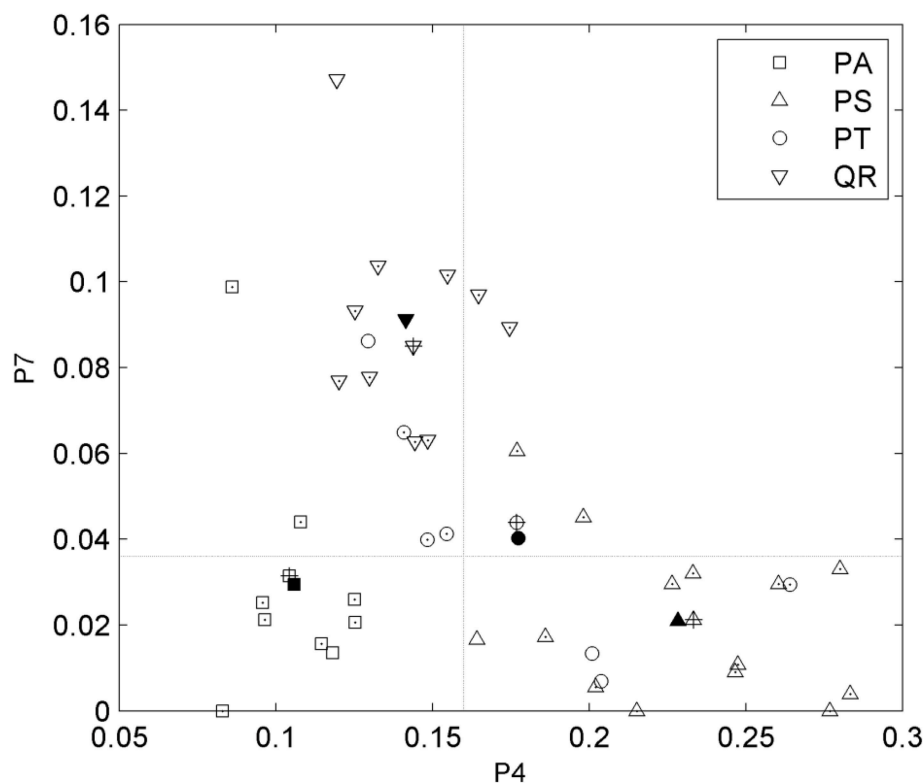


FIGURE 6 | Scatterplots of the P4 and P7 feature parameters for the four tree species and their value-pair distributions for determining the related HAMs. Markers with points indicate the correctly-classified tree specimens; markers filled in black indicate the centers of the parameter-pair distributions for all tree specimens; markers with pluses indicate the centers of the parameter-pair distributions for the correctly-classified tree specimens; and two dash lines segment the four local quadrants.

TABLE 4 | Tree growth habits in structure derived for the four tree species.

Tree specimens	Trunk	Branches	Growth
<i>Pinus sylvestris</i>	Monopodial, indeterminate	Plagiotropic, (main branches) produced in whorls	Rhythmic
<i>Populus tremula</i>	Monopodial	Orthotropic, morphogenetically equivalent to the trunk	Rhythmic
<i>Quercus robur</i>	Monopodial, indeterminate	Plagiotropic, monopodial, non-phyllomorphic	Continuous
<i>Picea abies</i>	Monopodial	Orthotropic, morphogenetically equivalent to the trunk	More or less continuous

Application Extending

First, the implications of this study require more elucidation, because the results in **Table 5** may be misleading, i.e., the results appear simple. Theoretically, the HAM related to each individual tree can be identified by the associated experts in the field (Negrón et al., 2013), and in such a sense, the development of a complicated method as shown in this study seems unnecessary. This reasoning, however, is only rooted in the results that were derived from the limited number of sample trees, without

taking the broader-sense context of this study into account. This study tried to advance the field of forest mapping based on ALS, which generally can cover large areas (Ferraz et al., 2012). In this situation, manual determination of the HAM for each single tree is unadvisable. Besides, the concerned targets in practical ALS-based forest mapping tend to be tree communities, seldom single trees (Koetz et al., 2007). The HAMs can give descriptive “plagiotropic or not,” instead of quantitative “plagiotropic degree,” and they are useful for learning the whole structural traits of communities. Hence, the combination of ALS and the HAMs can help people to have a sketchy understanding of tree structural attributes for forest stands; namely, the first implication of this work is to present a solution of reflecting large-scale properties of forest dynamics, further advancing studies on plant geography and spatial ecology.

Then, for each specimen of the four tree species measured by the ALS, its HAM identification means not only grasping its growth habits in structure but also learning the rules for its symbolic delineation and graphical representation (Prusinkiewicz and Remphrey, 2000). Specifically, the proposed feature parameters can be adopted as the geometric/topologic restrictions to briefly determine the sequencing of apices and the configuration of branches; then, tree geometric models can be derived from ALS data, a difficult task for long but of extensive interest (Calders et al., 2013). The developed tree geometric

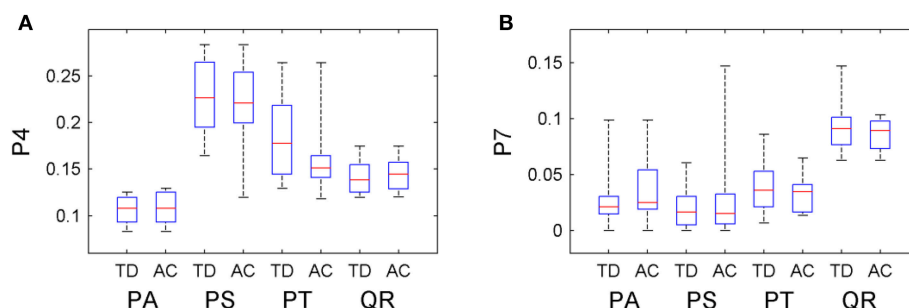


FIGURE 7 | Boxplots of the (A) P4 and (B) P7 CI parameter values after the classification and HAM identification (AC), compared with their values derived from the ground-truth data (TD).

TABLE 5 | Derivations of the (a) P4 and (b) P7 CI parameter (DP) values after the classification and HAM identification, compared with their values in the ground-truth (GT) data.

		DP		GT		Ratio (%)	
		μ	σ	μ	σ	$\Delta\mu$	$\Delta\sigma$
P4	PA	0.1058	0.0161	0.1071	0.0176	1.30	1.50
	PS	0.2280	0.0409	0.2224	0.0437	-2.31	1.12
	PT	0.1844	0.0485	0.1650	0.0510	-6.65	0.83
	QR	0.1415	0.0191	0.1439	0.0186	2.12	-0.49
P7	PA	0.0294	0.0285	0.0375	0.0333	4.72	2.82
	PS	0.0201	0.0185	0.0277	0.0362	6.89	15.89
	PT	0.0393	0.0260	0.0343	0.0189	-3.19	-4.55
	QR	0.0912	0.0245	0.0850	0.0156	-4.22	-6.06

More information is listed in the Supplementary Data file.

models can be effectively in favor of the simulation-based analyses of forest-environment interaction effects (Lovelock et al., 2006). The botanic knowledge derived from the HAMs, further, can help to push forward the stage of tree geometric modeling to tree comprehensive biophysical modeling for better understanding forest physiologies.

Influence Analysis

The performance of the proposed method is restricted to many influence factors, and analyses indicated that there are three kinds of primary ones. First, the calculated values of the proposed feature parameters were unavoidably attached with errors. The sources of these errors may involve biased geometric modeling, incomplete structure representations, and uncertainties of parameter derivations. When the errors are large enough, such feature parameters may play negative roles in tree species classification (Lin and Herold, 2016). This limitation is evidenced by the maximum total accuracy, which initially improved and then leveled off, as indicated by the top whiskers of the boxes in **Figure 5**. In other words, this effect occurs in the cases involving the CI and TE parameters separately (Supplementary Figures S2 and S4) and their optimal-accuracy-related combination (**Figure 5**), respectively.

Second, the proposed feature parameters maybe cannot efficiently reflect some particular features of the tree species of interest. This influence factor is exemplified by the *P. tremula* trees, which showed relatively lower accuracies in all of the three steps of determining their HAMs in comparison to the other three tree species. In fact, as illustrated in **Figure 1**, the ALS-represented morphologies of the *P. tremula* tree and the *Q. robur* tree are similar, and it is difficult to distinguish them even by visual interpretation. This just verifies the hypothesis that the feature parameters proposed in this study may better present the characteristics of *Q. robur* than *P. tremula*.

Third, one tree species may not be strictly related to one HAM. At its different growth stages, a tree may satisfy different HAMs. Some tree species show different developments in HAMs during their ontogeny (Millet et al., 1998). Some trees may behave as satisfying multiple HAMs. Although this diversity issue can be somehow handled in the context of ALS-based large-area forest mapping as mentioned above, errors may be introduced into the results of tree species classification. This suggests that this problem may limit extending the application range of the proposed method; in addition, in the following studies more attention shall be paid onto the CI feature parameters, which were proposed by following the characteristics of HAMs.

Potential Improvements

Note that although the derivations were on the basis that the used ALS system can represent trees in a relatively complete manner (as illustrated in **Figure 1**), the proposed CI and TE feature parameters are far from enough for representing the real structures of the targeted trees, let alone the scenarios of other tree species and other HAMs regarded. Compared to the parameters considered in this study, multiple other structural feature parameters can be extracted by exploring other morphological or topological features of the trees. For instance, the spatial distribution mode of the primary branch centers (Lin and Herold, 2016) can be extracted to supply more information. Simultaneously, the proposal of other structural parameters needs also to regard the characteristics of ALS-based tree representations. When a crown is dense, less information is available from the related trunk (Puttonen

et al., 2010); therefore, the parameters derived from such trunks cannot be directly compared in some cases. With such detailed influence factors as many as possible considered, the results will be improved. In addition, strengthening the automatic degree in deriving these feature parameters is highlighted.

The proposed approach for identification of the HAMs for different tree species did help to exploit the unready-observable tree growth habits in structure hidden in the ALS-collected point clouds, but this approach is still principally under qualitative interpretation. In other words, the derived knowledge of tree growth habits cannot be quantitatively used for precision managements of forests. Subsequent studies need to emphasize how to propose new feature parameters and to attempt other kinds of tree architecture models (e.g., Weber and Penn, 1995) for quantitative characterization of tree growth habits.

The used ALS system is a currently-pervasive airborne LiDAR, and its sampling density can reflect the main-stream laser scanning performance. Hence, the proposed method shall be highly applicable for practical applications. Nevertheless, many ALS systems produced at earlier times are still in use, and consequently, the sensitivity of the proposed method in response to laser point density needs to be further examined in the subsequent work, e.g., via checking its performance based on the ALS data of different sampling densities. The feasibility of the proposed method for different tree ages also needs to be examined, e.g., via applying it to the stands at varying growth stages and comparing its performance relative to different morphologies of trees.

REFERENCES

- Barthélémy, D., and Caraglio, Y. (2007). Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Ann. Bot.* 99, 375–407. doi: 10.1093/aob/mcl260
- Calders, K., Lewis, P., Disney, M., Verbesselt, J., and Herold, M. (2013). Investigating assumptions of crown archetypes for modelling LiDAR returns. *Remote Sens. Environ.* 134, 39–49. doi: 10.1016/j.rse.2013.02.018
- Calders, K., Newnham, G., Burt, A., Raunonen, P., Herold, M., Culvenor, D., et al. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods Ecol. Evol.* 6, 198–208. doi: 10.1111/2041-210X.12301
- Chang, C. C., and Lin, C. J. (2014). *LIBSVM: A Library for Support Vector Machines*. Available online at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (Accessed June 12, 2014).
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104
- Costes, E., and Guédon, Y. (2002). Modelling branching patterns on 1-year-old trunks of six apple cultivars. *Ann. Bot.* 89, 513–524. doi: 10.1093/aob/mcf078
- Costes, E., Lauri, P. E., and Regnard, J. L. (2006). Analyzing fruit tree architecture: implications for tree management and fruit production. *Hortic. Rev.* 32, 1–61. doi: 10.1002/9780470767986.ch1
- Côté J. F., Fournier, R. A., and Egli, R. (2011). An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. *Environ. Model. Softw.* 26, 761–777. doi: 10.1016/j.envsoft.2010.12.008
- Côté J. F., Fournier, R. A., Frazer, G. W., and Niemann, K. O. (2012). A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. *Agric. For. Meteorol.* 166, 72–85. doi: 10.1016/j.agrformet.2012.06.007
- Creber, G. T., and Chaloner, W. G. (1984). Influence of environmental factors on the wood structure of living and fossil trees. *Bot. Rev.* 50, 357–448. doi: 10.1007/BF02862630

AUTHOR CONTRIBUTIONS

YL: designed this study and developed the method; MJ: made the data analysis; PP: analyzed the results and revised the manuscript; JH: reviewed and revised the manuscript.

FUNDING

This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 31670718, 41471281, and 41401515) and in part by the SRF for ROCS, SEM, China.

ACKNOWLEDGMENTS

Thank the Mobile Mapping group at the Finnish Geospatial Research Institute for distributing the data. This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 31670718, 41471281, and 41401515) and in part by the SRF for ROCS, SEM, China. The cooperation was made possible by China SAFEA (State Administration of Foreign Experts Affairs) supported program (Grant No: 20161100067).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00220/full#supplementary-material>

- Dalponte, M., Bruzzone, L., and Gianelle, D. (2008). Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas. *IEEE Trans. Geosci. Rem. Sens.* 46, 1416–1427. doi: 10.1109/TGRS.2008.916480
- Davidson, C., Prusinkiewicz, P., and von Aderkas, P. (2008). Description of a novel organ in the gametophyte of the fern *Schizaea pusilla* and its contribution to overall plant architecture. *Botany* 86, 1217–1223. doi: 10.1139/B08-085
- Eitel, J., Vierling, L., and Magney, T. (2013). A lightweight, low cost autonomously operating terrestrial laser scanner for quantifying and monitoring ecosystem structural dynamics. *Agric. For. Meteorol.* 180, 86–96. doi: 10.1016/j.agrformet.2013.05.012
- Feng, L., de Reffye, P., Dreyfus, P., and Auclair, D. (2012). Connecting an architectural plant model to a forest stand dynamics model - application to Austrian black pine stand visualization. *Ann. For. Sci.* 69, 245–255. doi: 10.1007/s13595-011-0144-5
- Ferraz, A., Bretar, F., Jacquemoud, S., Gonçalves, G., Pereira, L., Tomé M., et al. (2012). 3-D mapping of a multi-layered mediterranean forest using ALS data. *Remote Sens. Environ.* 121, 210–223. doi: 10.1016/j.rse.2012.01.020
- Fisher, J. B. (1984). “Tree architecture: relationships between structure and function,” in *Contemporary Problems in Plant Anatomy*, eds R. A. White and W. C. Dickison (Orlando, FL: Academic Press), 541–589.
- Guédon, Y., Barthélémy, D., Caraglio, Y., and Costes, E. (2001). Pattern analysis in branching and auxiliary flowering sequences. *J. Theor. Biol.* 212, 481–520. doi: 10.1006/jtbi.2001.2392
- Hallé, F., Oldeman, R. A. A., and Tomlinson, P. B. (1978). *Tropical Trees and Forests: An Architectural Analysis*. Berlin: Springer-Verlag.
- Hauglin, M., Gobakken, T., Astrup, R., Ene, L., and Næsset, E. (2014). Estimating single-tree crown biomass of Norway Spruce by airborne laser scanning: a comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. *Forests* 5, 384–403. doi: 10.3390/f5030384

- Heinzel, J., and Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *Int. J. Appl. Earth Observ. Geoinform.* 18, 101–110. doi: 10.1016/j.jag.2012.01.025
- Holopainen, M., Kankare, V., Vastaranta, M., Liang, X., Lin, Y., Vaaja, M., et al. (2013). Tree mapping using airborne, terrestrial and mobile laser scanning – a case study in a heterogeneous urban forest. *Urban For. Urban Green.* 12, 546–553. doi: 10.1016/j.ufug.2013.06.002
- Hosoi, F., and Omasa, K. (2007). Factors contributing to accuracy in the estimation of the woody canopy leaf area density profile using 3D portable lidar imaging. *J. Exp. Bot.* 58, 3463–3473. doi: 10.1093/jxb/erm203
- Hyde, P., Dubayah, R., Peterson, B., Blair, J. B., Hofton, M., Hunsaker, C., et al. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: validation of montane ecosystems. *Remote Sens. Environ.* 96, 427–437. doi: 10.1016/j.rse.2005.03.005
- Ko, C., Sohn, G., and Rimmel, T. K. (2013). Tree genera classification with geometric features from high-density airborne LiDAR. *Can. J. Remote Sens.* 39, S73–S85. doi: 10.5589/m13-024
- Koetz, B., Sun, G., Morsdorf, F., Ranson, K. J., Kneubühler, M., Itten, K., et al. (2007). Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization. *Remote Sens. Environ.* 106, 449–459. doi: 10.1016/j.rse.2006.09.013
- Lehtomäki, M., Jaakkola, A., Hyypä J., Kukko, A., and Kaartinen, H. (2010). Detection of vertical pole-like objects in a road environment using vehicle-based laser scanning data. *Remote Sens.* 2, 641–664. doi: 10.3390/rs2030641
- Liang, X., and Hyypä J. (2013). Automatic stem mapping by merging several terrestrial laser scans at the feature and decision levels. *Sensors* 13, 1614–1634. doi: 10.3390/s130201614
- Lin, Y., and Herold, M. (2016). Tree species classification based on explicit tree structure feature parameters derived from static terrestrial laser scanning data. *Agric. For. Meteorol.* 216, 105–114. doi: 10.1016/j.agrformet.2015.10.008
- Lin, Y., and Hyypä, J. (2016). A comprehensive but efficient framework of proposing and validating feature parameters from airborne LiDAR data for tree species classification. *Int. J. Appl. Earth Observ. Geoinform.* 46, 45–55. doi: 10.1016/j.jag.2015.11.010
- Lin, Y., Hyypä, J., Kukko, A., Jaakkola, A., and Kaartinen, H. (2012). Tree height growth measurement with single-scan airborne, static terrestrial and mobile laser scanning. *Sensors* 12, 12798–12813. doi: 10.3390/s120912798
- Lovelock, C. E., Ball, M. C., Choat, B., Engelbrecht, B. M., Hobbrook, N. M., and Feller, I. C. (2006). Linking physiological processes with mangrove forest structure: phosphorus deficiency limits canopy development, hydraulic conductivity and photosynthetic carbon gain in dwarf Rhizophora mangle. *Plant Cell Environ.* 29, 793–802. doi: 10.1111/j.1365-3040.2005.01446.x
- Melgani, F., and Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* 42, 1778–1790. doi: 10.1109/TGRS.2004.831865
- Millet, F., Bouchard, A., and Edelin, C. (1998). Plagiotropic architectural development and successional status of four tree species of the temperate forest. *Can. J. Remote Sens.* 76, 2100–2118.
- Moorthy, I., Miller, J. R., Berni, J. A. J., Zarco-Tejada, P., Hu, B., and Chen, J. (2011). Field characterization of olive (*Olea europaea* L.) tree crown architecture using terrestrial laser scanning data. *Agric. For. Meteorol.* 151, 204–214. doi: 10.1016/j.agrformet.2010.10.005
- Mutke, S., Sievänen, R., Nikinmaa, E., Perttunen, J., and Gil, L. (2005). Crown architecture of grafted Stone pine (*Pinus pinea* L.): shoot growth and bud differentiation. *Trees* 19, 15–25. doi: 10.1007/s00468-004-0346-7
- Negrón, C., Contador, L., Lampinen, B. D., Metcalf, S. G., and DeJong, T. M. (2013). Systematic analysis of branching patterns of three almond cultivars with different tree architectures. *J. Am. Soc. Hortic. Sci.* 138, 407–415.
- Pfeifer, N., and Briese, C. (2007). Geometrical aspects of airborne laser scanning and terrestrial laser scanning. *Int. Arch. Photogramm. Remote Sens.* 36, 311–319.
- Pinard, M. A., Putz, F. E., Rumiz, D., Guzman, R., and Jardim, A. (1999). Ecological characterization of tree species for guiding forest management decisions in seasonally dry forests in Lomerio, Bolivia. *For. Ecol. Manage.* 113, 201–213. doi: 10.1016/S0378-1127(98)00426-5
- Pretzsch, H., Utschig, H., and Sadtke, R. (2006). “Applications of tree growth modelling in decision support for sustainable forest management,” in *Sustainable Forest Management*, ed H. Hasenauer (Berlin: Springer-Verlag), 131–149.
- Prusinkiewicz, P., and Remphrey, W. (2000). “Characterization of architectural tree models using L-system and Petri nets,” in *Larbre - The Tree 2000: Papers presented at the 4th International Symposium on the Tree*, ed M. Labrecque, 177–186.
- Puttonen, E., Suomalainen, J., Hakala, T., Räikkönen, E., Kaartinen, H., Kaasalainen, S., et al. (2010). Tree species classification from fused active hyperspectral reflectance and LIDAR measurements. *For. Ecol. Manage.* 260, 1843–1852. doi: 10.1016/j.foreco.2010.08.031
- Renton, M., Guédon, Y., Godin, C., and Costes, E. (2006). Similarities and gradients in growth unit branching patterns during ontogeny in ‘Fuji’ apple trees: a stochastic approach. *J. Exp. Bot.* 57, 3131–3143. doi: 10.1093/jxb/erl075
- Saxe, H., Cannell, M., Johnsen, Ø., Ryan, M., and Vourlitis, G. (2001). Tree and forest functioning in response to global warming. *New Phytol.* 149, 369–400. doi: 10.1046/j.1469-8137.2001.00057.x
- Shepherd, T. G. (2016). Effects of a warming Arctic. *Science* 353, 989–990. doi: 10.1126/science.aag2349
- Tang, Q. Y., and Zhang, C. X. (2013). Data processing system (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research. *Insect Sci.* 20, 254–260. doi: 10.1111/j.1744-7917.2012.01519.x
- Wang, C., and Glenn, N. F. (2008). A linear regression method for tree canopy height estimation using airborne lidar data. *Can. J. Remote Sens.* 34, S217–S227. doi: 10.5589/m08-043
- Weber, J., and Penn, J. (1995). “Creation and rendering of realistic trees,” in *ACM International Conference on Computer Graphics and Interactive Techniques* (New York, NY), 119–128.
- Yang, X., Strahler, A., Schaff, C., Jupp, D., Yao, T., Zhao, F., et al. (2013). Three-dimensional forest reconstruction and structural parameter retrievals using a ground-based full-waveform lidar instrument (Echidna®). *Remote Sens. Environ.* 135, 36–51. doi: 10.1016/j.rse.2013.03.020
- Yu, X., Hyypä J., Kukko, A., Maltamo, M., and Kaartinen, H. (2006). Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogramm. Eng. Remote Sens.* 72, 1339–1348. doi: 10.14358/PERS.72.12.1339

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lin, Jiang, Pellikka and Heiskanen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Not All Trees Sleep the Same—High Temporal Resolution Terrestrial Laser Scanning Shows Differences in Nocturnal Plant Movement

András Zlinszky^{1,2*}, Bence Molnár³ and Anders S. Barfod²

¹ Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences, Tihany, Hungary,

² Ecoinformatics and Biodiversity Section, Department of Bioscience, Aarhus University, Aarhus, Denmark, ³ Department of Photogrammetry and Geoinformatics, Budapest University of Technology and Economics, Budapest, Hungary

OPEN ACCESS

Edited by:

Yann Guédon,

Centre de Coopération Internationale
en Recherche Agronomique pour le
Développement (CIRAD), France

Reviewed by:

Franck Hétroy-Wheeler,

Université de Strasbourg, France

Frédéric Boudon,

Centre de Coopération Internationale
en Recherche Agronomique pour le
Développement (CIRAD), France

*Correspondence:

András Zlinszky

zlinszky.andras@okologia.mta.hu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 09 June 2017

Accepted: 05 October 2017

Published: 20 October 2017

Citation:

Zlinszky A, Molnár B and Barfod AS
(2017) Not All Trees Sleep the
Same—High Temporal Resolution
Terrestrial Laser Scanning Shows
Differences in Nocturnal Plant
Movement. *Front. Plant Sci.* 8:1814.
doi: 10.3389/fpls.2017.01814

Circadian leaf movements are widely known in plants, but nocturnal movement of tree branches were only recently discovered by using terrestrial laser scanning (TLS), a high resolution three-dimensional surveying technique. TLS uses a pulsed laser emitted in a regular scan pattern for rapid measurement of distances to the targets, thus producing three dimensional point cloud models of sub-centimeter resolution and accuracy in a few minutes. Here, we aim to gain an overview of the variability of circadian movement of small trees across different taxonomic groups, growth forms and leaf anatomies. We surveyed a series of 18 full scans over a 12-h night period to measure nocturnal changes in shape simultaneously for an experimental setup of 22 plants representing different species. Resulting point clouds were evaluated by comparing changes in height percentiles of laser scanning points belonging to the canopy. Changes in crown shape were observed for all studied trees, but clearly distinguishable sleep movements are apparently rare. Ambient light conditions were continuously dark between sunset (7:30 p.m.) and sunrise (6:00 a.m.), but most changes in movement direction occurred during this period, thus most of the recorded changes in crown shape were probably not controlled by ambient light. The highest movement amplitudes, for periodic circadian movement around 2 cm were observed for *Aesculus* and *Acer*, compared to non-periodic continuous change in shape of 5 cm for *Gleditschia* and 2 cm for *Fargesia*. In several species we detected 2–4 h cycles of minor crown movement of 0.5–1 cm, which is close to the limit of our measurement accuracy. We present a conceptual framework for interpreting observed changes as a combination of circadian rhythm with a period close to 12 h, short-term oscillation repeated every 2–4 h, aperiodic continuous movement in one direction and measurement noise which we assume to be random. Observed movement patterns are interpreted within this framework, and connections with morphology and taxonomy are proposed. We confirm the existence of overnight “sleep” movement for some trees, but conclude that circadian movement is a variable phenomenon in plants, probably controlled by a complex combination of anatomical, physiological, and morphological factors.

Keywords: circadian plant movement, laser scanning, tree canopy movement, 3-dimensional modeling, tree physiology, turgor, chronobiology, circadian rhythm

INTRODUCTION

Circadian movement of leaves includes the process of nyctinasty, already known to the ancient Greeks (Bretzl, 1903) but first described by Darwin (Darwin and Darwin, 1880). The leaves of many shrubs and trees actively change position during the night. In some fabaceous trees this phenomenon is particularly striking. In a recent series of experiments by Puttonen et al. (2016) terrestrial laser scanning (TLS) was applied to detect well-defined circadian movement rhythm of the branches and the trunk of the temperate tree *Betula pendula*. TLS (Dassot et al., 2011) allows for detailed measurement of the distance from sensor to a target object by measurement of the travel time of short laser pulses emitted from a given position in a known direction. Since the measurements are conducted during very short time (in the order of milliseconds), a high repetition rate is possible, which together with an equidistant scan pattern allows regular sampling of the target objects. The output of such a measurement is a set of points with positioning coordinates and the amplitude of the reflected signal as an attribute. By employing TLS, Puttonen et al. (2016) created detailed 3-dimensional models of trees at hourly intervals during the night. Their results show that tree branches moved up to 8 cm downwards during the night and returned to their original position by daybreak. By scanning two *Betula* trees nearly 2,000 km from each other, they proved that this phenomenon is probably an attribute which is unique to individual plant species. This seminal discovery opened up many new questions. What is the adaptive advantage of sleep movement (as Darwin called it and as we will also call nocturnal changes in plant shape)? What are the processes governing sleep movements and how are these linked to ambient light, temperature, humidity, and soil water content? Are the movements driven by an internal circadian clock or induced externally by photoperiodism? Are patterns of sleep movement characteristic for certain taxonomic groups?

Our aim is to gain an overview of the variability of plant circadian movement in seed plants. Our method measures movement as the concerted displacement of branches and leaves. Under conditions monitored for light, humidity, temperature, and (lack of) wind, we surveyed a wide taxonomic range of trees and shrubs grown in containers. The experiment was set up to compare sleep movement across taxonomic groups and different types of phenology and morphology. Advancing the methodology of nocturnal tree movement quantification was outside the scope of this mainly qualitative study.

MATERIALS AND METHODS

Measurement Setup

Timing

The study was carried out on 16–17 September 2016, close to the autumn solstice when day and night are equally long. The first laser scanning was conducted at 6:30 p.m. and repeated every 30 min until 8:00 p.m. Hereafter, the scans were repeated every hour until 4:00 a.m. the next day. Finally a series of half-hourly scans were conducted until 7:00 a.m.

Location and General Environment

The experiment was carried out at a gardening shop at Balatonfüred, Western Hungary (N46°56'16.91"E17°51'42.48"), which has developed an open-air nursery for small-sized container grown woody plants. The shrubs and young trees were kept for at least half a year in the nursery, where they were watered on a daily basis. All the plants included in this study have thus been exposed to exactly the same weather and water conditions for at least 6 months. To avoid the expected effect of wind action we conducted the experiment in a greenhouse under construction, which had a glass roof, but no walls. During the full duration of the experiment weather conditions were stable with no wind and no precipitation or condensation. The nearest streetlights were more than 1 km away. During the overnight experiment we made great effort to exclude all outdoor lights and restrict the use of torch light to a minimum. A full moon allowed us to operate the scanner and read the instruments without an additional light source.

Experiment Design and Studied Plants

The choice of plants was somewhat restricted to what was available to us at site. Nevertheless, we were able to include in the experiment a broad range of growth forms (trees, shrubs, bamboos, and palms), leaf phenologies (evergreen and deciduous), leaf morphologies (entire, dissected, needle-shaped, and scaly), taxonomic groups, and ecoclimatic ranges (tropical, subtropical, temperate). Although high priority was given to include tall and dense plants, the height was restricted by the size of the containers. For most species we were unable to include more than one individual.

Measurement Techniques

Ambient Sensors and Instruments

Ambient air temperature and humidity were recorded by placing six climate loggers on the trunks of six selected trees and near the LIDAR sensor (Easylog USB™, Lascar Electronics®, UK), and temperature loggers (HOBO Pendant™) in the soil containers of the same trees. These were set to collect data every 10 s. Wind speed was measured before each laser scanning measurement cycle using a simple hand-held ball anemometer (Dwyer Instruments, Michigan City, IND, US). At the same time ambient light was recorded using a LabSphere™ integrating sphere mounted on a Li-Cor™ data logger, within the greenhouse frame as well as outside.

Terrestrial Laser Scanning Instrument and Measurement Parameters

Laser scanning can produce a high-density point cloud that allows for dynamic 3D modeling of complex shapes in few minutes. For 3D data acquisition, a phase-based Terrestrial Laser Scanner (TLS) was applied. FARO™ Focus S120 (FARO Technologies, Lake Mary, FL, US) operates a Class 3 laser with 905 nm wavelength. Acquisition parameters were set to capture 7×10^9 points in 15 min with highest quality. Based on instruments data sheet (FARO, 2013), the distance measurement has 0.95 mm accuracy on a solid, well-defined target while angular deviation has ± 2 mm effect on accuracy at 25 m

Vernacular name	Scientific name	Height of studied specimen [m]	Crown diameter of studied specimen [m]	Taxonomy (Chase et al., 2016)	Origin	Phytoclimatic region	Phenology
Nootka cypress	<i>Chamaecyparis nootkatensis</i> (D.Don) Spach	2.75	2.2	Gymnosperm	NE US & Pacific Canada	Boreal	Evergreen
Japanese larch	<i>Larix kaempferi</i> (Lamb.) Carr.	2.8	0.9	Gymnosperm	SE Asia	Subalpine	Deciduous
Warty birch	<i>Betula pendula</i> Roth.	3.72	1	Fabid	Europe & SW Asia	Nemoral to boreal	Deciduous
Himalayan birch	<i>Betula utilis</i> D. Don	1.35	1.1	Fabid	Himalayas	Alpine	Deciduous
Southern magnolia	<i>Magnolia grandiflora</i> L.	2.75	1.4	Basal angiosp.	SE US	Subtropical	Evergreen
Austrian pine	<i>Pinus nigra</i> Arnold	1.9, 3.7 (two trees)	1.1, 1.9 (two trees)	Gymnosperm	Europe	Subalpine	Evergreen
Cedar of Libanon	<i>Cedrus libani</i> A. Rich	2.75	1.1	Gymnosperm	E. Mediterran.	Subalpine	Evergreen
Japanese maple	<i>Acer palmatum</i> Thunb.	3.8	2	Malvid	NE Asia	Temperate	Deciduous
Horse chestnut	<i>Aesculus hippocastanum</i> L.	3.75	1.1	Malvid	SE Europe	Temperate	Deciduous
White mulberry	<i>Morus alba</i> L.	2.75	1.5	Fabid	N China	Subtropical to temperate	Deciduous
Pedunculate oak	<i>Quercus robur</i> L.	3.25	1.3	Fabid	Europe	Nemoral	Deciduous
Chokecherry	<i>Prunus virginiana</i> L.	3.2	0.9	Fabid	N America	Temperate	Semi-evergreen
Olive	<i>Olea europea</i> L.	2.6	2.3	Lamid	E. Mediterran.	Subtropical	evergreen
Prage viburnum	<i>Viburnum x pragense</i>	1.5	1.5	Campanulid	n.a.	n.a.	semi- evergreen
Persian silk tree	<i>Albizia julibrissin</i> Durazz.	3.5	1.8	Fabid	SW & E Asia	Subtropical	deciduous
Honey locust	<i>Gleditsia triacanthos</i> L.	3.5	1.7	Fabid	N America	Subtropical to temperate	deciduous
Oleander	<i>Nerium oleander</i> L.	1.75	1.95	Lamid	SW Asia	Subtropical	evergreen
Pagoda tree	<i>Styphnolobium japonicum</i> (L.) Schott (syn. <i>Sophora japonica</i> L.)	1.8	0.7	Fabid	China	Subtropical to temperate	deciduous to evergreen
Red tip photinia	<i>Photinia x fraseri</i>	2.7	1.4	Fabid	n.a.	n.a.	semi- evergreen
Umbrella bamboo	<i>Fargesia murielae</i> T.P.Yi.	2.65	1	Monocot	China	Temperate to subtropical	evergreen
European fan palm	<i>Trachycarpus fortunei</i> (Hook)	2.3	2	Monocot	W. Mediterran.	Subtropical	evergreen

object distance (Note: in our case the furthest tree had an object distance of 15 m). Each raw scan requires 3 GB storage space, however, only ~10% of the points are representing trees. Therefore, the first step of post-processing was to reduce the amount of data manually based on object distance and direction, removing all points that were outside the greenhouse frame. Reducing data amount allowed efficient data handling and comparison. The second step was to co-register the time series of the measurements. All measurements were captured from the same scanner position; however, the scanner battery and storage device (SD card) had to be replaced during the measurement session. Because this could potentially cause minimal changes in the scanner position relative to the object, we choose target based registration with reference spheres, but also including

the steel girder frame of the greenhouse as an additional source of reference surfaces. Altogether 18 scans were processed; maximum mean tension of registration between the reference (first) scan and any other scans is 2.3 mm. The first scan was selected as a fixed position, but global registration was applied for network adjustment. Assuming that all the scans were captured from the same position, no error propagation was expected from the network.

Data Analysis and Visualization

After removal of terrain and roof points we viewed the point clouds belonging to individual plants in 3D using FugroViewer[®] and in 2-dimensional height models which were obtained by rasterizing the maximum height in each cell in 10 cm intervals.

The 2D outlines were digitized in GIS with a margin of at least 10 cm, and further refined with the point cloud viewer to exclude all points belonging to the greenhouse structures and the plant container. Wherever the canopy started more than ~1 m above the soil level of the container, we also excluded the lower part of the trunk from subsequent analysis via a minimum height limit in order not to misinterpret canopy height changes. For each plant, this frame defined by the horizontal and the vertical limits of the canopy was kept constant for individual exports of the point clouds acquired during the scannings. No further editing of the raw data was undertaken. We used the procedure described by Puttonen et al. (2016) to subdivide individual point clouds into height percentiles that were exported to Excel. The height of each percentile during the first survey (6:30 p.m.) was used as baseline, and the changes of percentile heights with respect to this baseline were plotted during the course of the experiment. All changes in percentile height exceeding 0.5 cm were interpreted as movement. Percentiles are height thresholds that separate the point cloud horizontally into sets of equal counts of points amounting always to 10% of the total point count for each survey. Where the branches and/or leaves of a plant are displaced, this is followed by the height distribution of the point cloud and thus also the height percentiles. However, depending on the crown shape, the effect of canopy displacement on the changes in height percentiles may vary considerably. For trees with a uniform (near-cylindrical) crown shape extending from the ground right upwards, the interpretation of percentile height changes is straightforward. For trees with other shapes, the narrower parts such as the trunk or crown apex have the same number of points between percentiles, but spread across a wider height interval. Height changes in any part of the canopy affect all percentiles, and such narrow parts amplify these effects, creating artifacts of high displacement of these percentiles. To keep track of such effects, we indicated in gray the curves pertaining to the percentiles of the narrow tree parts. To render figures comparable we excluded data from percentiles pertaining to the most narrow parts of the tree, where the extent across was less than 50% of the maximum crown diameter. The movement for the percentiles that had more than 50% of the maximum crown diameter was also averaged to create a mean displacement value for each studied point in time. The extent and nature of canopy movement was visually interpreted from these graphs, checking time distances between local and general minimum and maximum points and the vertical extent of displacement, without further numeric data analysis, in accordance with Puttonen et al. (2016).

RESULTS

Ambient Conditions

Wind speed was measured to be 0 m/s during all measurements. Ambient light level inside the greenhouse was 63 mmol/s/m² at 6:30 p.m., 9.3 at 7:00 p.m., and dropped to 0.05 at 7:30 p.m., remaining at or below this value until 6:30 a.m., when 2.4 mmol/s/m² was measured, increasing to 33 mmol/s/m² by the last measurement at 7:00 a.m. Air temperature decreased gradually from 24°C at 6:30 a.m. to 20°C at midnight and further

to 17°C around 3:00 a.m., where it slowly increased to 18°C by 7:00 a.m. All air temperature sensors recorded the same temperature within $\pm 1^\circ\text{C}$. Relative humidity was around 65% for all sensors placed on tree trunks (and 55% for the sensor near the LIDAR), and increased continuously throughout the studied timeframe with small fluctuations of 1–2% that were closely similar for all trees, up to 95% at 7:00 a.m.

Observations of Tree Position Changes

All the plants measured showed movement of at least part of the crown that exceeded our noise threshold defined here as 0.5 cm. Most trees showed periodic movement. Either they revealed true sleep movement by returning to the early evening position by next morning, or they showed shorter term periodic movement, or both. Additionally, some trees showed unidirectional movement and did not return to their starting position within the measurement timeframe.

The most striking result obtained in this study was the large variation revealed in nocturnal tree movement patterns across the seed plants. *Betula utilis* (Figure S7) and *Olea europea* (Figure 4) displayed the least movement, where the mean of all crown height percentiles never exceeded the preset threshold of 0.5 cm for noise. The legume tree *Gleditsia triacanthos* (Figure 14) represented the opposite extreme by moving part of its crown up to 9 cm during the scannings. Practically all of the remaining trees showed oscillations in movement and shape of the crown of 0.5–1 cm with cycles that were shorter than the observed timeframe, typically a few hours.

Nerium oleander (Figure 1) revealed a pattern reminiscent of the overnight movement recorded in uncultivated *B. pendula* trees by Puttonen et al. (2016). Our scanning captured a slight downwards movement between 6:30 p.m. and 7:00 p.m. followed by a phase of upwards movement peaking at 9:30 p.m. Hereafter, the crown moved downwards with minor oscillation until 4:30 a.m. The remaining part of the morning the crown moved upward to resume the position recorded at the start of the experiment. Although only minor movements below 1 cm were captured during the scanning, the recorded patterns were very consistent across the crown height percentiles. It should be noted that the overall crown shape of this shrub was relatively even with no pronounced narrow parts.

Acer palmatum (Figure 2) showed a pronounced circadian cycle as well, but in the exact opposite direction: crown height percentiles were observed to move upwards reaching the highest point (+1.5 cm) at 4:30 a.m. Hereafter, the percentiles gradually moved downwards until next morning when the canopy had almost resumed the starting position.

The scanning of *Aesculus hippocastanea* (Figure 3) revealed a similar pattern of upward movement in the evening peaking at 04:00 a.m. and gradual return to the start position at 07:00 a.m. However, this overall pattern was interfered by rapid changes in height, which caused the mean of the crown height percentiles to shift more than 1 cm downwards between 04:00 a.m. and 05:00 a.m.

Although sleep movements were detected in the individual crown height percentiles of *O. europea* (Figure 4) the displacement of the mean of all percentiles never exceeded

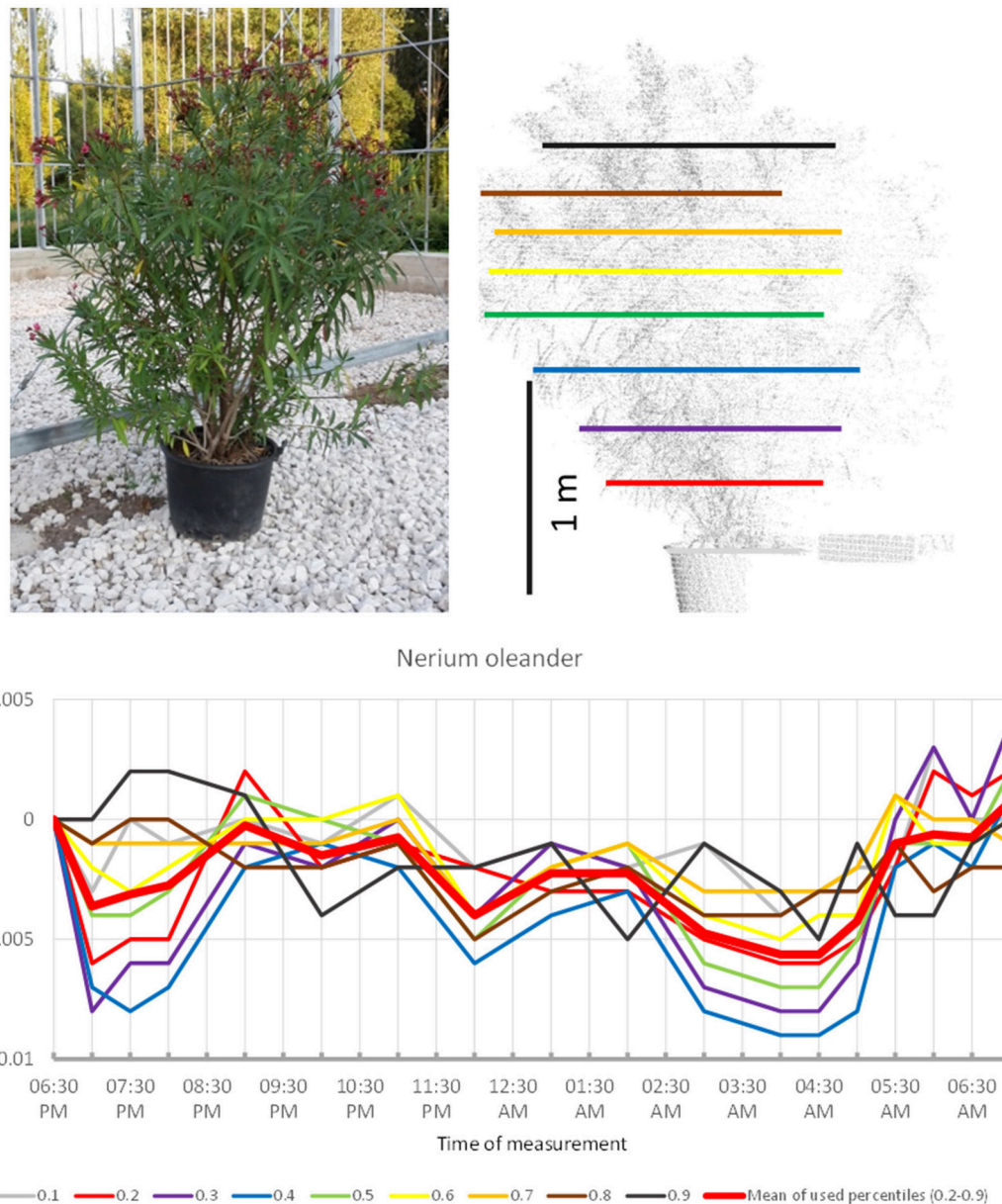


FIGURE 1 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Nerium oleander*. Percentiles not used for calculating mean movement are shown in gray. Note that movement pattern resembles typical sleep motion with some oscillation superimposed.

our preset detection threshold of 0.5 cm. The lower branches moved downwards before 07:00 p.m., upwards between 09:00 p.m. and 11:00 p.m., downwards between 01:00 a.m. and 05:30 a.m. and upwards again hereafter. The upper branches moved in quite the opposite pattern, reaching the highest point at 05:00 a.m., and resuming the starting position in the morning. The mean of the percentiles revealed an overall upward trend in the early evening, a more or less constant height during the night and returned to the start by morning; however, it should be noted that the amplitude was only 0.3 cm.

Viburnum pragnense (Figure 5) moved similarly but with an even longer upward phase and perhaps some short-term

periodicity superimposed. The lower branches moved downward whereas the upper branches seemed to move up almost continuously until around 4:30, when the movement was reversed and both the lower and the upper branches of the canopy returned close to their starting position. This results in the displacement amplitude of the mean of all used percentiles barely reaching 0.5 cm.

Photinia x fraseri (Figure 6) also showed movements similar to the upward sleep pattern detected in the species above: most crown height percentiles increased in height until 11:00 p.m. (compared to 10:00 p.m. recorded for most other plants), then slowly moved downwards until a minimum was reached between

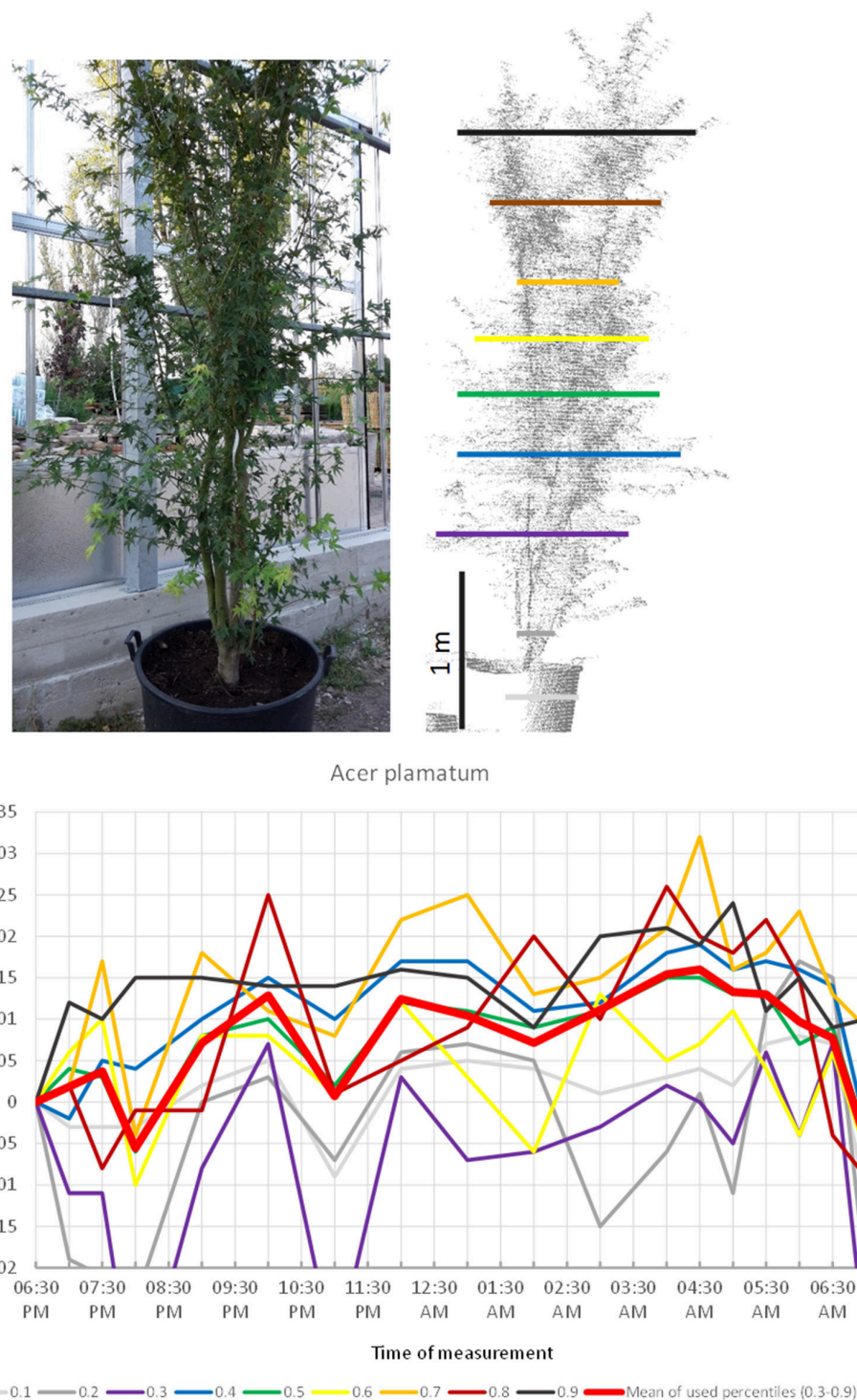


FIGURE 2 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *A. palmatum*. Percentiles not used for calculating mean movement are shown in gray. Circadian movement starts upwards and returns close to starting position.

4:00 and 5:00. However, prior to 8:00 p.m. and after 04:00 a.m., we recorded much more variation in point heights, which suggests that this shrub might have a shorter period of only about 8 h of sleep movement.

Quercus robur (**Figure 7**) seemed to carry out two cycles of movement. The first one started at 07:00 p.m., peaked at 11:00 p.m., and returned to the starting position around midnight. The second cycle started at midnight, peaked

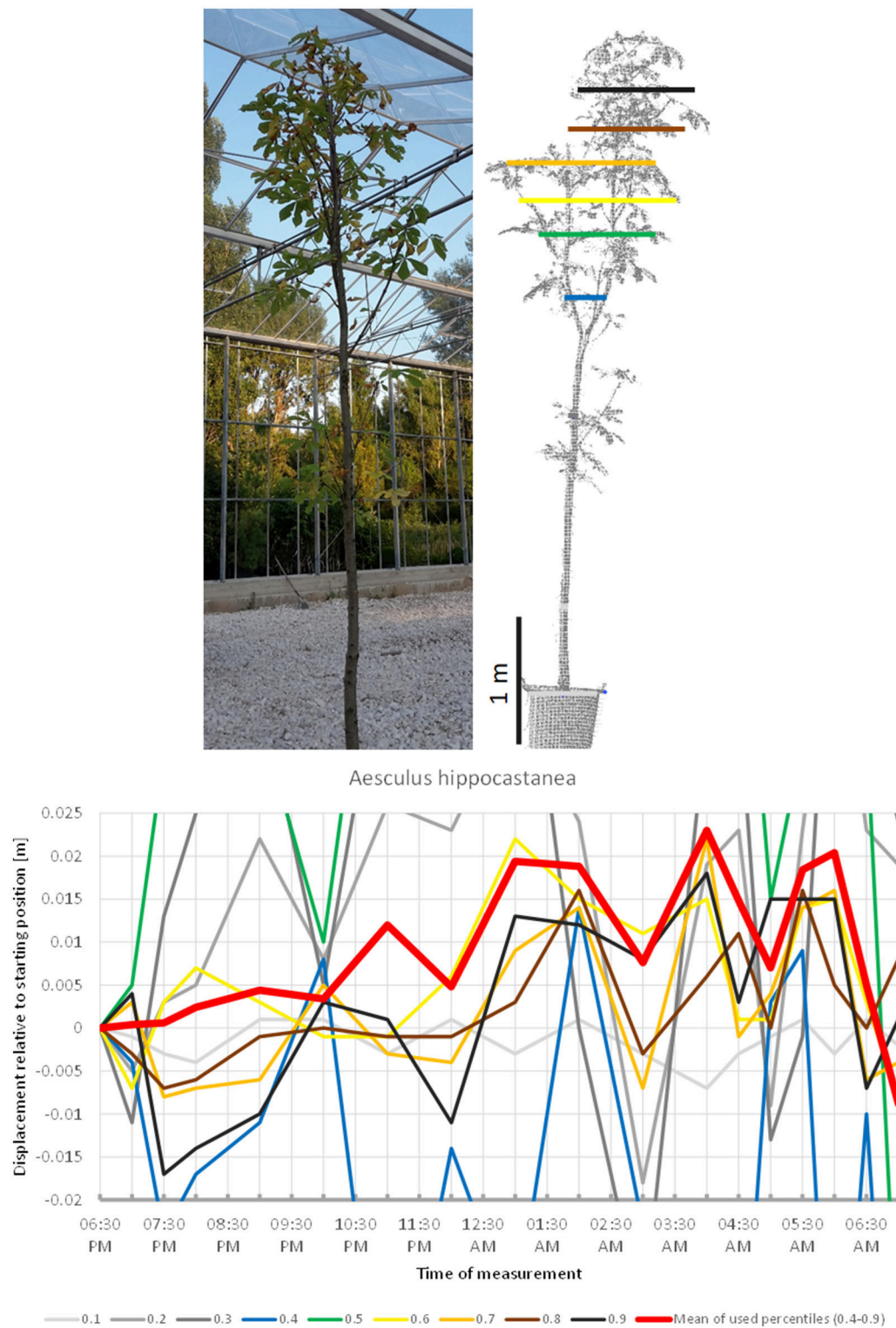
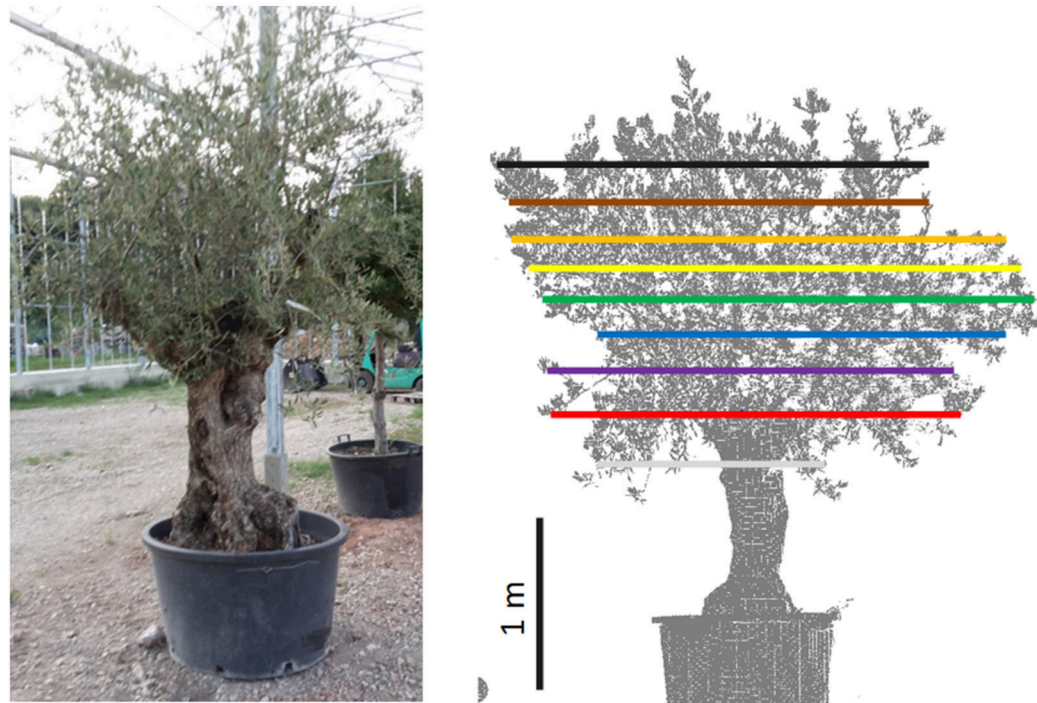


FIGURE 3 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Aesculus hippocastanea*. Percentiles not used for calculating mean movement are shown in gray. Circadian movement is upwards, with intensive oscillation superimposed.



Olea europaea

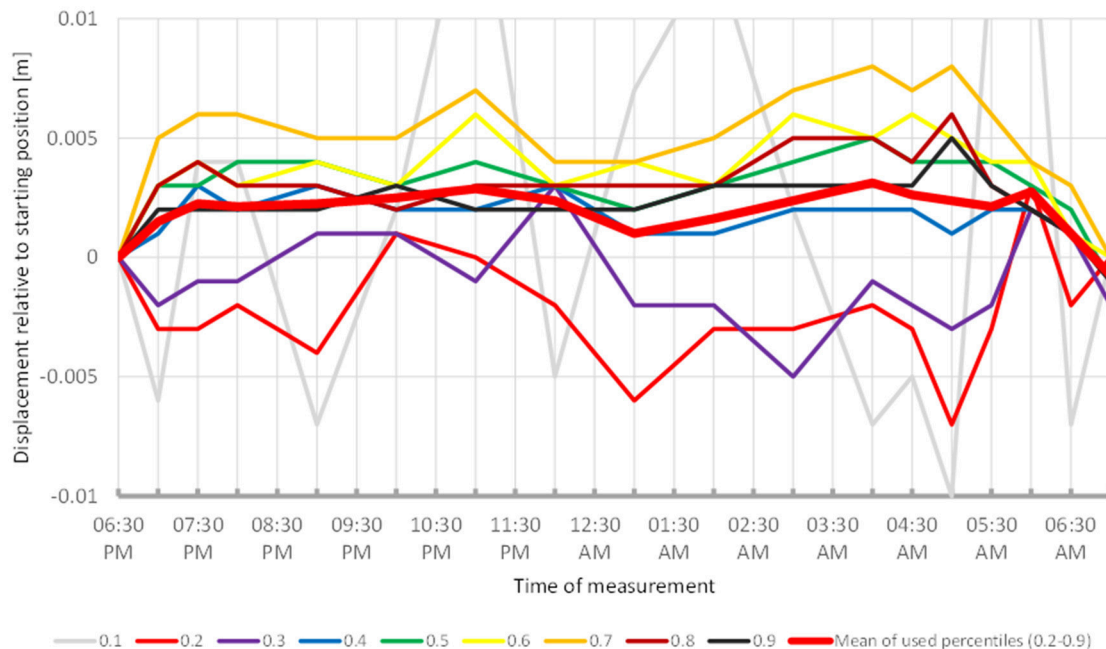


FIGURE 4 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Olea europaea*. Percentiles not used for calculating mean movement are shown in gray. Slight circadian movement starts upwards and returns to starting position, oscillation of mean is not prominent.

at 5:00 a.m. and resumed the starting position around 7:00 a.m.

Magnolia grandiflora (Figure 8) apparently did not follow a diurnal cycle but moved in cycles of ~ 3 h, peaking at 8:00 p.m., midnight, and 4:00 a.m. respectively. This pattern

was quite distinct from the other species included in the experiment. The fact that the start and end positions of the branches were similar, and that the movement of the individual height percentiles was more or less synchronized strongly suggests that this recorded pattern was not just an

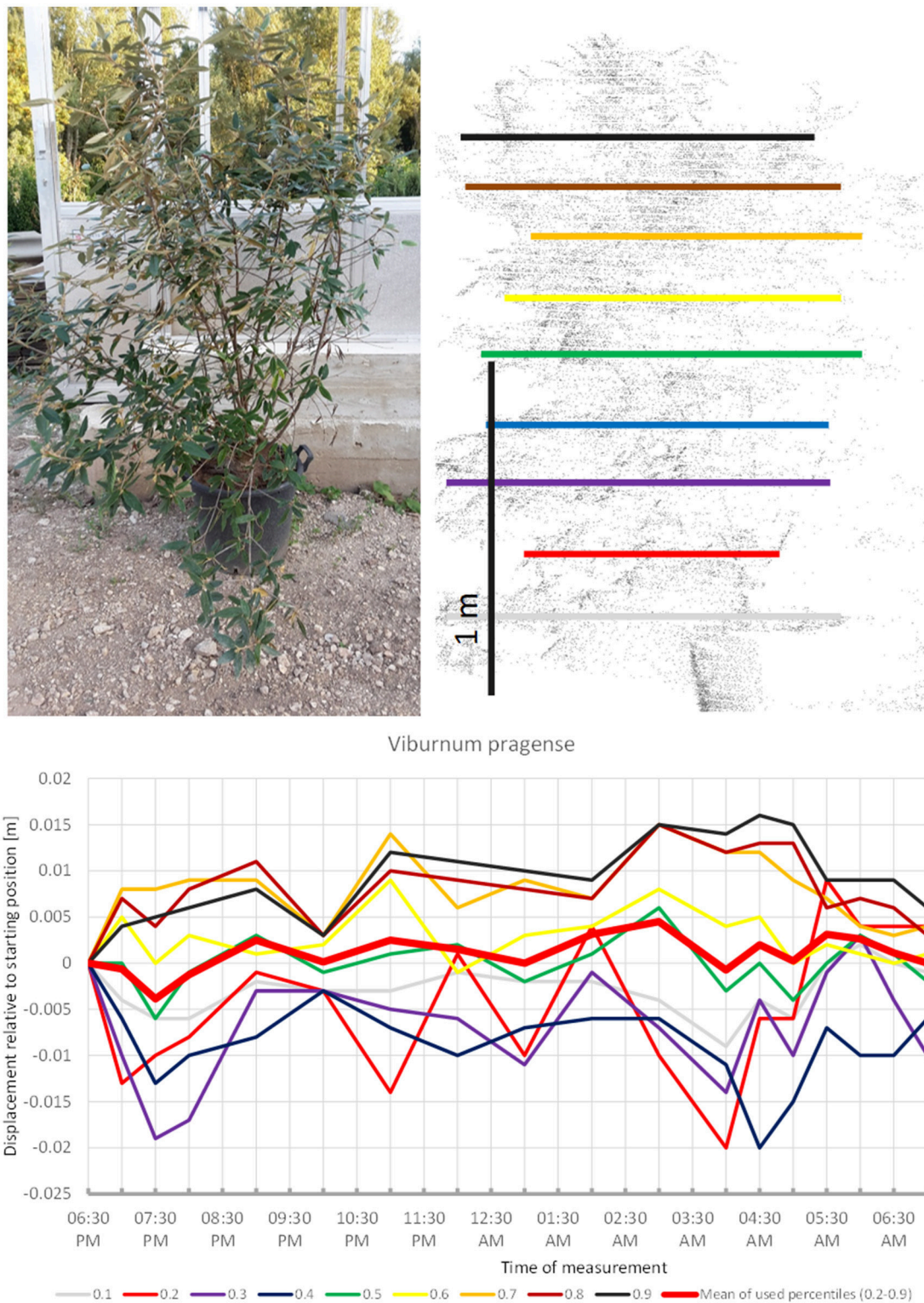
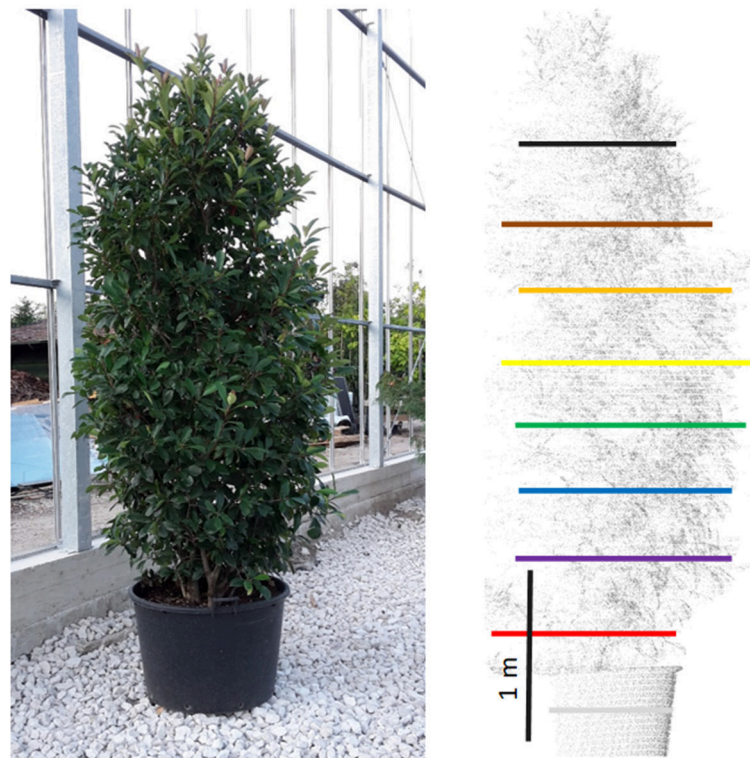


FIGURE 5 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Viburnum pragense*. Percentiles not used for calculating mean movement are shown in gray. Note that lower percentiles move downwards while higher ones move up.



Photinia x fraseri

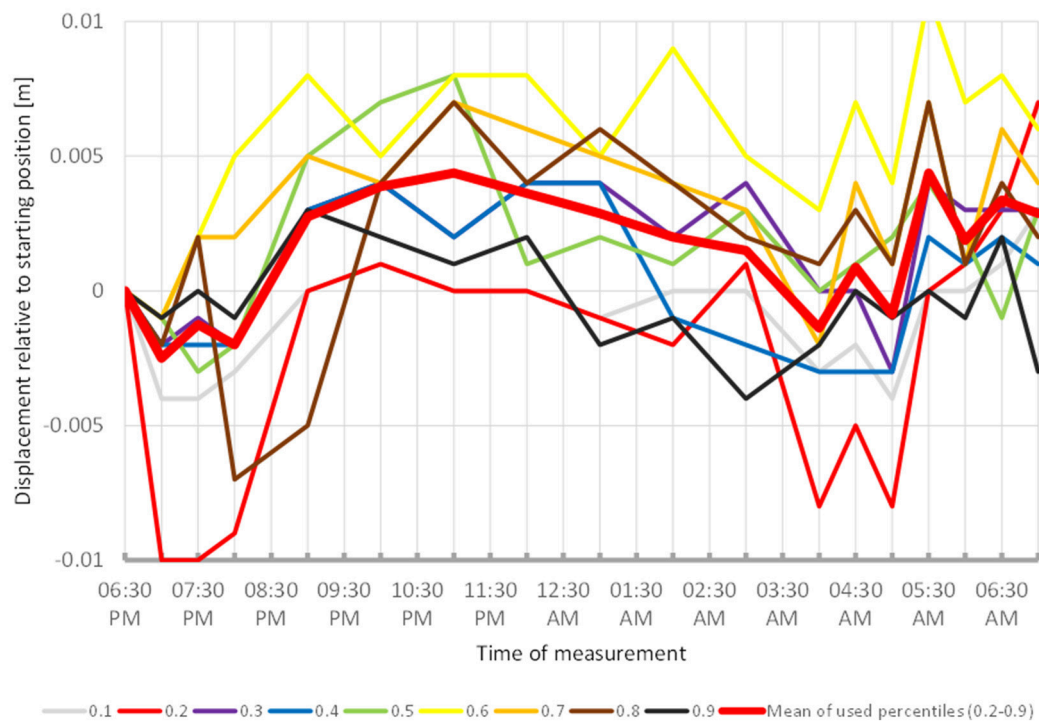


FIGURE 6 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Photinia x fraseri*. Percentiles not used for calculating mean movement are shown in gray. This plant seems to have an overnight phase shorter than 12 h.

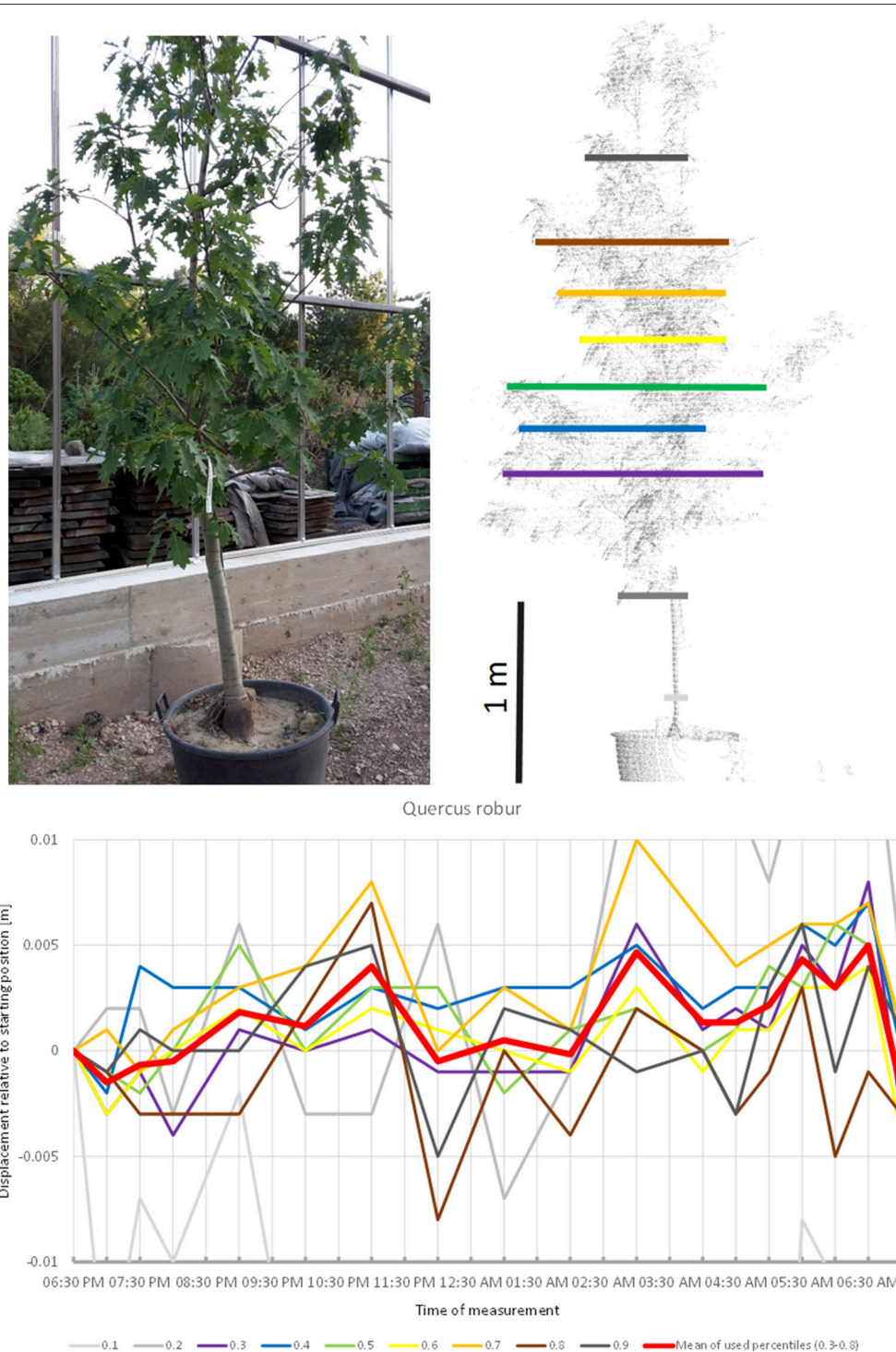


FIGURE 7 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Quercus robur*. Percentiles not used for calculating mean movement are shown in gray. This movement pattern can be interpreted as two cycles, always returning to the mean position.

artifact of noise. The maximum displacement of the mean of all percentiles is exactly -0.5 cm, which equals to our noise threshold. However, this is partly an effect of the lowest percentile, which moved in the opposite direction relative

to the rest of the tree, thus resulting in lower average movements.

Among the Gymnosperms, *Larix kaempferi* (Figure 9) showed upward movement during the first hours of the

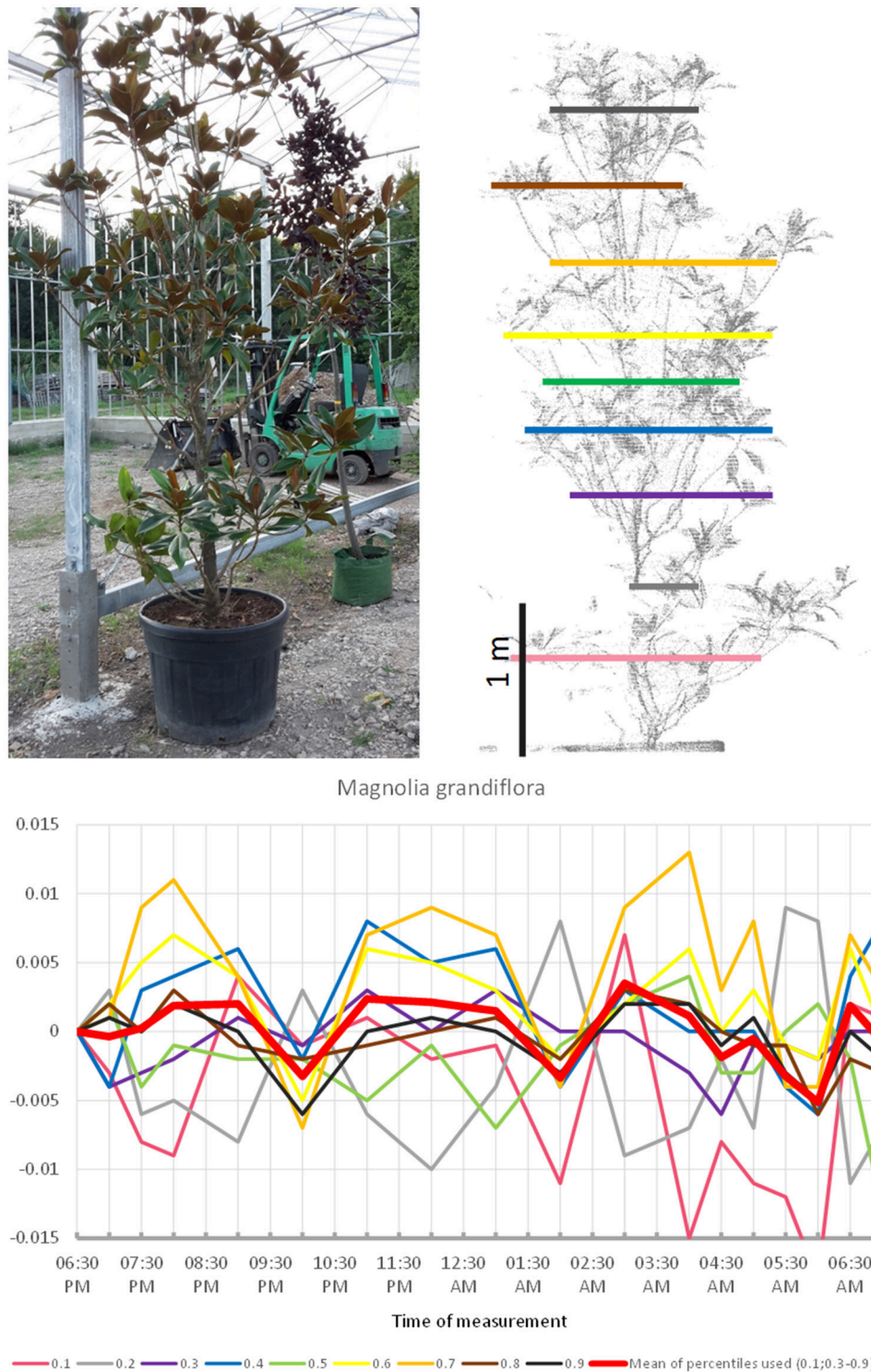


FIGURE 8 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Magnolia grandiflora*. Percentiles not used for calculating mean movement are shown in gray. Movement pattern implies domination of short-term oscillation, with three cycles during the measurement period.

survey and then continued downward for the rest of the night, with two short peaks at 01:00 a.m. and 06:00 a.m. It returned relatively close to its start position by the end of

the experiment in the morning. Nevertheless, the recorded movement pattern seemed to deviate from the sleep patterns of the trees above.

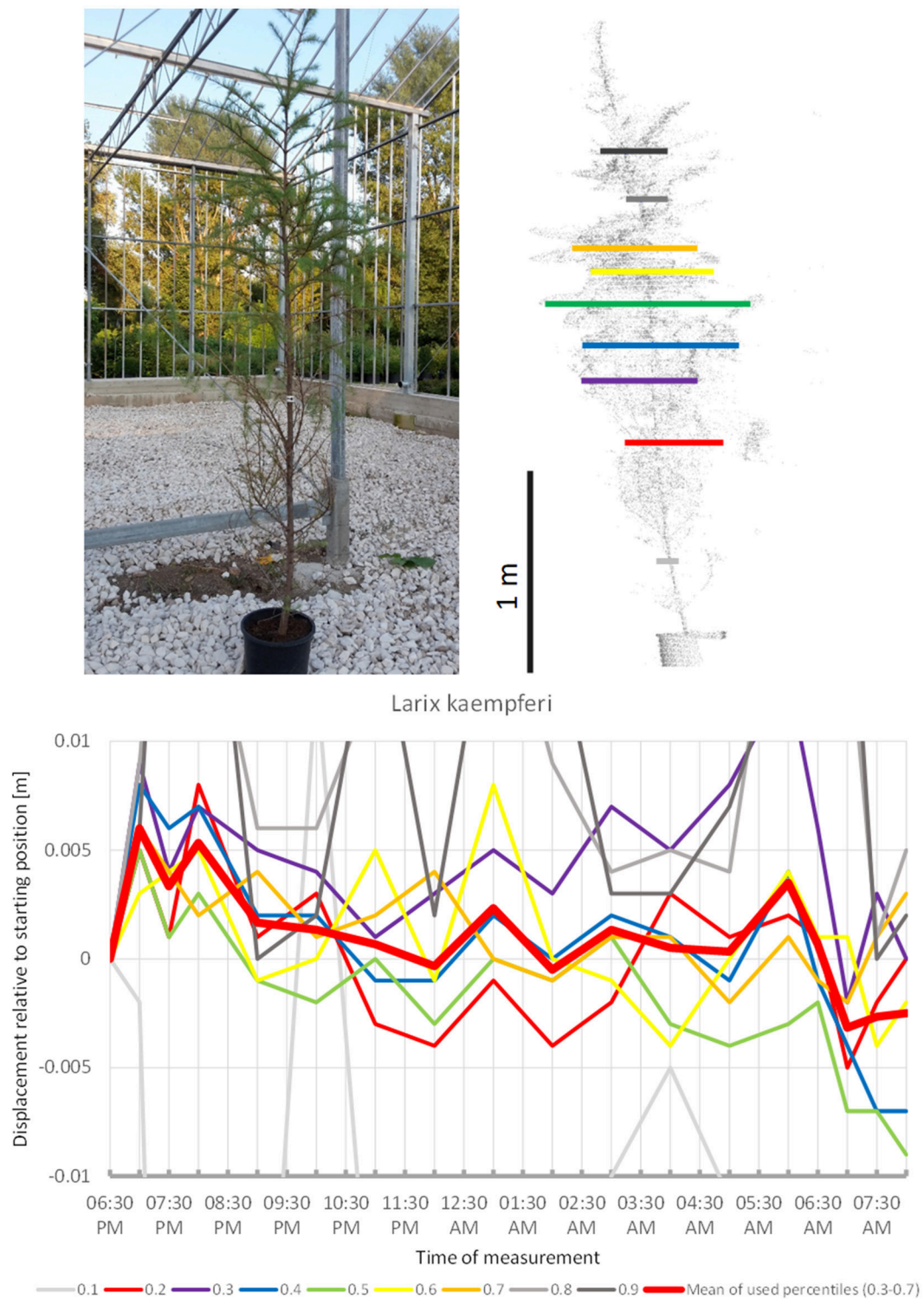


FIGURE 9 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Larix kaempferi*. Percentiles not used for calculating mean movement are shown in gray.

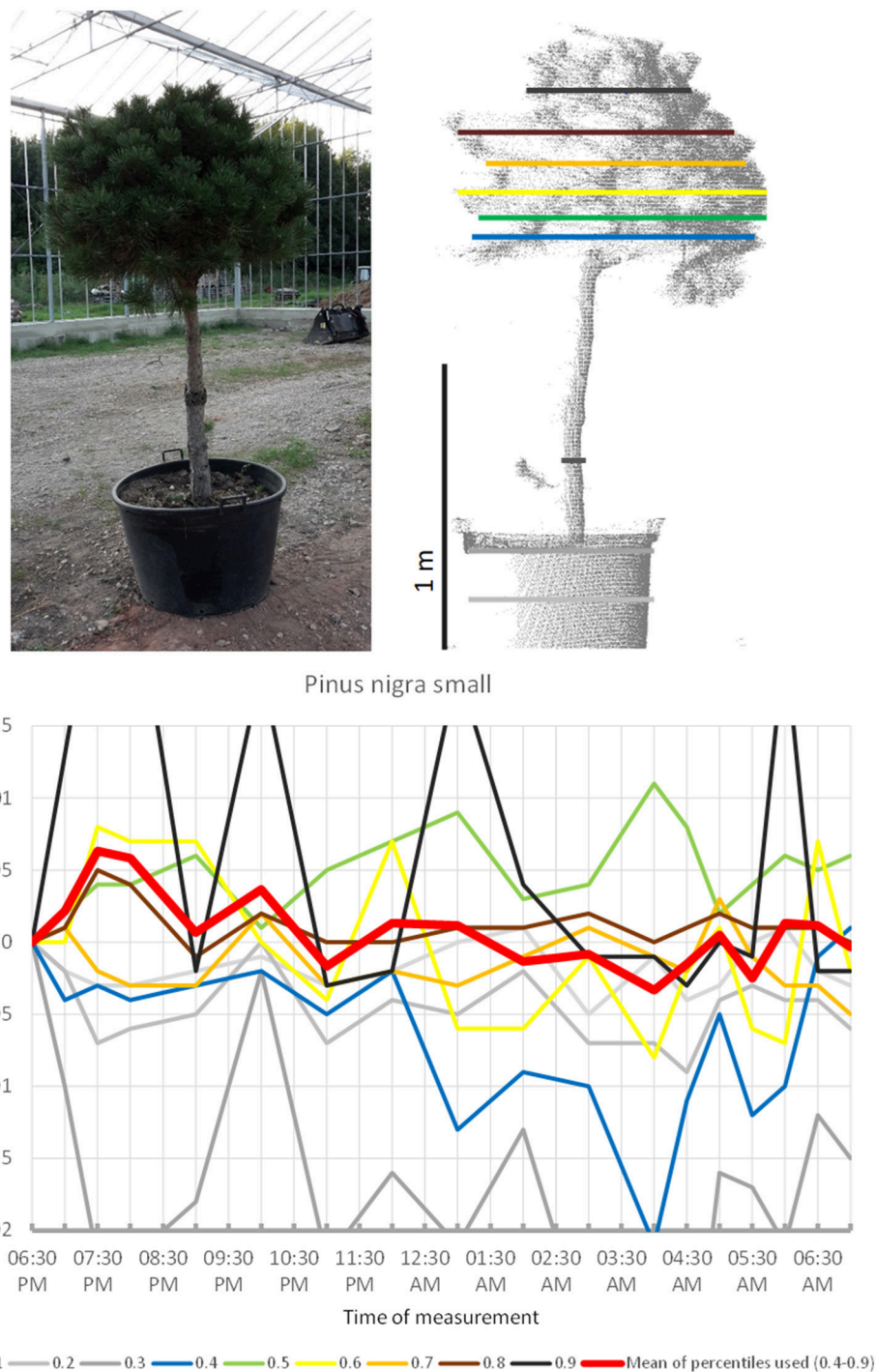


FIGURE 10 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for the smaller *Pinus nigra* tree. Percentiles not used for calculating mean movement are shown in gray. Fluctuation around a mean returns to the original position by the end of the survey in the morning.

The smaller of the two *Pinus nigra* individuals included in the experiment, *P. nigra* 2 (**Figure 10**), moved upwards until 7:30 p.m., downward until 04:00 a.m., and then returned back to the original position, with some oscillation superimposed. More scannings are needed to clarify whether this pattern form part of a

sleep movement or rather represents random oscillations around a mean.

The larger *P. nigra* individual, *P. nigra* 1 (**Figure S1**), showed extreme fluctuations in the beginning of the scanning. We suspect this to be an artifact of mechanical disturbance at the

beginning of the experiment and therefore not go into further detail. Compared to the other species included in the experiment the percentiles are highly variable in between scanings. This may be explained by the large size and long branches of the selected individual. The mean of the percentiles revealed that the tree kept a relatively constant height during most of the experiment, with two peaks at 8:00 p.m. and 05:00 a.m. respectively. It should be noted that this particular tree was almost twice as tall as the other trees included in the study and therefore had to be placed outside the greenhouse. For this reason great caution should be taken when making comparisons with the other species.

Cedrus libanica (Figure S2) fluctuated around a mean as well, with peaks around 07:00 p.m. and 3:00 a.m. It stood out, however, by not returning to its original position at the end of the measurement; in fact, the largest displacement from the starting position was observed at the last scan in the morning.

B. pendula (Figure 11) showed some abrupt downward movements compared to a seemingly constant baseline. Low points are occurred at 8:00 p.m., 04:00 a.m., and 07:00 a.m. This did not resemble classical downward sleep movement. Movement of most crown height percentiles was congruent. The tree did not return to its original position by 07:00 a.m.

Many plants belonging to the family Fabaceae with pinnately or double pinnately compound leaves display striking sleep movement of their leaves (nyctinasty). Among the trees included in our experiment *Albizia julibrissin* (Figure 12) showed the most conspicuous movement pattern, even though the leaves entered the sleep position before the scanning was initiated. A substantial part of this movement may be explained by the concerted action of the pulvini of consecutive orders. The scanning revealed upward movement of the crown until ca. 9:00 p.m., followed by downward movement until it reached the lowest point at 03:00 a.m. Hereafter, the crown gradually ascended until the morning. The leaves were observed to have reached their daylight configuration around 6:30 a.m., however, the tree reached its 06:30 p.m. start position already 1 h earlier (5:30 a.m.), and continued to move upward.

Another fabaceous tree included in the study, *Styphnolobium japonicum*, (Figure 13) stood out by having a pendulous crown, with an overall more or less cylindrical shape. The circadian movement of the leaves contributed only moderately to the overall crown movement in this species. The lower percentiles of the tree revealed most movement, but the broader lower part of the canopy top also moved more than 2 cm down. The crown movement had two low points, one at midnight and one between 4:00 and 5:00. It was not clear whether the night phase of the circadian cycle of this tree was longer than the duration of our experiment and the crown would return to its start position later, or whether the movement was not circadian.

Prunus virginica (Figure S3) showed a similar pattern, but the crown moved in the opposite, upward direction apparently right from the beginning of the experiment, with the highest point reached around 4:00 by most of the crown height percentiles. A return to the starting position was, however, not observed for the

mean and some percentiles and it remains uncertain whether the branches would have resumed the starting position later during the day or whether the movement was induced by other factors than circadian rhythm.

Trachycarpus fortunei (Figure S4) stands out in its growth form by having an unbranched stem. Most crown movement must therefore be attributed to the leaves. This species was characterized by a short downward movement, followed by upward movement of most crown height percentiles until 01:00 a.m., then downward movement until 4:30 and finally upward movement again. It is not clear whether this pattern reflects a short circadian cycle similar to *P. x fraseri*, or whether it is the result of a longer cycle with a return to the starting position later during the day.

Some trees showed more or less unidirectional movement through the duration of the experiment. In such cases the observed patterns were not sufficiently clear to conclude whether the movements represented sleep or not. The leaves of the third fabaceous tree included in the experiment, *G. triacanthos* (Figure 14) were observed to have entered sleep configuration before the beginning of the experiment. The crown height percentiles of the canopy moved slightly downwards between 06:30 p.m. and 07:00 p.m., then moved upwards by 9:00 p.m. until 01:00 a.m., when a slight downward movement was detected, after which the percentiles showed a clear increase in height again. The last position scanned was situated 5 cm away from the starting position and we did not observe any signs that this would be resumed later.

The crown of *Morus alba* (Figure S5) moved constantly upward during the experiment, with only a slight downward movement around midnight. The duration of the experiment was not sufficiently long to determine whether this pattern was due to sleep movement initiated before the scanning started at 06:30 p.m. or some other processes. However, since this tree appeared to be generally unhealthy, the observed movements might have been determined by other factors than sleep movement.

The bamboo, *Fargesia murielae* is characterized by having multiple stems, each with a uniform modular structure and slender first order branches carrying the leaves (Figure S6). For the lower two crown height percentiles (0.5 and 0.6) we recorded a pattern similar to other plants with sleep motion: a short downward movement between 06:30 a.m. p.m. and 07:00 a.m. p.m., followed by upward movement until 9:00 p.m., down until 04:30 a.m. and back up, but with the original position reached at 5:30 and downwards movement afterwards. This species was excluded from further analysis since the uppermost part of the crown was partly in contact with and supported by the greenhouse frame.

In some of the tress studied, the movement observed can be interpreted as non-cyclic movement superimposed on an overall circadian cycle. This particularly applies to *Chamaecyparis nootkatensis* (Figure 15), which moved upwards during most of the night. At 10:00 p.m. the apical part of the canopy showed an abrupt downward shift (the crown height percentiles 0.6, 0.7, 0.8, and 0.9). An hour later at 11:00 p.m., it had returned to the pre 10:00 p.m. position. The tree reached the highest position at 02:00 a.m. and

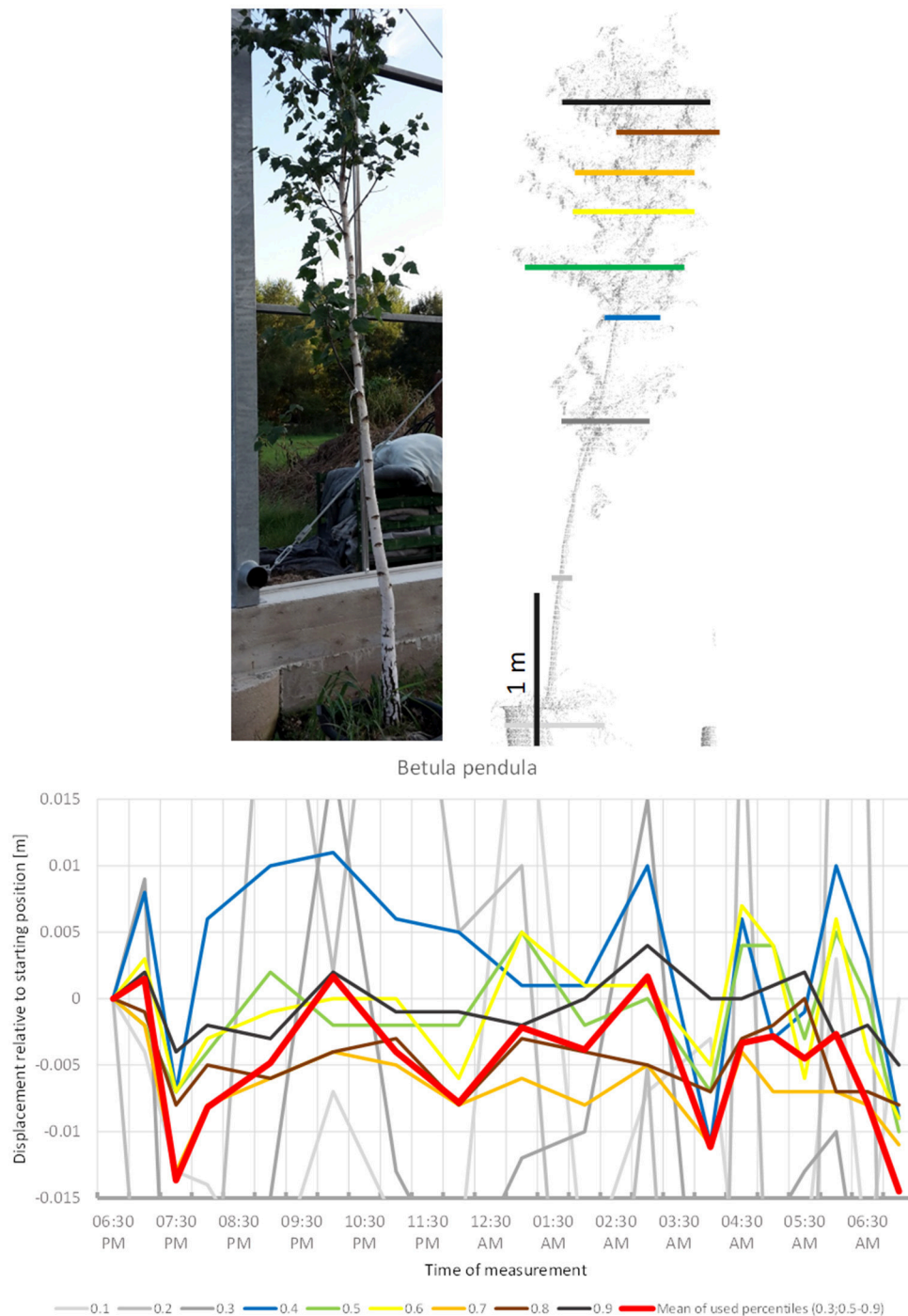


FIGURE 11 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Betula pendula*. Percentiles not used for calculating mean movement are shown in gray. Movement of the tree seems to be dominated by short-term periodicity and does not return to the original position.

moved downwards until 5:00 a.m. Hereafter, it moved upwards again but did not resume the starting position. It should be emphasized that this particular tree has small scaly leaves and that the patterns observed were solely a result of branch movements.

Similarly, the crown of *B. utilis* (Figure S7) moved upward as expected until 8:00 p.m. The lowest point was reached at 02:00 a.m., and the highest point at 06:00 a.m. with slight downward movement afterwards. The crown did not return to its original position.

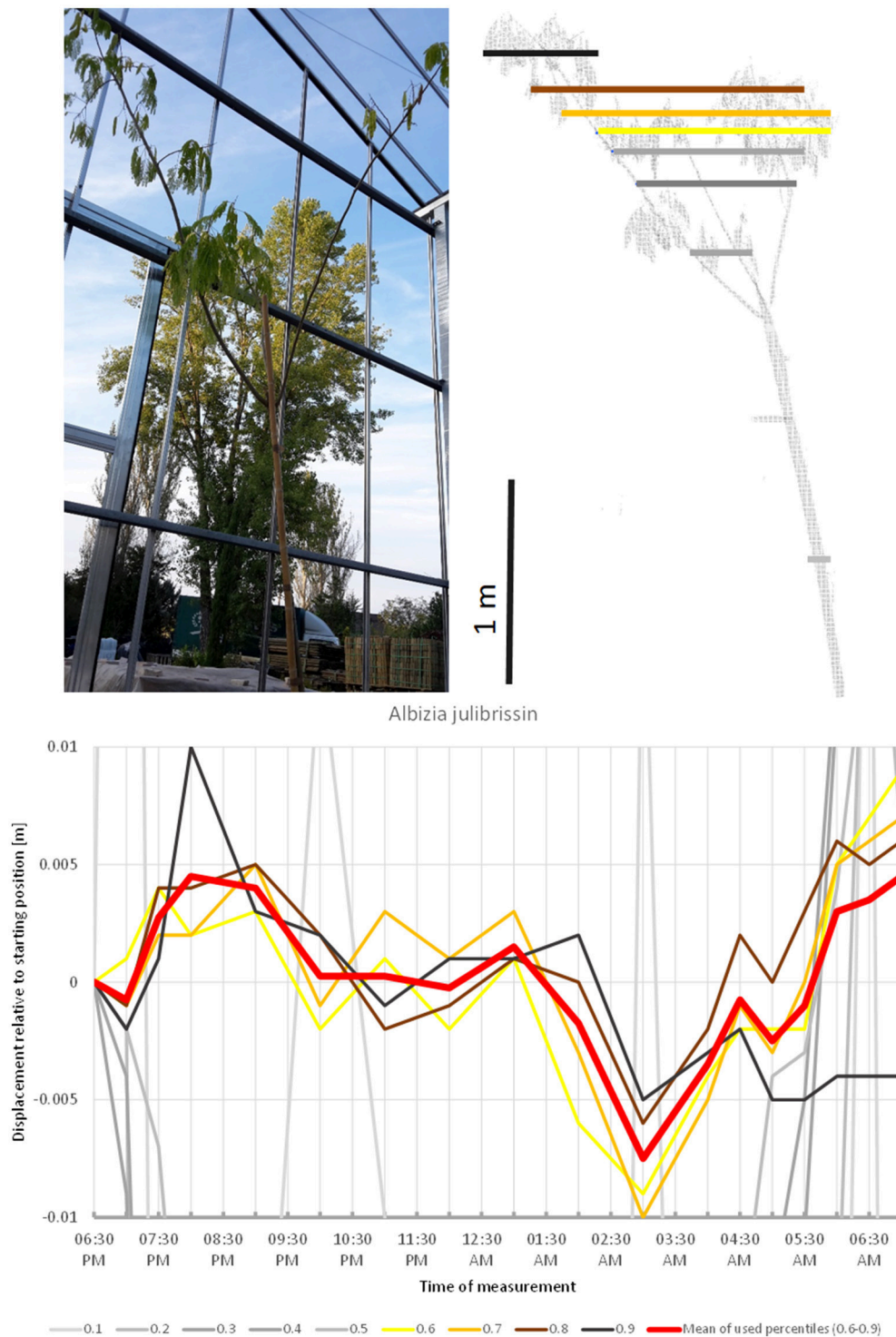
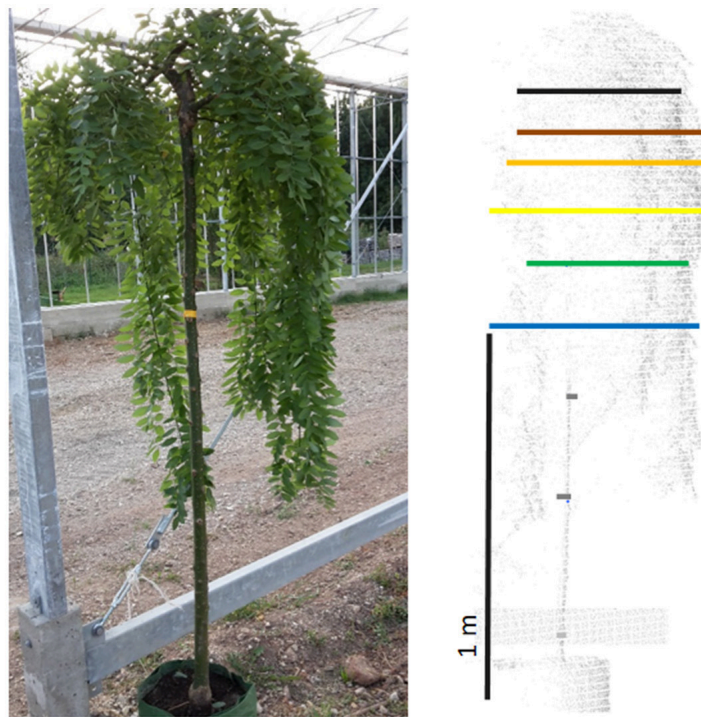


FIGURE 12 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Albizia julibrissin*. Percentiles not used for calculating mean movement are shown in gray. Movement of this tree might be interpreted as a circadian cycle returning to the original position before the end of the measurement.



Styphnolobium japonicum

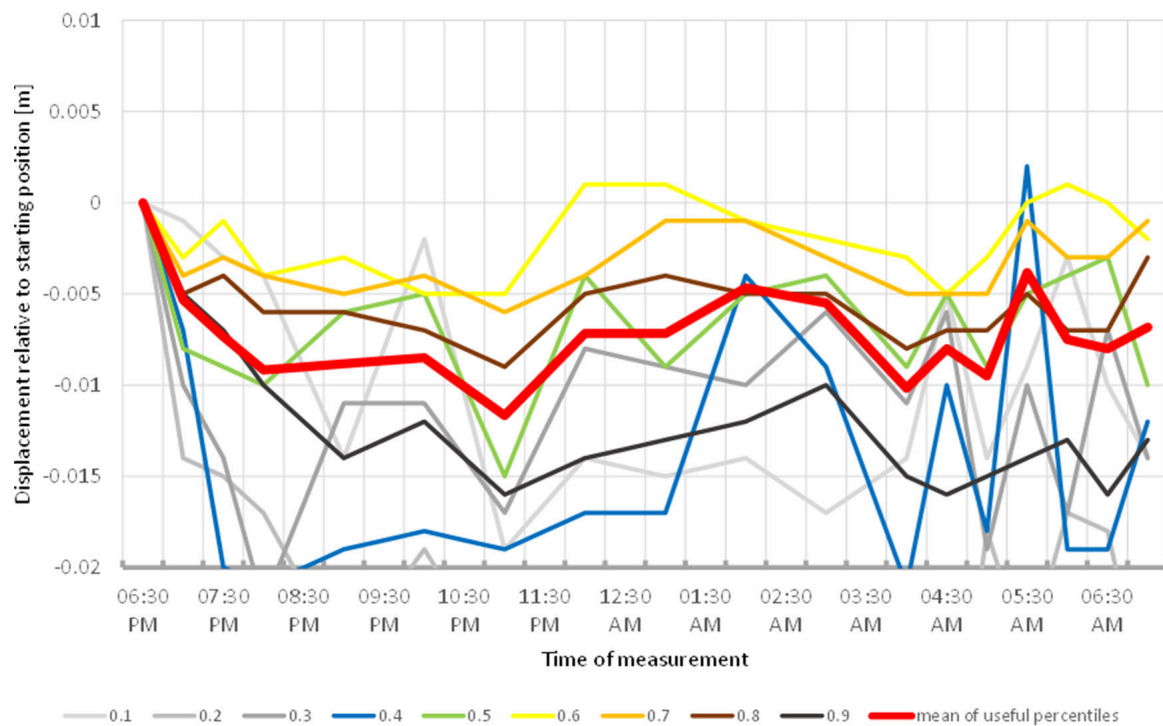
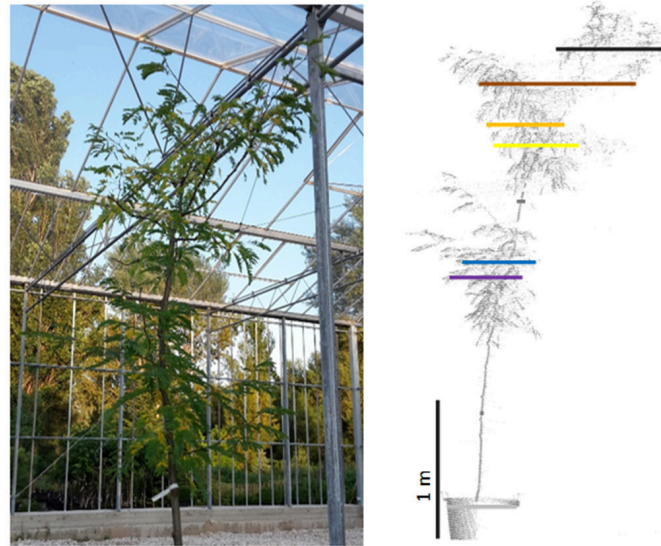


FIGURE 13 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Styphnolobium japonicum*. Percentiles not used for calculating mean movement are shown in gray. It is not clear whether the branch positions would have returned to the start later during the day.



Gleditsia triacanthos

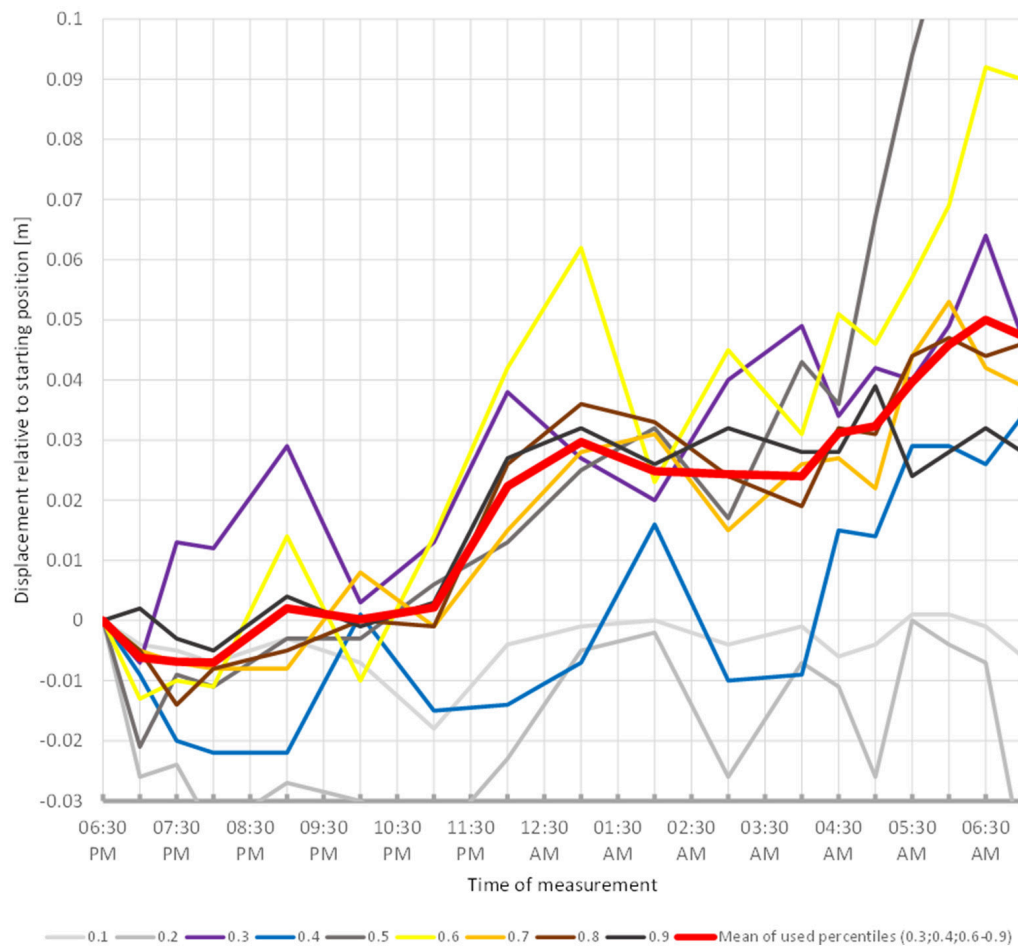


FIGURE 14 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Gleditsia triacanthos*. Percentiles not used for calculating mean movement are shown in gray. Movement of this tree seems to be dominated by aperiodic displacement.

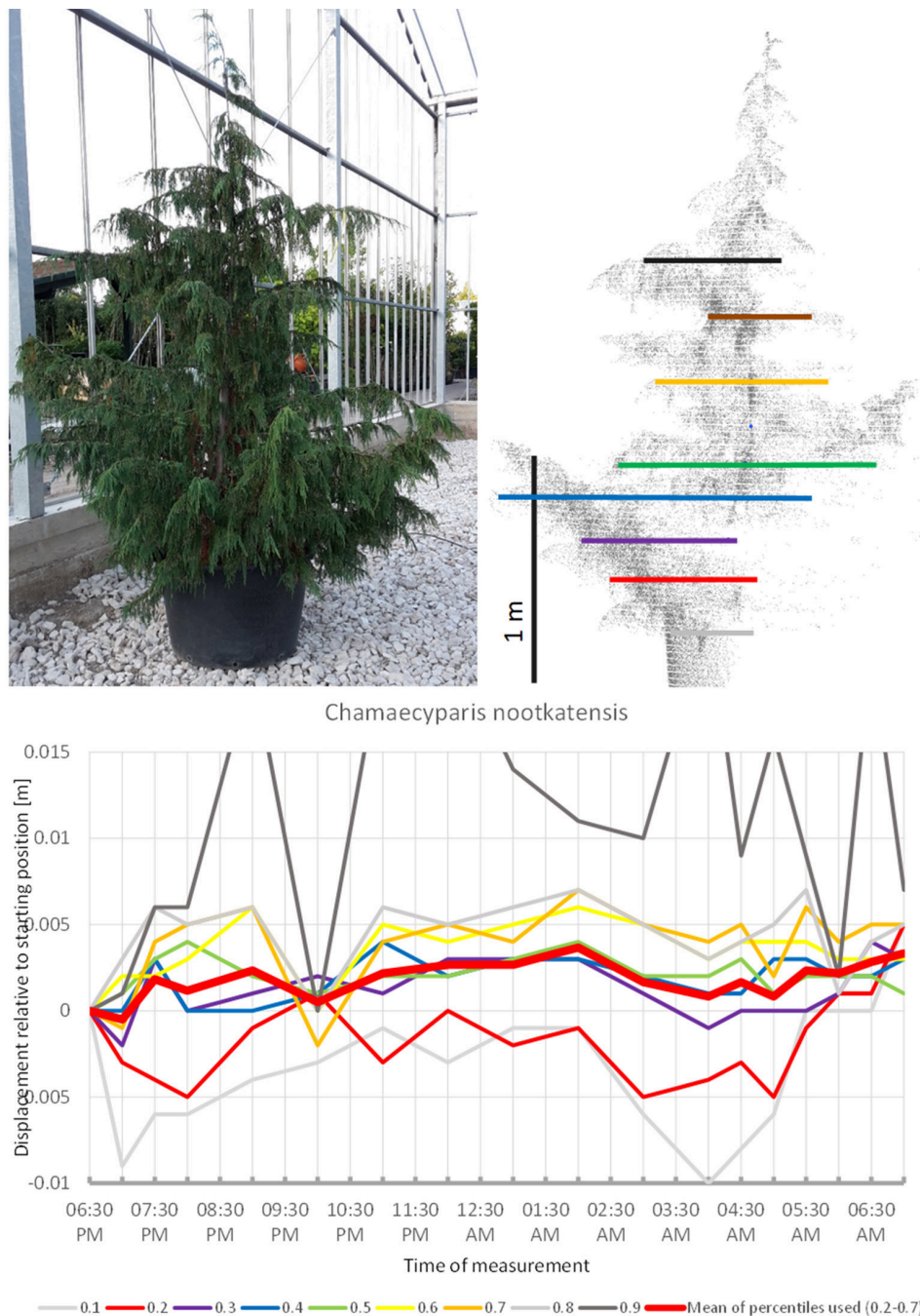


FIGURE 15 | Point cloud height percentiles and their displacement (m) in time relative to the starting position, for *Chamaecyparis nootkatensis*. Percentiles not used for calculating mean movement are shown in gray.

Summary of Results

Changes in crown shape were observed for all studied trees, but clear and well-measurable sleep movements are apparently rare. Ambient light conditions were continuously dark between 7:30 p.m. and 6:00 a.m., but most changes in movement direction occurred during this period, suggesting that they were not controlled by ambient light. *Aesculus* and *Acer* had the

highest periodic circadian movement amplitudes of the crown amounting to 2 cm, whereas the highest aperiodic change in crown shape was observed for *Gleditschia* (5 cm) and *Fargesia* (2 cm). We detected some shorter-period oscillation with cycles of 2–3 h and amplitudes between 1 and 0.5 cm, close to the margin of our measurement accuracy. Movement patterns recorded can be attributed to in three major types: circadian

movement in an 8–12 h cycle (*Nerium*, *Acer*, *Aesculus*, *Viburnum*, *Photinia*), periodic movement between 6 and 2 h (*Quercus*, *Magnolia*, *Pinus*, *Cedrus*), and non-periodic, unidirectional movements (*Styphnolobium*, *Prunus*, *Gleditschia*, *Morus*). The remaining trees studied (*Olea*, *Betula*, *Larix*, *Trachycarpus*, *Fargesia*, *Chamaecyparis*) either show a combination of these basic movement forms or low-amplitude movements close to the margin of our accuracy or an apparently random pattern.

DISCUSSION

Measurement Accuracies and their Effects on the Results

The scanning method used here has a nominal ranging accuracy of 0.003 m on solid, fixed targets. This signifies that theoretically the results presented here can be influenced by errors in position measurement. However, since these effects are probably random, their influence can be reduced by ensuring a large number of points in each height percentile. The differences in the number of points acquired for each tree during the individual scanning sessions fluctuate slightly ($\pm 1\%$) mainly due to ghost points. This may constitute an additional source of error. Finally, the co-registration accuracy of the individual scans falls within 2.3 mm, including both horizontal and vertical errors. Horizontal co-registration errors have no effect on crown height percentile calculations since the point clouds representing the individual trees were cut out with the same 2D and height boundaries each time, allowing for at least 10 cm distance around the points observed to belong to the tree. Vertical co-registration errors might have an effect, but that is well characterized by the error observed at the co-registration targets since the reference spheres and the greenhouse structure were further from the sensor than the trees were. Co-registration errors would reveal themselves by spatial autocorrelation, which we did not detect in our data. In this study we followed the methodological approach used by Puttonen et al. (2016) as close as possible to make the results comparable. The relatively low number of measurement points on the trees limited more sophisticated analysis of position change, e.g., by segmentation or the iterative closest point algorithm (Glira et al., 2015).

The phase-based TLS successfully recovered sleep movement in some of the trees species studied, whereas no movement was detected in other of the trees. The small size of the trees included in the experiment probably constrained the magnitude of the movement. The scanings were not duplicated for practical reasons (for most species, the gardening shop had only one individual in stock) except for the two individuals of *P. nigra*, thus we cannot exclude that some of the canopy movements were undetected (false negatives). Also due to the small size of the trees it was decided not to affix standard reflectors to the branches that would have allowed for a more accurate tracking of movement. The draw-back of using such a method would have been that the weight of these reflectors could have influenced the movement patterns. Although the half-hour intervals between some scans represent a step forward compared to Puttonen et al. (2016), the timing of the survey was still not optimal. The duration of the

experiment should ideally be 24 h, but for the interpretability of the results the scanings should at least have been initiated a couple of hours before and run for a few hours longer. Despite the exploratory nature of this study we have nevertheless revealed substantial variation of nocturnal plant movement across seed plants, which opens up a number of new research questions specifically addressing the dynamics of sleep movements in trees.

A Conceptual Framework for Interpretation of Plant Movement

Our results clearly show that small trees move their canopy during the night in complex ways and that it is difficult to identify general patterns. Some of the movements detected were close to the preset noise threshold and although we find it unlikely, we cannot exclude effects caused by random noise. Based on the results of the present exploratory study we propose a framework that can be used for future reference of research focusing on the morphological and physiological implications of canopy movements. We propose that all movements are broken down in four basic and more or less interacting patterns, (1) sleep (circadian, light or internal clock controlled), (2) drift (non-circadian, a trend within the window of observation), (3) oscillation (short term fluctuation, typically superimposed on other types of movements) and, (4) noise (random fluctuation around a mean of zero within the expected noise limits).

As an initial rule of thumb we differentiated between canopies with and without circadian movement (“sleep”) respectively by comparing the starting and end positions of these. Traditionally sleep movement has been defined as vertical, cyclic with approximately 12 h as period length (Darwin and Darwin, 1880; Puttonen et al., 2016). We classified movement as sleep if the tree returned to its starting position 8–12 h after the start of the experiment, and if one or two clear minimum/maximum height positions could be identified. Most of the trees included in this study returned to their evening positions the morning after thereby moving full circle, whereas others showed unidirectional movement and never resumed their position at the beginning of the experiment. Sleep movement governed by an internal circadian clock has already been described e.g., in *Arabidopsis thaliana* (Barak et al., 2000; McClung, 2013). Canopy movement may be an adaptation to light exposure, and may occur in both upward or downward directions during the night. The movement pattern during the day is expected to be the opposite (Mullen et al., 2006). The movement itself is linked to turgor pressure in cells situated in places of the plant body where their effect on organ motility is optimal. This applies e.g., to the bulliform cells in grass leaves, the lodicules in grass flowers and to cells in the pulvinus found in plant families such as Fabaceae and Maranthaceae that controls movement of the leaf lamina. When it comes to movement of branches, much less is known but this is most likely a consequence of the hydraulic dynamics of the conducting tissues. The simplified view that phloem and xylem constitute separate transport systems with exclusive functions has recently been abandoned and today, we have a more refined understanding of the complex interaction between tissues of the inner bark region and the xylem, facilitated by radially aligned

rays (Pfautsch et al., 2015). The palm included in this study is of course excepted since it is a monocot without rays. It has been shown that the diurnal pulse of stem expansion and contraction is driven by the daily cycle of transpiration and to a lesser degree by photosynthesis (Zweifel et al., 2001). The mechanisms behind the minute nocturnal branch movements revealed in this study are probably more complex and linked to physiological dynamics.

The second type of movement we propose here is “drift,” which implies unidirectional movement. Observed movement was classified as drift if no periodic changes in movement direction occurred and the position farthest from the starting point was measured at one of the last scans in the morning. Since we only ran our experiment for 12 h we do not know whether these movements form part of a long-term change in the tree shape. Such long-term changes may be linked to stress caused by disease, herbivory, or lack of water and will eventually lead to wilting. It may also be a result of the plants preparations for hibernation. The study was conducted near autumnal equinox to assure an almost even length of night and day and we cannot exclude that some of the trees were already preparing for hibernation.

The third type of movement, which we refer to as “oscillation” includes short-term movements in the canopy that occur in cycles of a couple of hours or shorter. Movement was interpreted as oscillation if at least three local maxima or minima could be observed, with comparable amplitudes. Although the biomechanics and ecophysiological basis of these are unknown, movements are most likely linked to tissue specific changes in turgor pressure caused by action of membrane borne ion pumps. We can only speculate of the molecular mechanism leading to these rather fast changes in turgor pressure, yet one possibility could be pulses in the way osmotically active metabolites are stored and released. In that case the movements could be regarded as *consequence* of physiologically processes that are vital to the plants and the movement *per se* can therefore be regarded as having no adaptational implications.

Lastly random noise should also be considered as a driver of movement recorded. Here, we used a preset threshold of 5 mm to filter out the noise, well knowing that we could lose valuable information in this way. For convenience and lack of knowledge we used the same filter across all the plants studied. Future studies of plant movement should consider other ways of reducing noise especially in those species that show moderate changes in the canopy.

Observed movements are a result of a combination of these four types, therefore their interpretation might still be ambiguous. Combinations of two or three main movement types may be superimposed and can in such cases only be interpreted through the main trend and local deviations from this trend with shorter or longer periodicity.

Interpretation of our Observations Based on the Conceptual Framework

Among the species included in this studies *A. palmatum* and *A. hippocastanea* most clearly represent the sleep type of

movement. In both species drift is negligible, whereas oscillations are superimposed on the general sleep pattern. The latter is most obvious in *A. hippocastanea*. Interestingly, both species are classified in the *Sapindaceae* in the malvid clade, and both have palmate, deciduous leaves. However, the leaves of *A. hippocastanea* are much larger and therefore contribute more to the movement of the crown. *V. pragnense* and *O. europea* are also primarily showing sleep movement. The crown moves upwards in the early hours of the night and back down by morning. We were unable to distinguish drift or oscillations from the background noise. This was also the case for *P. x fraseri*, which nevertheless seems to have a short period of active oscillation movement before and especially after an 8-h period of sleep. In our study, *N. oleander* showed downward circadian movement, reminiscent of the movement identified by Puttonen et al. (2016) in *B. pendula*. The crown of *Albizia* shifted upwards until 8:30 p.m. and then downwards until it reached the lowest point at 03:00 a.m. After it had resumed the baseline position at 5:30 a.m., it continued upwards. This overall pattern in combination with leaf movement suggests that this species has a pronounced sleep cycle, but that the window of observation in our study was too short to completely capture it. These findings suggest that sleep movement can occur in trees both in downward direction as revealed by Puttonen et al. (2016) in *B. pendula*, as well as in the upward direction, or downward for part of the branches and upward for the rest as we observed for *Olea* and *Viburnum*. We can only speculate about the mechanisms leading to these differences and more plant physiological and biomechanical studies are needed to fully understand the driving mechanisms.

Clear cases of drift were revealed in *Morus alba* and *F. murielae*, with *Morus* moving upwards and *Fargesia* moving downwards. We cannot rule out that these plants later returned to their starting position since the experiment only lasted 12 h. The *Morus alba* tree revealed several signs of being unhealthy, which suggests that the unidirectional movement was a result of wilting of the plant. However, this was not evident in the *Fargesia murielae* and furthermore this bamboo showed signs of reversal to the starting configuration of the crown between 4:30 a.m. and 5:30 a.m. The fabaceous tree *Gleditschia* started moving downward in the evening and reached a low point at 8:00 p.m. Then it moved upward during the remaining part of the night and reached the most elevated position in the morning. This movement pattern can best be explained by downward sleep such as in *A. julibrissin*, but combined with an upward drift. Similarly the results suggest that *P. virginica* has upwards sleep movement. During early morning, the leaves started resuming the evening position. It is not clear whether they complete the full circle of movement, or whether the upward movement is repeated resulting in overall drift. In case of the palm tree *T. fortunei* the movement of the crown can be entirely attributed to the leaves. Our observations suggest an upward sleep movement for the upper part of the crown combined with a downwards drift of some of the lower part.

The sleep cycle of *B. utilis* is reminiscent of that described for *B. pendula* by Puttonen et al. (2016). It differs from this by having a barely distinguishable upwards drift superimposed.

Some trees revealed cyclic movements of shorter duration than the sleep movement, which we refer to here as oscillation. In two species we captured crown movements in 3 or 4 h cycles. This was most evident in *Magnolia alba* and somewhat obscure in *Q. robur*, being barely above the preset level set for background noise.

The gymnosperms mainly differed from angiosperms trees by the short-period fluctuations, which were more pronounced and spatially decoupled in the sense that the crown height percentiles did not show concerted movement such as it is more or less the case with angiosperms, but moved rather independently of each other. The results indicate that “sleep” movements may be weak or entirely absent in gymnosperms. The large individual of Austrian pine (*P. nigra* 1) revealed the clearest patterns of short-period movement of which individual pulses reached 0.6 cm. The small Austrian pine tree (*P. nigra* 2) had a similar pattern of erratic pulses deviating from the starting configuration at night. The individual percentiles of *C. libani* showed distinct peaks, whereas the mean of the percentiles was less clear and difficult to interpret. *L. kaempferi* was the only gymnosperm with movements reminiscent of “sleep” cycle. Superimposed on this pattern we captured an oscillation effect, which exceeded the preset background noise level for the individual percentiles as well as for the mean of all percentiles.

Comparison with Taxonomy, Morphology, and Anatomy

The two representatives for the malvid clade, *A. hippocastanea* and *A. palmatum* showed similar sleep patterns. Superimposed on this pattern *A. hippocastanea* showed conspicuous oscillation, which we attribute to the larger leaves of this species as compared to *A. palmatum*. *Olea* and *Viburnum* both have rigid, sclerophyllous leaves, which may explain why no short-period movement was observed in these plants. *Photinia* also seems to match this pattern of no short-period movement, at least during its sleep period, which is shorter than for other trees. The three Fabaceous trees probably have downward sleep cycles, but the duration of the experiment was probably too short so the drift observed may actually instead be part of a longer sleep cycle. This may also apply to the drift recorded in *P. virginica* and *B. utilis*, which may have sleep cycles lasting longer than the duration of the experiment.

Apparently gymnosperms have no sleep movement, except for the single case of the smaller *L. kaempferi* in which sleep movement cannot be entirely ruled out. *L. kaempferi* stands out among the gymnosperm species investigated by being deciduous and having needles with less supportive tissue. In general we saw a gradual dimming of pulsation movement from *P. nigra* through *C. libani* to *C. nootkatensis*, which is probably correlated to decreasing length of the needles. Interestingly, *M. grandiflora* is the only angiosperm where short-term oscillation is overriding the circadian rhythm. This species represents a basal lineage within the angiosperms and as such has an isolated taxonomic position.

To sum up: sleep is variable in both direction (upward or downward) and extent and apparently determined by ancestry to

a certain extent. Oscillation, which may be a universal feature, is probably controlled by the amount of supportive tissue in the leaves. Finally, drift is probably related to physiological phenomena such as, growth, senescence, and health of the plant.

New Finding and Hypotheses

The demonstration by Puttonen et al. (2016) of sleep movements in two provenances of *B. pendula* left the question outstanding whether it is a widespread phenomenon across seed plants. Our results indicate that several types of circadian crown movement occur and that some trees move more than others. Some of the movement patterns detected were quite unique but at the same time so close to the preset detection threshold that we abstain from sweeping conclusions, but instead encourage further studies. The fact that our experiment only lasted 12 h may in some instances have prevented us from capturing the full extent of the sleep cycles and instead led us to interpret the observed pattern as drift since the tree did not resume the starting position at the end of observation. Puttonen et al. (2016) hypothesize that the observed branch movements are independent from sunlight and governed by the internal circadian clock of the plant. Here, we also show ample cases of canopy movement taking place during the dark period of night after sunset and before sunrise. The results of the present study suggest that movements are somehow linked to ancestry. Biomechanics and ecophysiology are probably key drivers that interplay in complex ways, but this needs to be further explored before speculations are made.

Nevertheless nocturnal crown movement is a general characteristic of the plants monitored in this study. More specifically in nearly all the plants included we observed the same crown movement cycles of a few hours duration, which were also revealed (but not discussed) by Puttonen et al. (2016). Circadian movements in plants have been explained by attributes of the diurnal rhythm of photosynthesis and the resulting fluctuation in turgor pressure (Solomon et al., 2010). Does this imply that shorter cycles of movements can be attributed to short term fluctuations in turgor pressure of yet unknown causes? Sap flow measurements suggest that fluctuations in turgor pressure more or less follow the diurnal rhythm in light exposure of the plant (Chapotin et al., 2006; Steppe et al., 2015). However, Ehrenberger et al. (2012) revealed some instances of short cycle fluctuations in turgor pressures during the night.

Future Outlook

Our experiment is a trade-off between small-sized trees grown in containers under controlled conditions and fully developed trees growing in their natural environment where larger movements are expected due both to physiological changes and to wind. Due to the extremely high point densities the datasets generated in this study allow for more sophisticated analysis and as a consequence we have made all point clouds available online for further studies. Based on our results we clearly recommend that the duration of future experiments is more than 12 h, ideally 24 h. A draw-back of the laser scanning technique that we employed here is that the point cloud is rather unstructured. At the given resolution it is not trivial to keep track of individual parts of the trees

and compare these across the species studied. Puttonen et al. (2016) solve this by affixing standard measurement targets to some branches, which are then automatically located in post-processing. Alternatively, several processing techniques have been developed with the purpose of creating a 3-dimensional model from point clouds of trees obtained by TLS (Xu et al., 2007; Raunonen et al., 2013; Hackenberg et al., 2015), and these could potentially allow determining and tracking similar parts of the tree in consecutive scans to track their movement. However, such methods typically focus on modeling the branch structure without the leaves, which might be a limitation for detecting canopy movement. Alternatively, leaf segmentation approaches are also promising for detecting change maybe even at the individual leaf level (Béland et al., 2014; Koma et al., 2017). It seems feasible to alternatively use reflective prism foil for marking specific points on the branches, which can then be detected based on signal amplitude. Such a setup would furthermore open up for using a geodetic total station that can automatically locate and accurately measure the position of the reflective markers only. This would probably be a cheaper and more practical approach while delivering data that allows for similar or even more detailed spatial statistical analysis.

From a plant physiological point of view, the most daunting challenge is to explain the short-term cycles of movement. This can be done by using non-invasive methods to measure oscillations in water potential. The question is whether these are sufficiently sensitive to measure relatively small differences within a short time span. Another option is semi-invasive, measurement of conductivity between electrodes inserted into the trunk has been proposed as a metric of sap flow (Koppán et al., 2000). High precision measurements of tree girth may also give an indication of sap flow and pressure regulation but the question is again whether it is sufficiently precise enough to reveal any patterns (Pesonen et al., 2004). Finally, more invasive methods could be considered involving anatomical sectioning of fresh wood sampled at different phases of crown movement and immediately frozen in liquid nitrogen for later microanalysis of sugar contents at the tissue level.

Recent advances in laser technology open up for detailed studies of spatial dynamics in plants such as presented here.

REFERENCES

- Barak, S., Tobin, E. M., Green, R. M., Andronis, C., and Sugano, S. (2000). All in good time: the *Arabidopsis* circadian clock. *Trends Plant Sci.* 5, 517–522. doi: 10.1016/S1360-1385(00)01785-4
- Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A., and Verstraete, M. M. (2014). On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. *Agric. For. Meteorol.* 184, 82–97. doi: 10.1016/j.agrformet.2013.09.005
- Bretzl, H. (1903). *Botanische Forschungen des Alexanderzuges*. B.G.Teubner, Leipzig.
- Chapotin, S. M., Razanameharizaka, J. H., and Holbrook, N. M. (2006). Water relations of baobab trees (*Adansonia* spp. L.) during the rainy season: does stem water buffer daily water deficits? *Plant Cell Environ.* 29, 1021–1032. doi: 10.1111/j.1365-3040.2005.01456.x

Since studies like this are still in their infancy they raise more questions that they answer. However it is very clear that plant movement is understudied, and that a set of promising methods exist that, in combination with TLS, can be expected to provide fundamental new knowledge on the physiology of plants. Our results give an impression of complex relationships that are controlled by many drivers and perhaps when mapped in detail reveal information on the plants adaptation to the environment and its interactions with other species. We have tried to reduce this complexity by establishing a rough conceptual framework within which phenomena can more easily be described.

AUTHOR CONTRIBUTIONS

AZ planned and coordinated the study, performed data processing and analysis and wrote most of the main text. BM carried out the laser scanning measurements and the pre-processing of the point clouds, and wrote the methodology of the scanning. AB contributed to the analysis and interpretation of the results and to the discussion section. All authors worked together to produce the final version of the text.

FUNDING

AZ was supported by NKFIH OTKA PD 115833 and joint work on data analysis was allowed by a study visit to Aarhus University funded by the Department of Ecoinformatics and Biodiversity and OTKA ERC_16_M 122670.

ACKNOWLEDGMENTS

We gratefully acknowledge the collaboration of the “Sylvestris” garden shop staff in support of our measurements, and the assistance of Katharina Prager (BOKU, Vienna) during the experiment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.01814/full#supplementary-material>

- Chase, M. W., Christenhusz, M. J. M., Fay, M. F., Byng, J. W., Judd, W. S., Soltis, D. E., et al. (2016). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* 181, 1–20. doi: 10.1111/boj.12385
- Darwin, C., and Darwin, F. (1880). *The Power of Movement in Plants*. London: John Murray.
- Dassot, M., Constant, T., and Fournier, M. (2011). The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. *Ann. For. Sci.* 68, 959–974. doi: 10.1007/s13595-011-0102-2
- Ehrenberger, W., Rüger, S., Fitzke, R., Vollenweider, P., Günthardt-Goerg, M., Kuster, T., et al. (2012). Concomitant dendrometer and leaf patch pressure probe measurements reveal the effect of microclimate and soil moisture on diurnal stem water and leaf turgor variations in young oak trees. *Funct. Plant Biol.* 39, 297–305. doi: 10.1071/FP11206

- FARO (2013). *FARO Focus3D Features, Benefits and Technical Specifications*. Lake Mary, FL: FARO Technologies.
- Glira, P., Pfeifer, N., Brieke, C., and Ressler, C. (2015). A Correspondence Framework for ALS Strip Adjustments based on Variants of the ICP Algorithm Korrespondenzen für die ALS-Streifenausgleichung auf Basis von ICP. *Photogramm. Fernerkund. Geoinformation* 2015, 275–289. doi: 10.1127/pfg/2015/0270
- Hackenberg, J., Spiecker, H., Calders, K., Disney, M., and Raunonen, P. (2015). Simpletree—an efficient open source tool to build tree models from TLS clouds. *Forests* 6, 4245–4294. doi: 10.3390/f6114245
- Koma, Z., Rutzinger, M., and Bremer, M. (2017). Automated estimation of leaf distribution for individual trees based on TLS point clouds. *Presented at the EGU General Assembly Conference Abstracts*, 5329.
- Koppán, A., Szarka, L., and Westergom, V. (2000). Annual fluctuation in amplitudes of daily variations of electrical signals measured in the trunk of a standing tree. *CR Acad. Sci. III* 323, 559–563. doi: 10.1016/S0764-4469(00)00179-7
- McClung, C. R. (2013). Beyond arabidopsis: the circadian clock in non-model plant species. *Semin. Cell Dev. Biol.* 24, 430–436. doi: 10.1016/j.semcdb.2013.02.007
- Mullen, J. L., Weinig, C., and Hangarter, R. P. (2006). Shade avoidance and the regulation of leaf inclination in Arabidopsis. *Plant Cell Environ.* 29, 1099–1106. doi: 10.1111/j.1365-3040.2005.01484.x
- Pesonen, E., Mielikäinen, K., and Mäkinen, H. (2004). A new girth band for measuring stem diameter changes. *For. Int. J. For. Res.* 77, 431–439. doi: 10.1093/forestry/77.5.431
- Pfautsch, S., Hölttä, T., and Mencuccini, M. (2015). Hydraulic functioning of tree stems—fusing ray anatomy, radial transfer and capacitance. *Tree Physiol.* 35, 706–722. doi: 10.1093/treephys/tpv058
- Puttonen, E., Brieke, C., Mandlbürger, G., Wieser, M., Pfennigbauer, M., Zlinszky, A., et al. (2016). Quantification of overnight movement of birch (*Betula pendula*) branches and foliage with short interval terrestrial laser scanning. *Front. Plant Sci.* 7:222. doi: 10.3389/fpls.2016.00222
- Raunonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., et al. (2013). Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sens.* 5, 491–520. doi: 10.3390/rs5020491
- Solomon, O. L., Berger, D. K., and Myburg, A. A. (2010). Diurnal and circadian patterns of gene expression in the developing xylem of Eucalyptus trees. *South Afr. J. Bot.* 76, 425–439. doi: 10.1016/j.sajb.2010.02.087
- Steppe, K., Sterck, F., and Deslauriers, A. (2015). Diel growth dynamics in tree stems: linking anatomy and ecophysiology. *Trends Plant Sci.* 20, 335–343. doi: 10.1016/j.tplants.2015.03.015
- Xu, H., Gossett, N., and Chen, B. (2007). Knowledge and heuristic-based modeling of laser-scanned trees. *ACM Trans. Graph.* 26:19. doi: 10.1145/1289603.1289610
- Zweifel, R., Item, H., and Häslér, R. (2001). Link between diurnal stem radius changes and tree water relations. *Tree Physiol.* 21, 869–877. doi: 10.1093/treephys/21.12-13.869

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer FB and handling Editor declared their shared affiliation.

Copyright © 2017 Zlinszky, Molnár and Barford. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Aerial Images and Convolutional Neural Network for Cotton Bloom Detection

Rui Xu¹, Changying Li^{1*}, Andrew H. Paterson², Yu Jiang¹, Shangpeng Sun¹ and Jon S. Robertson²

¹ Bio-Sensing and Instrumentation Lab, College of Engineering, University of Georgia, Athens, GA, United States, ² Plant Genome Mapping Laboratory, Department of Genetics, University of Georgia, Athens, GA, United States

OPEN ACCESS

Edited by:

Norbert Pfeifer,
Technische Universität Wien, Austria

Reviewed by:

Maria Balota,
Virginia Tech, United States
Patrick Tutzauer,
University of Stuttgart, Germany

*Correspondence:

Changying Li
cyli@uga.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 30 September 2017

Accepted: 19 December 2017

Published: 16 February 2018

Citation:

Xu R, Li C, Paterson AH, Jiang Y,
Sun S and Robertson JS (2018) Aerial
Images and Convolutional Neural
Network for Cotton Bloom Detection.
Front. Plant Sci. 8:2235.
doi: 10.3389/fpls.2017.02235

Monitoring flower development can provide useful information for production management, estimating yield and selecting specific genotypes of crops. The main goal of this study was to develop a methodology to detect and count cotton flowers, or blooms, using color images acquired by an unmanned aerial system. The aerial images were collected from two test fields in 4 days. A convolutional neural network (CNN) was designed and trained to detect cotton blooms in raw images, and their 3D locations were calculated using the dense point cloud constructed from the aerial images with the structure from motion method. The quality of the dense point cloud was analyzed and plots with poor quality were excluded from data analysis. A constrained clustering algorithm was developed to register the same bloom detected from different images based on the 3D location of the bloom. The accuracy and incompleteness of the dense point cloud were analyzed because they affected the accuracy of the 3D location of the blooms and thus the accuracy of the bloom registration result. The constrained clustering algorithm was validated using simulated data, showing good efficiency and accuracy. The bloom count from the proposed method was comparable with the number counted manually with an error of -4 to 3 blooms for the field with a single plant per plot. However, more plots were underestimated in the field with multiple plants per plot due to hidden blooms that were not captured by the aerial images. The proposed methodology provides a high-throughput method to continuously monitor the flowering progress of cotton.

Keywords: cotton, flower, bloom, unmanned aerial vehicle, point cloud, convolutional neural network, phenotyping

INTRODUCTION

Improving cotton yield is one of the main objectives in cotton breeding projects and cotton growth and production management. Yield can be defined as the product of the number of bolls produced per unit area and the mass of lint per boll (a common measure of boll size). Cotton yield is associated with many physiological variables and environmental factors. However, an increase in cotton yield is generally associated with an increase in the number of bolls regardless of genotype or environment (Wells and Meredith, 1984; Pettigrew, 1994). Flower and boll retention will affect the final number of bolls produced. Therefore, processes that affect flower and boll retention will have a significant impact on yield.

Flowering is important to cotton yield because pollinated flowers form cotton bolls. Within a given genotype, seasonal flower production per unit area is more closely related to yield than boll retention percentage (Heitholt, 1993, 1995). Since a cotton flower is white (cream-color for some upland germplasms and yellow for Pima) on the first day it opens and turns pink within 24 h, it is unlikely to mistake an old bloom for a new bloom on separate days. Therefore, if bloom counts are obtained daily then the seasonal total counts can be calculated. The flowering time (the time of the first flower opening) and the peak flowering time can be determined accordingly, both of which are critical to production management. The timing and duration of the flowering stage also reflect the difference in growth habits of different genotypes, which can help breeders select specific genotypes, for example, short-season or long-season genotypes. Although manual counting is perhaps the simplest and most reliable way to count flowers, it is tedious and inefficient and requires a massive amount of labor, which is not practical for large fields.

Imaging methods can improve the efficiency of manual counting. Many studies have been done on flower detection and classification using color images (Adamsen et al., 2000; Siraj et al., 2010; Hsu et al., 2011; Biradar and Shrikhande, 2015; Seeland et al., 2016; Thorp et al., 2016). Flowers usually have distinct color and texture from the background; therefore, traditional image processing methods such as color and texture analysis can be used to segment flower pixels (Thorp et al., 2016). The number of flowers can be calculated using morphological operations on the segmented flower pixels or correlated with flower pixel percentage (Biradar and Shrikhande, 2015; Thorp et al., 2016). Flower features can be extracted from flower images and are used to recognize flower species (Hsu et al., 2011). Machine learning techniques can be used to classify different flower types, which could be useful to determine the age of cotton flowers based on differences in color and shape (Seeland et al., 2016). Deep learning methods such as the convolutional neural network (CNN) have been demonstrated to be effective in recognizing flower species (Liu et al., 2016). CNN showed advantages over traditional machine learning methods because it does not require extraction of image features.

Although imaging methods are efficient in detecting flowers, the image collection throughput limits its usage in agriculture because agriculture usually deals with large fields. To improve the data collection throughput, the use of an unmanned aerial vehicle (UAV) is usually preferred over a ground vehicle or a robot because a UAV can provide superior data collection speed and larger spatial coverage. UAVs also do not interact with the plants, so constant data collection will not cause soil compaction and plant damage, which can happen when using a ground vehicle. Although many researchers have used UAVs for agriculture studies in recent years, only a few used UAVs to count flowers for crops. For example, aerial multispectral imaging has been used to calculate flower fraction in oilseed rape and monitor the peach flower (Fang et al., 2016; Horton et al., 2017). However, those studies only segmented flowers from the canopy rather than counting the flowers. To our knowledge, no study has been done to count flowers using aerial images. The main reason for this is

the low image resolution compared to images taken from ground platforms, making image processing challenging.

In this paper, we aimed to develop a methodology for counting cotton blooms using aerial color images. The overall objective of this paper was to develop a data processing pipeline to detect and count the number of newly opened cotton flowers using aerial images. Specific objectives were to: (1) build and train a CNN to classify flowers, (2) construct dense point clouds from raw images and evaluate the quality, (3) develop an algorithm to register flowers, and analyze its accuracy and efficiency using simulated data, and (4) evaluate the performance of the data processing pipeline compared with manual counting. To avoid ambiguity, we used the term bloom to refer to a newly opened flower to distinguish it from its other growth stages.

MATERIALS AND METHODS

Test Fields

Two test fields were used, both of which were located in the Plant Science Farm in Watkinsville, GA (33°43'37.80"N, 83°17'57.52"E) (**Figure 1**). Field 1 had 132 plots with a single cotton plant in each plot, arranged in 12 rows and 11 columns with a 1.5 m (5 ft) distance in both row and column directions. The cotton in field 1 was planted on May 25, 2016. Field 2 had 128 plots of cotton, arranged in 16 rows and 8 columns (**Figure 1**). Each plot was 3 m (10 ft) long and 1.5 m (5 ft) wide, with a 1.8 m (6 ft) alley between each plot. Fifteen cotton plants were planted in each plot with equal spacing of 0.15 m (6 inches) on June 13, 2016. The plant density varied because of the different number of plants germinated, resulting in some empty plots in field 1 and field 2. The number of germinated plants in each plot for field 2 was recorded 2 weeks after planting (June 8, 2016) to be used to calculate the average number of blooms for each plant.

Data Collection

Aerial color images of the test fields were collected using an octocopter (s1000+, DJI, China) with a color camera (Lumix DMC-G6, Panasonic, Japan) on August 12, August 19, August 26, and September 9, 2016 (**Table 1**). The color camera was directly mounted on the bottom of the drone with the lens facing downward. An inertial measurement unit (IMU) with GPS was mounted on the drone to record the location and orientation of the camera. Raspberry Pi 2 was used to trigger the camera and record the IMU/GPS measurement. During the flight, the camera took images at a frequency of 1 Hz. The exposure time and aperture were manually set for the color camera according to the light conditions of the field, and auto-settings were used for other parameters. The drone was flown at a height of 15 m above ground level (AGL) to reduce the ground sample distance (GSD). One flight was used to collect data from both of the test fields.

The number of blooms in each plot was manually counted as a reference for the image method. To count the blooms in each plot, only white flowers were counted in each plot on the same day that aerial images were collected. Although manual counting is reliable overall, it is possible for some blooms to be counted multiple times or not counted. The judgment of whether a flower is white is subjective and varies from person to person, which

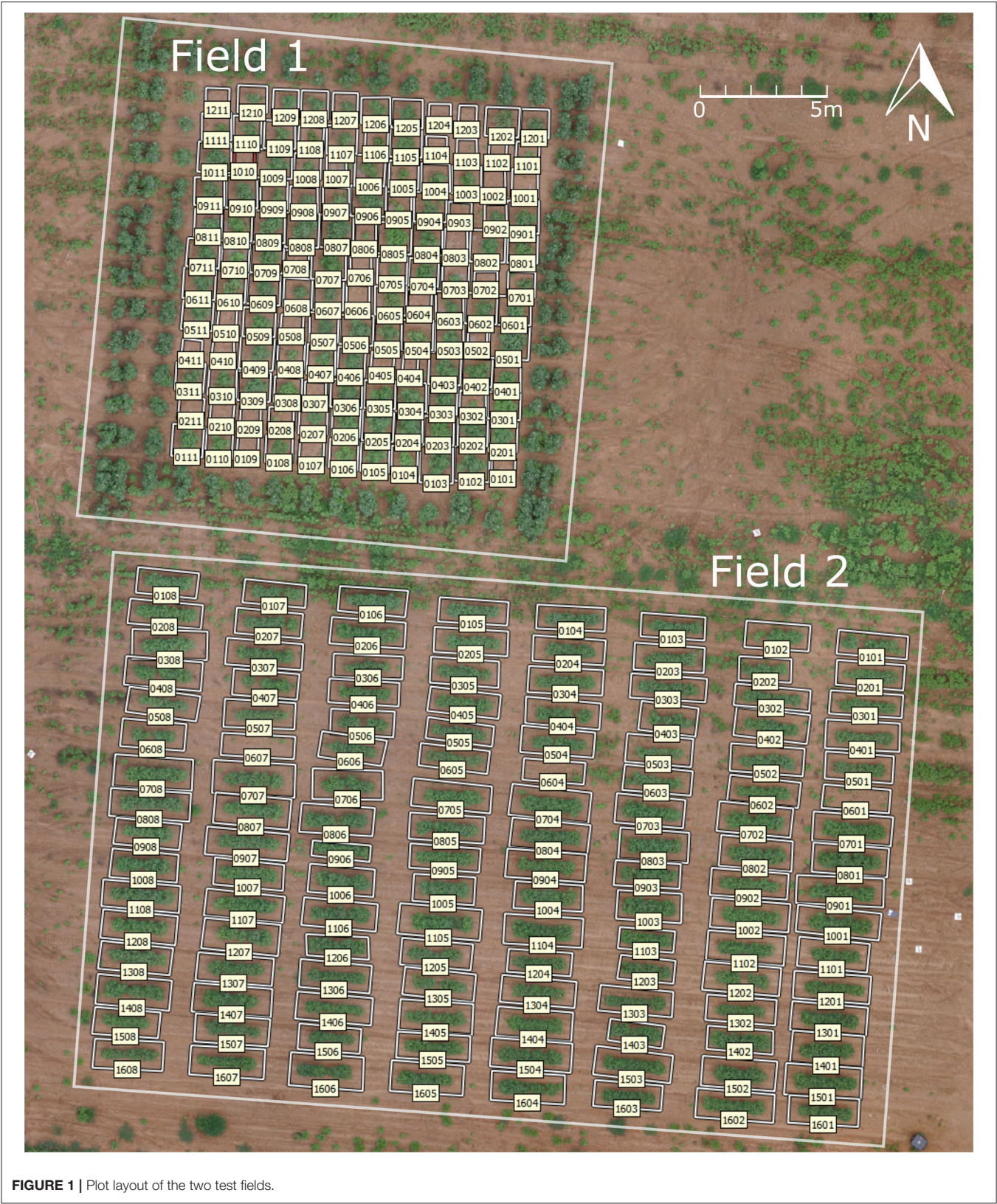


FIGURE 1 | Plot layout of the two test fields.

TABLE 1 | Data collection summary.

Date	Time (p.m.)	Flight height (m)	Flight speed (m/s)	Focal length (mm)	Ground pixel size (mm)
8/12/2016	12:09	15	2.5	18	3.17
8/19/2016	1:43	15	1.5	20	2.23
8/26/2016	1:35	15	3	20	2.69
9/9/2016	1:12	15	2.5	22	2.36

could cause human counting errors. For field 1, blooms in all of the plots were counted on all 4 days. For field 2, blooms were counted in all plots on August 12, and in 32 random plots on the other 3 days.

Data Processing Pipeline

The overall data processing pipeline for bloom counting can be divided into four key steps: (1) dense point cloud generation, (2) plot images extraction, (3) bloom detection, and (4) bloom registration (**Figure 2**). The final output of the pipeline is the bloom count for each plot. Other information such as bloom position can be obtained too.

Dense Point Cloud Generation

A dense point cloud of each test field was generated in PhotoScan (PhotoScan Professional 1.3.1, Agisoft, Russia) using the raw color images and the IMU/GPS measurements. The high accuracy and reference preselection settings were used for photo alignment, the high-quality setting for dense point cloud build, and default for other settings. After photo alignment, PhotoScan constructed a tie point cloud—often called a sparse point cloud in other software—from the feature points detected in the images. Before constructing the dense point cloud, the tie point cloud was used to calculate the accuracy of the dense point cloud.

After building the dense point cloud, it was found that field 1 was not fully covered by the images collected on August 19 and those uncovered plots were excluded from further data analysis. Additionally, some plots had an incomplete point cloud for the canopy due to low side-overlap on August 19, August 26, and September 9, when the focal length was enlarged to reduce the ground pixel size. Point cloud coverage, which is the percentage of the plot constructed with valid point cloud, was used to evaluate the completeness of the point cloud. The plots with point cloud coverage less than 80% were excluded from further analysis. The method to calculate point cloud coverage is explained in section Quality of the dense point cloud.

Plot Images Extraction

The main purpose of the plot images extraction step is to export plot images from the raw images such that one plot image contains only one plot or part of the plot. Since one plot can be imaged several times, one plot has multiple plot images from different raw images. This is helpful for detecting blooms inside the canopy because different raw images provide views of each plot from different angles, which greatly improves the chance of one bloom being imaged compared to the ortho-image, which only shows the top view of the canopy. To extract the plot

images, plot boundaries were first manually drawn for each plot and stored as quadrangle shapes in PhotoScan. Then the four vertices of each shape were projected to each raw image to get the pixel location of the vertices in each image. One image was considered to cover part of the plot if two or three vertices were in the image (pixel location is within the boundary of the image), and the whole plot if four vertices were in the image. The pixel location of the four vertices was recorded and plot images were extracted from the raw image accordingly. The step was processed in PhotoScan using built-in Python functions. The image file name, plot number, and pixel locations of the four vertices were saved in a text file and imported into MATLAB (MATLAB 2017a, MathWorks, USA) to extract plot images from raw images.

Bloom Detection

The bloom detection process contains two steps. The first step is to screen out the locations that most likely contain cotton blooms based on the fact that cotton blooms usually have the highest pixel intensity because of their white color. The plot image was first transferred into CIELAB color space and then the screening was performed using the L channel. A threshold of 0.75 on the normalized L channel (normalized by the maximum value in L channel) was used to extract bright objects. The number of pixels for each object was counted and noisy small objects less than 15 pixels were removed. Objects separated by a distance of less than the diameter of a flower—which is about 20 pixels for our dataset—were combined into one object because some flowers were split into two objects by the leaves. Images with a size of 36 by 36 pixels around the center of the remaining objects were extracted from the plot image. Those images were treated as potential bloom images.

The second step is to classify the potential bloom images into bloom and non-bloom class using the pre-trained CNN (detailed information on the CNN and training process is described in section Convolutional neural network). After classification, the images classified as bloom class were kept and the pixel location of those images in the raw color images was recorded.

Bloom Registration

Since the same bloom can be detected in several images, it is necessary to register the bloom before counting the blooms to prevent counting the same bloom more than once. Registration methods based on image features were not useful because very few features could be detected from the bloom images due to the resolution limit. Therefore, we first projected the pixel location of each detected bloom into the 3D location in the dense point cloud, and then clustered the blooms based on the locations because the same bloom should have the same 3D location in the dense point cloud. The projection was done in PhotoScan. Because each pixel in the image has a corresponding image ray, any point in the image ray will appear as the same pixel in the image, and the 3D location of that pixel in the point cloud is the intersection between the image ray and the point cloud. If there is no intersection between the image ray and the point cloud, PhotoScan will return the closest point to the image ray.

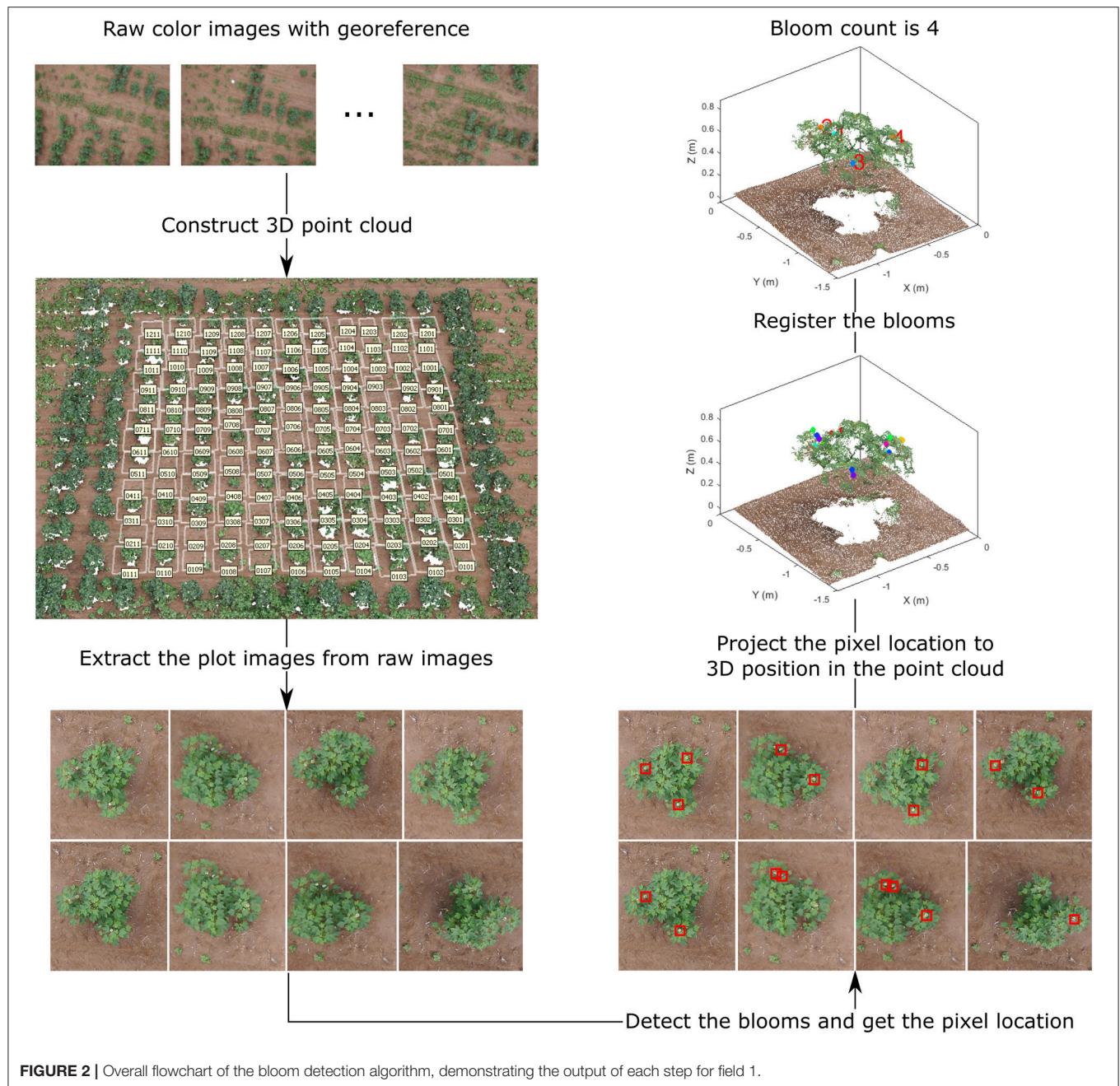


FIGURE 2 | Overall flowchart of the bloom detection algorithm, demonstrating the output of each step for field 1.

Due to the accuracy of the point cloud and pixel location, the 3D location of the bloom may deviate from the true location, which increases the chance of incorrect bloom registration solely based on the location. However, since multiple blooms detected in the same image are different blooms and cannot be in the same cluster, the bloom registration based on the 3D position can be generalized as a constrained clustering problem as follows:

Given the set of data points and the set of their corresponding classes, form the data points into clusters so that no data points in the same class will be in the same cluster and the distance between any two clusters are not smaller than a threshold.

In the case of bloom registration, the data point is the bloom position and the class is the image number to indicate in which image the bloom was detected. Although existing constrained clustering algorithms can solve the problem, most of the algorithms are for general clustering problems and are not efficient for this specific problem. Therefore, inspired by hierarchical clustering, a deterministic clustering algorithm was designed specifically for the bloom registration problem. The algorithm initializes with each data point as one cluster. The algorithm involves cluster selection and merging. First, for each cluster i , a set of clusters, S , was selected from all of the clusters such that no data points in S had the same class as any data

point assigned to cluster i . Second, the distance between cluster i and every cluster in \mathcal{S} was calculated and cluster i was merged to the closest cluster in \mathcal{S} if their distance was smaller than the threshold λ . The algorithm repeats the selection and merging until no merging happens. The algorithm has a similar effect as the regular hierarchical clustering algorithms when the distance between two clusters is measured using Equation (1).

$$\text{dist}(\mu_i, \mu_j) = \begin{cases} \text{dist}'(\mu_i, \mu_j), & \text{if } E_i \cap E_j = \emptyset \\ +\infty, & \text{otherwise} \end{cases} \quad (1)$$

because the distance calculation is only performed on a subset of the clusters.

Euclidean distance was used to measure the cluster dissimilarity. However, because of the large error of the 3D position projection on the z-axis, the Euclidean distance was modified by adding a weight term on the z-axis to adjust the influence of z-axis on the distance. The weighted Euclidean distance between $a = (a_x, a_y, a_z)$ and $b = (b_x, b_y, b_z)$ was calculated using Equation (2). The final program used 0.5 for w .

$$\text{dist}(a, b) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2 + w^2 (a_z - b_z)^2} \quad (2)$$

Bloom registration algorithm

Input: Set of data points $\mathcal{D} = \{x_i\}_i^n$, class $E = \{e_i\}_i^n$, and distance threshold λ .

Result: The assignment variables for the data points, $z = [z_1, z_2, \dots, z_n]$.

Initialization: each point is assigned as one cluster, $z_1 = 1, z_2 = 2, \dots, z_i = i, \dots, z_n = n$

while not converged do

for $i = 1 : n$ **do**

Find cluster i and its center μ_i

if cluster i exists **then**

$\mathcal{S} \leftarrow \emptyset$;

Given the current assignment of points, find the set of classes of the data points assigned to cluster i ,

$E_i = \{e_k \in E | z_k = i\}$;

for $j = 1 : n$ **do**

Find cluster j and its center μ_j

if cluster j exists **then**

Given the current assignment of points, find the classes of the data points assigned to cluster j ,

$E_j = \{e_k \in E | z_k = j\}$;

if $E_i \cap E_j = \emptyset$ **then**

// Add the current cluster to the selected clusters.

$\mathcal{S} \leftarrow \{\mathcal{S} \cup \mu_k\}$;

end

end

end

end

$[d_{min}, i_{min}] \leftarrow \min \{\text{dist}(\mu_i, s_1), \dots, \text{dist}(\mu_i, s_{|\mathcal{S}|})\}$;

if $d_{min} < \lambda$ **then**

//Merge cluster i with i_{min} and update the assignment variables;

Assign i_{min} to the assignment variables for the data points in cluster i ;

end

end

for $k = 1 : n$ **do**

Given the current assignment of points, find the set of points assigned to cluster k , $\mathcal{D}_k = \{x_i \in \mathcal{D} | z_i = k\}$;

if $|\mathcal{D}_k| > 0$ **then**

$\mu_k \leftarrow \frac{\sum_{x_i \in \mathcal{D}_k} x_i}{|\mathcal{D}_k|}$

else

Remove the cluster and apply the changes to the related variables ;

end

end

end

where μ_i and μ_j are the center of the cluster i and j , $\text{dist}'(\mu_i, \mu_j)$ is the distance metric used in the hierarchical clustering algorithms, and E_i and E_j are the set of classes of the data points assigned to cluster i and j . Compared to the regular hierarchical clustering algorithms, the efficiency of our algorithm is improved

Convolutional Neural Network

The CNN was used to classify potential bloom images. The network had seven layers, one input layer, two convolutional layers, two max pooling layers, and two fully connected layers (Figure 3). Since the classification was performed on

the potential bloom images, the potential bloom images were first extracted from individual data sets and manually labeled into bloom and non-bloom classes. Then the labeled potential bloom images were used to construct the image database. In total, ~14,000 images for the bloom class and ~60,000 images for the non-bloom class were extracted. Two thirds of the bloom images (roughly 9,000 images) and the same number of non-bloom images were randomly selected as training samples. The remaining images were used as testing samples. The CNN was trained for 30 epochs. The learning rate was set to 0.01 at the first epoch and decreased by a factor of 10 every 10 epochs. The mini batch size was set to 256. The regularization factor was set to 0.01 to prevent overfitting.

Quality of the Dense Point Cloud

The quality of the dense point cloud of the canopy greatly affects the accuracy of the 3D position of each bloom, and thus the clustering results, and ultimately the bloom numbers. The quality of the dense point cloud can be evaluated from two aspects—the accuracy and the completeness of the dense point cloud.

The position accuracy of each point in the point cloud is difficult to calculate because the geometry of the canopy is unknown. Instead, the projection and reprojection error were calculated for the tie points to estimate the overall accuracy of the dense point cloud. The projection error was calculated as the distance between the projection of the feature points to the point cloud and their corresponding tie points. The reprojection error was calculated as the pixel distance between the projection of the tie points on the images and their original corresponding feature points. The projection error was analyzed on each single axis (easting, northing, and elevation) and the summation of three axes.

The completeness of the dense point cloud quantifies how completely the point cloud represents the real object or scene, and can be measured by the density of the point cloud. The completeness of the point cloud on the easting-northing

plane was calculated since the camera only captures the top view. First, the point cloud was rasterized into a 2D elevation map using grid steps of 1.3 cm without interpolating the empty cells. The elevation map was divided into each plot. For each plot, the ground surface was calculated using the Maximum Likelihood Estimation Sample Consensus (MLSAC) (Torr and Zisserman, 2000). The plot was divided into ground and canopy using the elevation map based on the distance to the ground surface with a threshold of 0.1 m. The point cloud coverage (PCC) for the canopy is defined in Equation (3).

$$PCC = \frac{\text{Area of the canopy}}{\text{Area of the canopy} + \text{Area of the empty cells}} \quad (3)$$

The area of the empty cells was included in the denominator because the empty cells were usually part of the canopy. The PCC estimates the completeness with which the canopy is represented by the dense point cloud, which can greatly affect the bloom registration result.

Evaluation of the Bloom Registration Algorithm

To evaluate the efficiency and accuracy of the bloom registration algorithm, artificial data was generated in three steps. First, a data point with three dimensions was generated by randomly drawing integers from 0 to 9 for each axis. This step was repeated until 125 different data points were generated. The same class number was assigned to the 125 data points to simulate that they were from the same image. Second, the first step was repeated 10 times, generating in total 1,250 data points. A new class number was assigned to the 125 data points at each repeat to simulate that the data points from each repeat were from different images. The data points with the same coordinates were set in the same cluster, which is the true clustering result. Third, a random error that followed $\mathcal{N}(0, \frac{\sqrt{3}}{3}\sigma)$ was added to each axis of each data point, thus the position error (vector summation of the errors of the

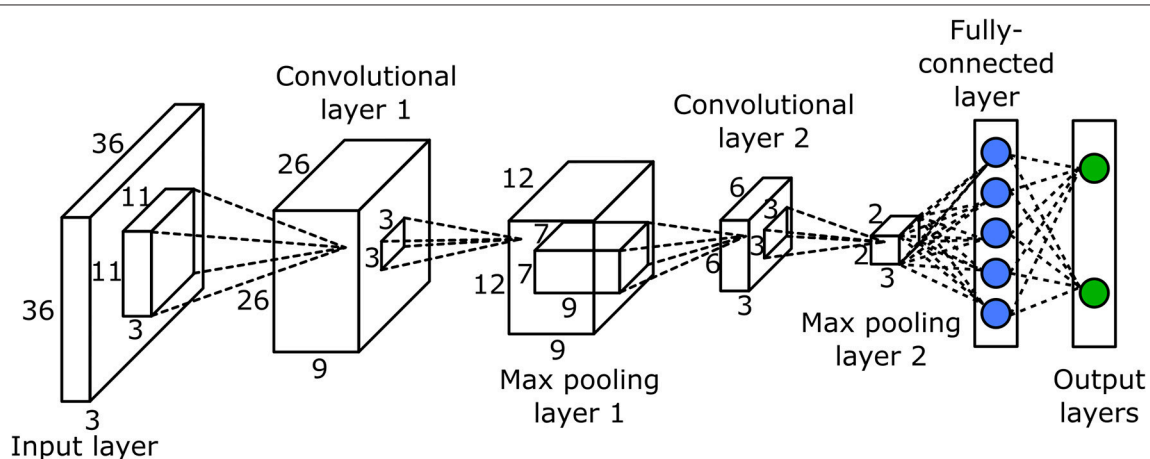


FIGURE 3 | Structure of the convolutional neural network.

three axis) followed $\mathcal{N}(0, \sigma)$. Different σ was used to simulate the noise level of the 3D position of the bloom. The 1,250 data points were clustered using the bloom registration algorithm and compared with the true clustering result.

Since the clustering result will depend on the noise level of the data and threshold for the clustering algorithms, 10 different σ values from 0.1 to 0.5 with an interval of 0.05, and 5 threshold values were tested. The threshold values were σ , 1.5σ , 2σ , 2.5σ , and 3σ and they covered from 68 to 99.7% of the position error. To acquire statistics of the result, the third step of generating artificial data was repeated 10 times to generate 10 sets of artificial data for each noise level and threshold. The mean and standard deviation of the clustering error and runtime on the 10 data sets were analyzed. The misclustering rate was defined using equation 4.

$$\text{misclustering rate} = \frac{\text{number of misclustered points}}{\text{total number of points}} \quad (4)$$

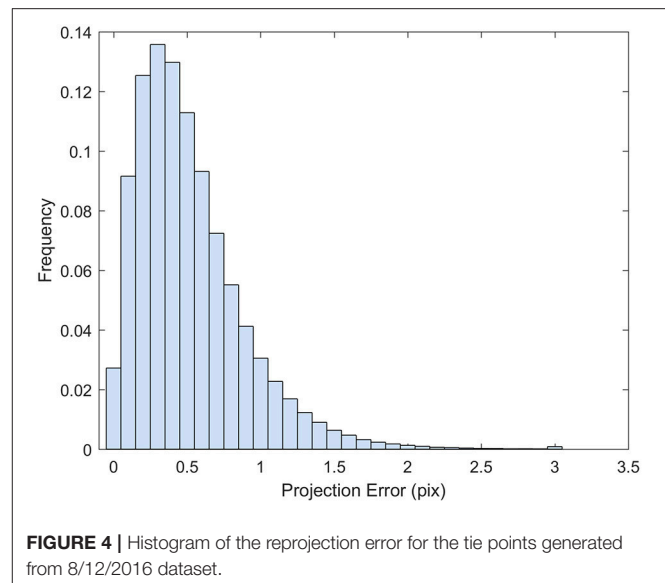
Since it is difficult to relate the clustering result with the ground truth, the misclustered points were considered as the points that were in the same cluster from the clustering result but not in the same cluster from the ground truth. The number of misclustered points was calculated with two approaches. The first approach—referred to as misclustering rate by reference—used the true clustering result as the ground truth, and the second approach—referred to as misclustering rate by result—used the clustering results as the ground truth. The first approach was good at evaluating the clustering error when the clustering algorithm over clustered the data points, while the second approach was suitable for under-clustering errors.

RESULT

Quality of the Dense Point Cloud

The point cloud accuracy was analyzed using the dataset collected on 8/12. More than 99% of the reprojection errors of the tie points were less than 1.7 pixels and only a few were larger than 2 pixels (Figure 4). The mean reprojection error was 0.5 pixels. The tie points had overall larger projection error on the elevation than the easting and northing (Figure 5). The mean projection error for the easting and northing was 0 m because the positive and negative errors were canceled out, whereas the mean projection error for the elevation was 0.014 m, which is reasonable since the depth generated from multi-view stereo is usually the least accurate. The elevation also had larger variation than the easting and northing. The mean projection error for the three axes combined was 0.022 m, and 99% of the errors were between 0 and 0.127 m (Figure 5D). The large projection error on the elevation validates the rationale to assign a weight to the elevation when calculating the cluster distance in bloom registration.

In the event of an incomplete dense point cloud for the canopy, some blooms may not be able to get 3D point positions and thus cannot be counted. If the blooms equally distribute over the canopy, the point cloud coverage can be considered as the probability of a bloom having a valid point in the dense point cloud. Assuming a canopy has n blooms and point cloud coverage



of p , then k ($k = 1, 2, \dots, n$) blooms not having valid points (which cannot be counted using the imaging method) follow a binomial distribution $B(n, 1 - p)$. The mean count error is $n(1 - p)$ and the relative error is $(1 - p)$. Therefore, the point cloud coverage should be close to 1 to make the counting error small.

Figure 6 showed that the 8/12 dataset had good point cloud coverage (> 0.9) on most of the plots for both test fields. However, the other three datasets showed low point cloud coverage on both fields, except that the 8/19 dataset had some plots with good point cloud coverage. The low point cloud coverage was mainly due to the insufficient image side-overlap ($< 60\%$); the PhotoScan was unable to construct valid 3D points for areas that were covered by less than three images. Therefore, increasing the image overlap can improve the point cloud coverage. The depth filter inside the PhotoScan removed noisy points due to the movement of the plants and image noise, which was another reason for the low coverage. To achieve low image counting error, the plots with point cloud coverage lower than 0.8 were excluded from data analysis, which removed 16 plots in field 1 for the 8/12 dataset, 18 plots in field 1 and 103 plots in field 2 for the 8/19 dataset, all the plots in field 1 and 112 plots in field 2 for the 8/26 dataset, and 127 plots in field 1 and 97 plots in field 2 for the 9/9 dataset.

Training Result of the Convolutional Neural Network (CNN)

With more than 28,000 training images and mini batch size of 256, the CNN was trained about 114 iterations every epoch. Because of the large training sample size and relatively simple structure, the training process converged quickly and the training accuracy reached 0.94 in the first epoch. As shown in Figure 7, the training accuracy for bloom and non-bloom class increased in the first few epochs. The training accuracy for the non-bloom class reached a maximum on the 8th epoch, and the training

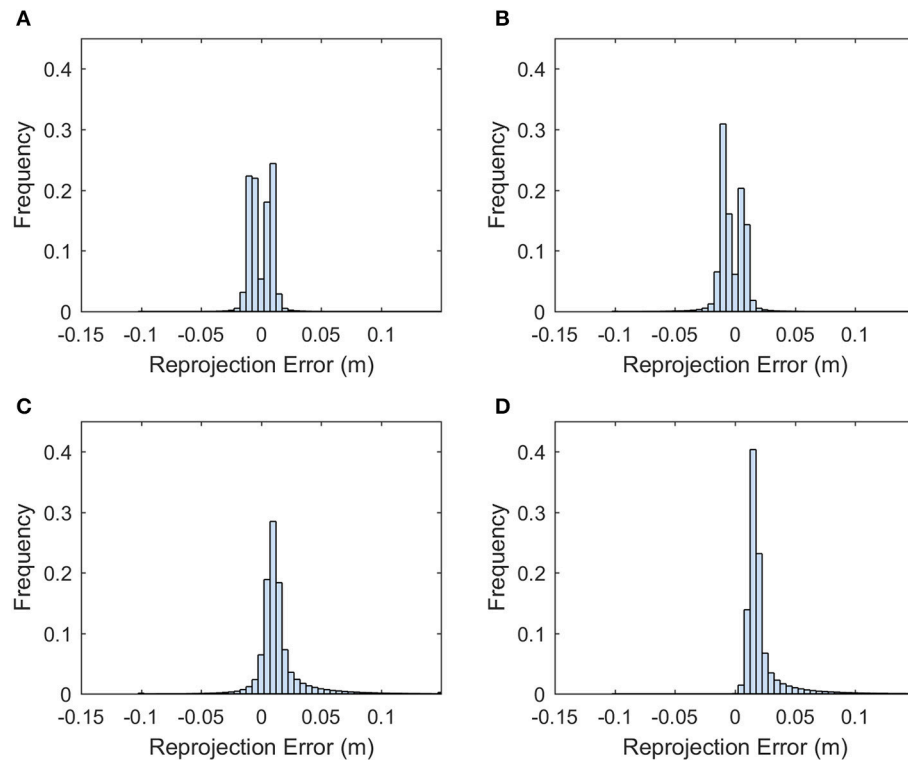


FIGURE 5 | Histogram of the projection error for the tie points generated from the 8/12/2016 dataset. **(A)** Error histogram for easting. **(B)** Error histogram for northing. **(C)** Error histogram for elevation. **(D)** Error histogram for the summation of the three axes.

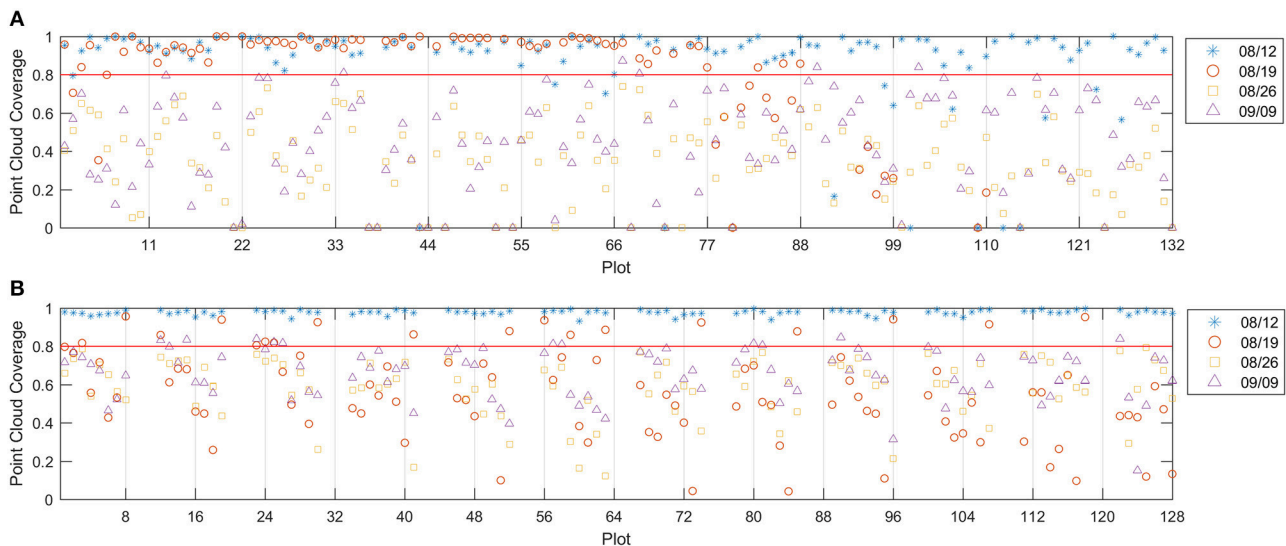
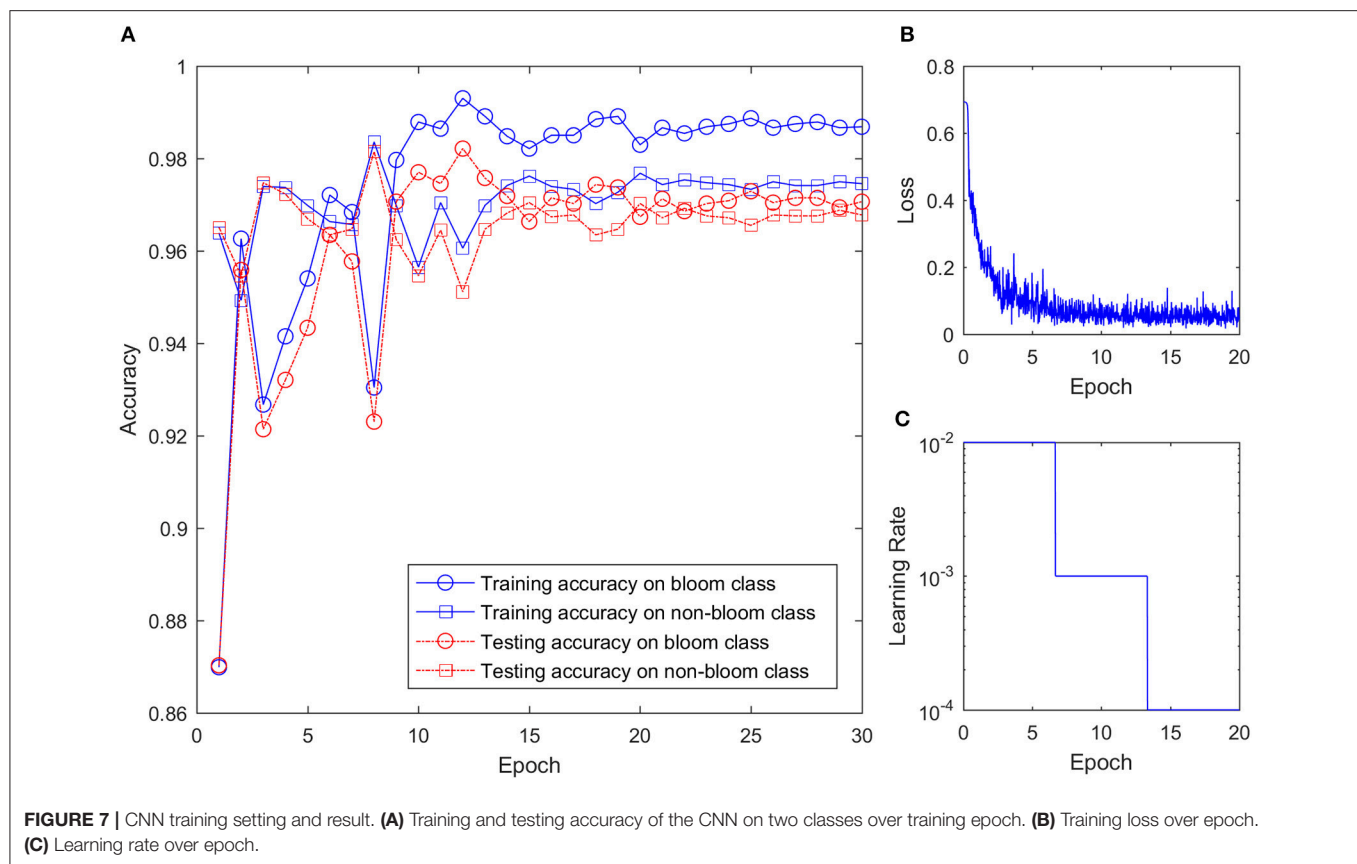


FIGURE 6 | Point cloud coverage for field 1 **(A)** and field 2 **(B)**. The red line indicates the 0.8 threshold.

accuracy for the bloom class reached a maximum on the 12th epoch. The accuracy for both classes decreased after reaching the maximum value, fluctuated a bit, but then reached a stable accuracy with small variation after 20 epochs. The training loss decreased quickly over the first few epochs, which indicates

the quick convergence of the training. After 10 epochs, where the learning rate changed from 0.01 to 0.001, the training loss kept decreasing but the change rate was small. After 20 epochs, the training loss reached the minimum with a certain level of oscillation, which indicates that the training process should



stop at the 20th epoch because no further improvement of the CNN can be gained from the last 10 epochs. Therefore, the training result at the 20th epoch was used to classify the potential bloom images. The testing accuracy was smaller than the training accuracy for both classes after 7 epochs but the difference between them was less than 0.01, which showed the CNN worked well on both the training and testing sets.

The classification result for each dataset using the trained CNN varied due to the different lighting conditions and flowering stages (Table 2). The classification result showed high precision (>0.9) for both classes across all datasets. However, the recall (true positive rate) for the bloom class was low for the 8/19 and 9/9 datasets. The CNN can cause overestimation or underestimation when comparing the number of predicted blooms and the number of actual blooms. For example, in the 8/12 dataset, the number of predicted bloom images was 789, but the actual number of bloom images was 719, thus the classification result overestimated the number of blooms by 10%. Similarly, the classification result overestimated the number of blooms by 23 and 29% in the 8/19 and 9/9 datasets, and underestimated the number of blooms by 2% in the 8/26 dataset. Objects such as cotton bolls, specular highlights on leaves, and pink flowers were misclassified as blooms because their shape and color appeared like a bloom in the aerial image due to the limited resolution (Figure 8). The misclassified blooms caused by objects (e.g., label sticks) on the ground could be removed based on the height from the ground, but the misclassified blooms caused

by the plants (e.g., leaves, pink flowers, and cotton bolls) were difficult to eliminate using the height. Small blooms or partly hidden blooms can easily be misclassified as non-blooms because of their size. Blooms in the shade can also be misclassified as non-blooms because their intensity is reduced. Including those misclassified images into the training set may further improve the CNN performance.

Efficiency and Accuracy of the Bloom Registration Algorithm

To cluster the 1,250 data points, the bloom registration algorithm used 13.3 to 48.4 s, which is 0.0106 to 0.0387 s per data point (Figure 9A). Using a larger threshold and larger noise level reduced the runtime because they resulted in a smaller number of clusters and thus the number of distance calculations. The bloom registration algorithm produced a near-zero misclustering rate by result at smaller noise level no matter what threshold was used because the distance between two reference clusters was still larger than the spread of the data points due to the noise (Figure 9B). As the noise level increased, the misclustering rate by result increased and larger threshold values had larger misclustering rates. This was mainly because of over-clustering when the noise of the points became large enough that certain portions of two clusters overlapped with each other and the algorithm could cluster them into one cluster. The misclustering rate by reference did not change significantly as the noise level

TABLE 2 | Classification result of the potential bloom images extracted from individual dataset for field 1.

		8/12/2016				8/19/2016				8/26/2016				9/9/2016			
		Bloom	Non-bloom	Precision	Recall	Bloom	Non-bloom	Precision	Recall	Bloom	Non-bloom	Precision	Recall	Bloom	Non-bloom	Precision	Recall
Actual class	Bloom	709	10	0.90	0.90	813	37	0.96	0.78	731	82	0.90	0.91	558	28	0.95	0.74
	Non-bloom	80	4876	0.98	0.997	234	23208	0.99	0.998	68	8543	0.99	0.99	191	3946	0.95	0.99

increased (Figure 9). Smaller thresholds had larger misclustering rates by reference.

At low noise level, the under-clustering takes the main effect because the bloom registration algorithm can split one cluster into smaller clusters at low noise level, generating a larger number of clusters compared to the real cluster number (Figure 9D). As the noise increased, a smaller number of clusters than the real number of clusters was generated, resulting in over-clustering. The bloom registration algorithm had a smaller cluster number with larger threshold at the same noise level, indicating that it is more prone to under-clustering but less prone to over-clustering.

Bloom Count Result

This section shows the bloom count results after removing the plots with point cloud coverage less than 0.8. The image count and manual count had the same trend for both fields (Figure 10). The error of the image count was between -4 and 3 for field 1, and the between -10 and 5 for field 2, showing that field 2 had more underestimated plots than field 1 (Figure 11). This may be due to the single plant per plot layout in field 1. The single plant per plot layout allows the plant to be seen from all directions without being blocked by other plants in the plot. Therefore, blooms in field 1 were more likely to be captured in the aerial images.

The overestimation was caused for two major reasons. One reason is the classification error of the CNN. For example, some leaves with specular highlights were classified as blooms (Figure 12A). The other reason is the quality of the point cloud, which causes the error of the bloom's 3D position and thus make the bloom registration incorrect (the same bloom from different images was registered as different blooms) (Figure 12B). The underestimation was caused by hidden blooms that were not shown in the aerial images, or blooms that were shown in the images but were not classified correctly by the CNN.

With enough datasets, it is possible to monitor the development of the flowers over time, which is one of the advantages of the proposed method. Figure 13 demonstrates the trend of the bloom development. It shows that the number of blooms was low at the early flowering stage, reached the peak in the middle flowering stage, and then decreased at the late flowering stage. This trend is also consistent with the trend from manual counting (Figures S1-S5).

DISCUSSION

This paper demonstrated a high-throughput methodology to count cotton blooms using aerial color images. Continuously monitoring the cotton blooms over time can provide information about the cotton growth status, such as the flowering time and peak flowering time, which can be used for production management and yield estimation. It is also helpful for breeding programs to identify short season or long season genotypes.

Unlike sorghum or corn, whose flowers open at top of the stem, cotton flowers open from the bottom and progress up the plant. At the early cotton flowering stage, blooms at the bottom can be covered by the leaves and may not be able to be captured by the aerial images. Therefore, it is expected that image counting

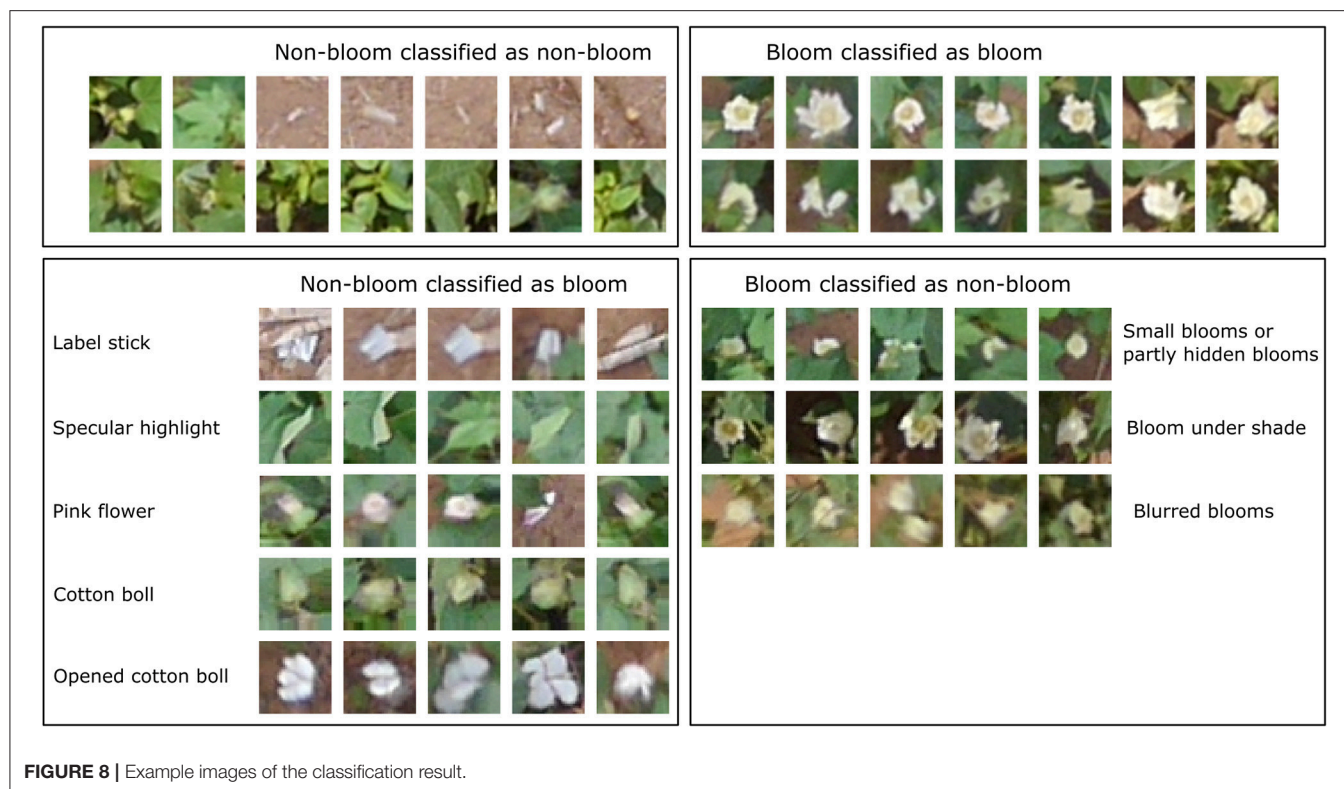


FIGURE 8 | Example images of the classification result.

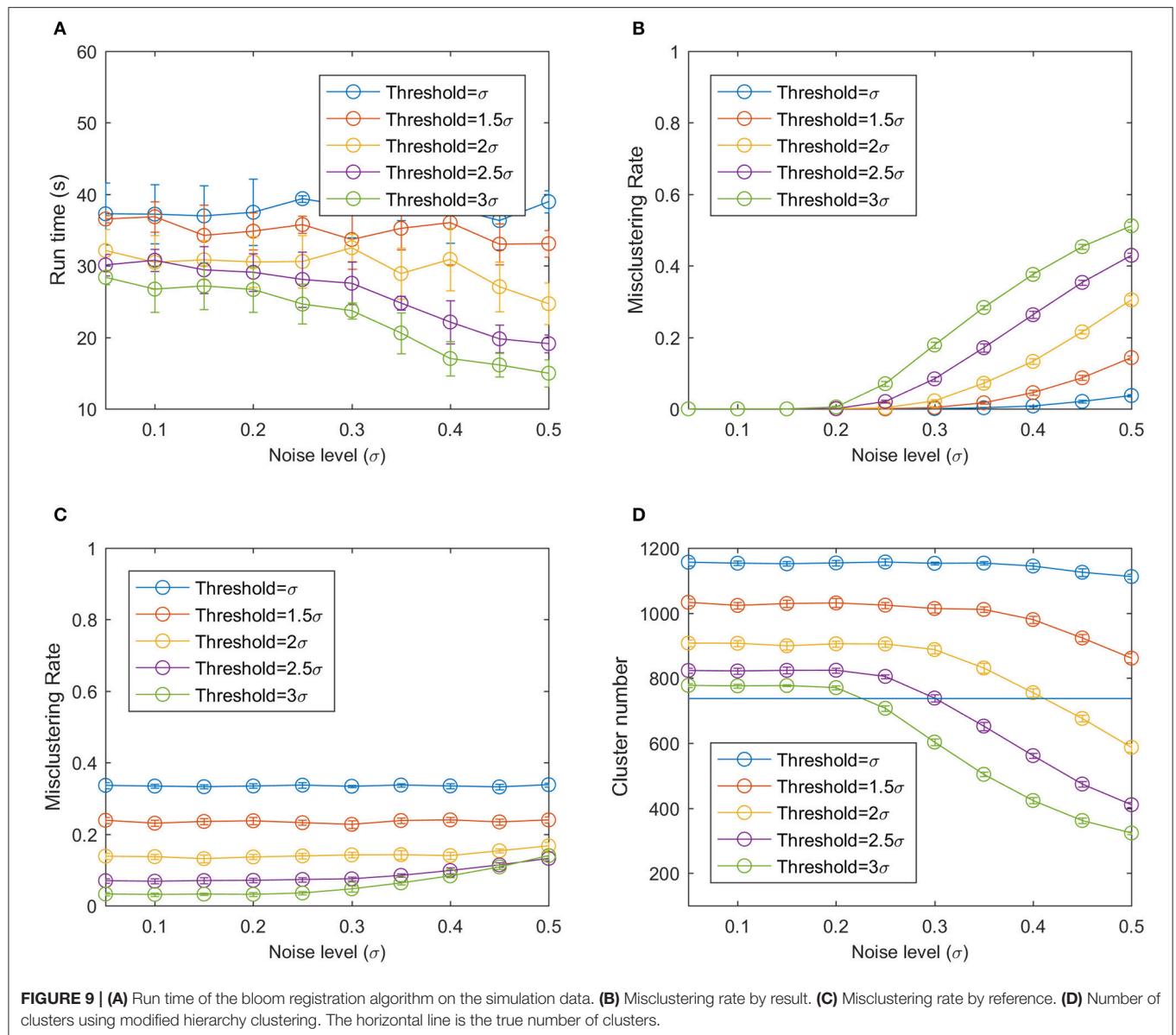
can underestimate the real bloom count. Instead of using the ortho-images which only take the top view of the canopy, we utilized all the raw images that take different views of a plot to get the bloom count in order to improve the underestimation to a certain extent. The inability to detect hidden blooms inside the canopy from aerial images is the major limitation of the proposed methodology. This issue could be addressed by using oblique aerial images or ground side-view images. The underestimation could be improved at the middle and later flowering stages since the flowers that open at the middle and top of the canopy are more likely to be imaged in the aerial images.

Besides the inability to detect hidden flowers, the proposed bloom counting method can generate errors from two aspects: the bloom detection error and the bloom registration error. The bloom detection error was affected by the image quality, the threshold to select the potential bloom images, and the accuracy of the CNN. The image quality was affected by the illumination condition at the time of data collection. Ideally, the best image quality can be obtained when the illumination is uniform from all directions so there are no shadows or specular highlights. The ideal condition is hard to obtain in the field, but the approximate ideal condition can be obtained on a cloudy day. When the sunlight directly shines on the canopy without being scattered by clouds, shadows and specular highlights can be found in the aerial images, which cause non-uniformed intensity changes of the plot. Those changes will affect the selection of the potential bloom positions.

The threshold to select the potential bloom positions was arbitrary and mainly based on the images. A high threshold can

eliminate some blooms that have low intensity due to shadows. A low threshold can include more non-bloom objects (such as the specular highlights), which can increase the processing time and be misclassified by the CNN. The classification of the potential bloom images relied on the CNN but misclassifying a bloom image as a non-bloom image affected the result differently from misclassifying a non-bloom image. When a bloom was classified as a non-bloom in one image, it was possible to be correctly classified as a bloom in other images given a different perspective of the flower; therefore, this flower could be included in the final result. However, when a non-bloom image was classified as a bloom, this false bloom was counted and there was no approach to remove it in the current method. The training samples for the CNN were selected from images collected on only 4 different days, so the trained CNN may not be suitable for data collected in different growth stages, especially when the color of the cotton leaves changes over the season. Therefore, including more training samples from different dates over the growth season could improve the accuracy and robustness of the CNN. The different appearances of the cotton flower caused by genotype differences—for instance, the Pima cotton has yellow bloom in contrast with white bloom for some upland cotton varieties—should also be considered when constructing the training samples.

The accuracy of the bloom registration was affected by the 3D position of the bloom and the registration algorithm. The 3D position of the bloom was affected by the pixel location and the dense point cloud. Therefore, the accuracy of the dense point cloud has an important impact on the accuracy of a bloom's 3D



position, and thus affects the bloom registration results. A larger error of the 3D position can cause the bloom registration to under-cluster or over-cluster the blooms. Under-clustering can count the same bloom more than once and over-clustering may count several blooms as one bloom, making the bloom count unreliable. The completeness of the point cloud also affects the registration result. If the dense point cloud cannot cover the whole canopy, some blooms that are detected in the images may not have a valid projection on the dense point cloud. Those blooms will not be registered and underestimation of the bloom count will occur. Therefore, adequate image overlap is critical in data collection to capture the whole canopy. Increasing the image overlap can improve the completeness, but also requires more data collection and processing time. Oblique imagery can also improve the completeness of the point cloud by providing more views of a plot and potentially can image more occluded flowers.

Although the proposed bloom counting methodology usually provides an underestimated bloom count compared to manual counting, it has the advantage in throughput over manual counting. It saves manual labor and makes continuously monitoring the flowering possible. Without such throughput, it is impossible to continuously monitor the flowering stage and determine the flowering time, peak flowering time, and seasonal bloom count. For farmers and breeders, it is helpful to estimate the fiber yield because some studies have shown the fiber yield is correlated with seasonal bloom count (Heitholt, 1993, 1995). However, additional studies on how well the proposed methodology can estimate fiber yield are needed. The methodology is also helpful for differentiating the growth behavior among different genotypes, which can be used to select certain genotypes, such as short-season or long-season genotypes. Compared to other flower detection methods that are

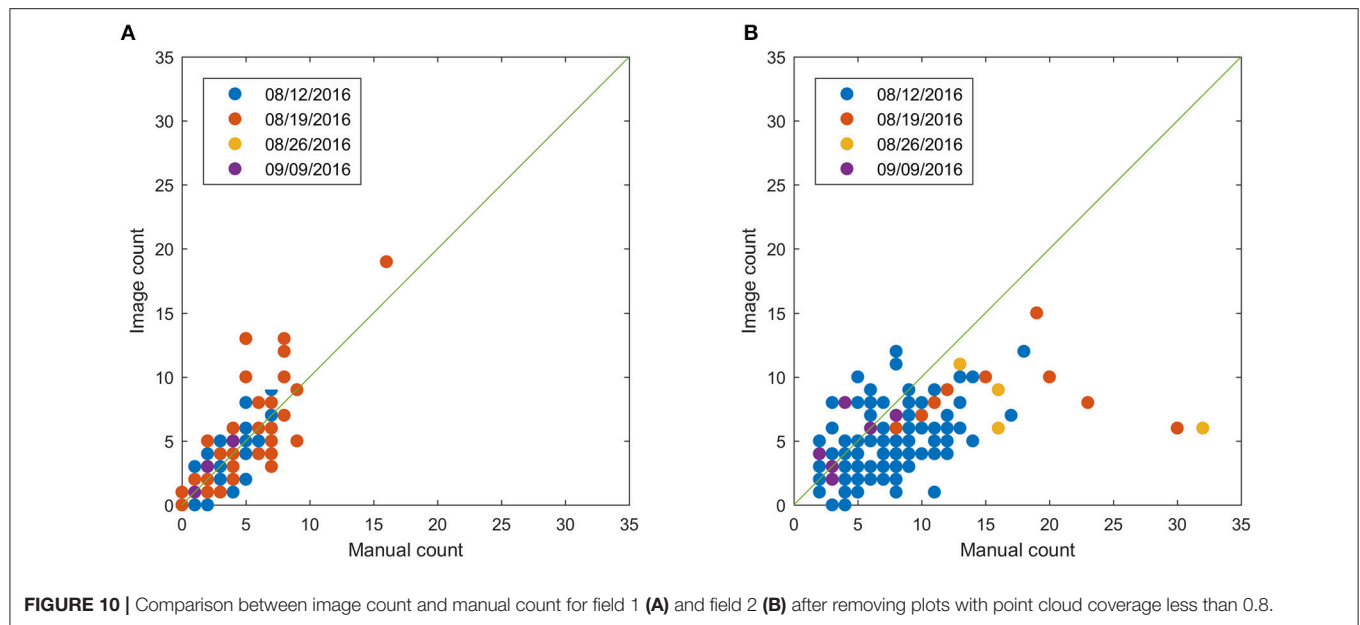


FIGURE 10 | Comparison between image count and manual count for field 1 (A) and field 2 (B) after removing plots with point cloud coverage less than 0.8.

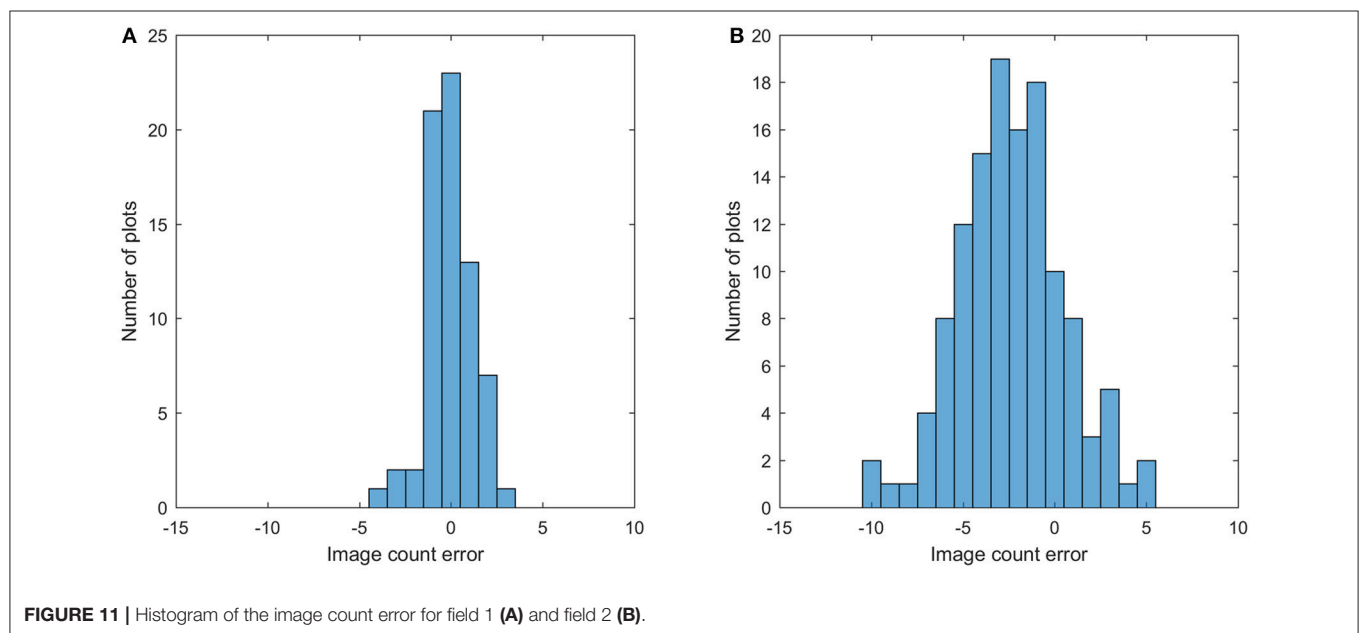


FIGURE 11 | Histogram of the image count error for field 1 (A) and field 2 (B).

based on the percentage of flower pixels, the method proposed in this study can directly provide flower count without exploring the correlation between pixel percentage and flower count (Adamsen et al., 2000). The proposed method also can provide the locations of flowers as byproducts, which could be used to correlate with the cotton bolls (Figure 14).

To implement the proposed method, farmers and breeders can use commercial aerial photogrammetry systems or build custom systems to collect aerial images. The data processing pipeline can be used as long as the image quality (such as the ground resolution) meets the requirements. This study used small fields of 920 m² (0.22 acre) and the data collection throughput is not

enough for commercial farms or breeding programs with large fields. Although the throughput can be increased by increasing the flight altitude, the reduced ground resolution may not meet the requirement for the pipeline to correctly recognize cotton flowers. Alternative solution is to use high-resolution cameras to maintain the ground resolution when imaging at higher altitudes.

CONCLUSION

This study developed a high throughput methodology for cotton bloom detection using aerial images, which can be potentially

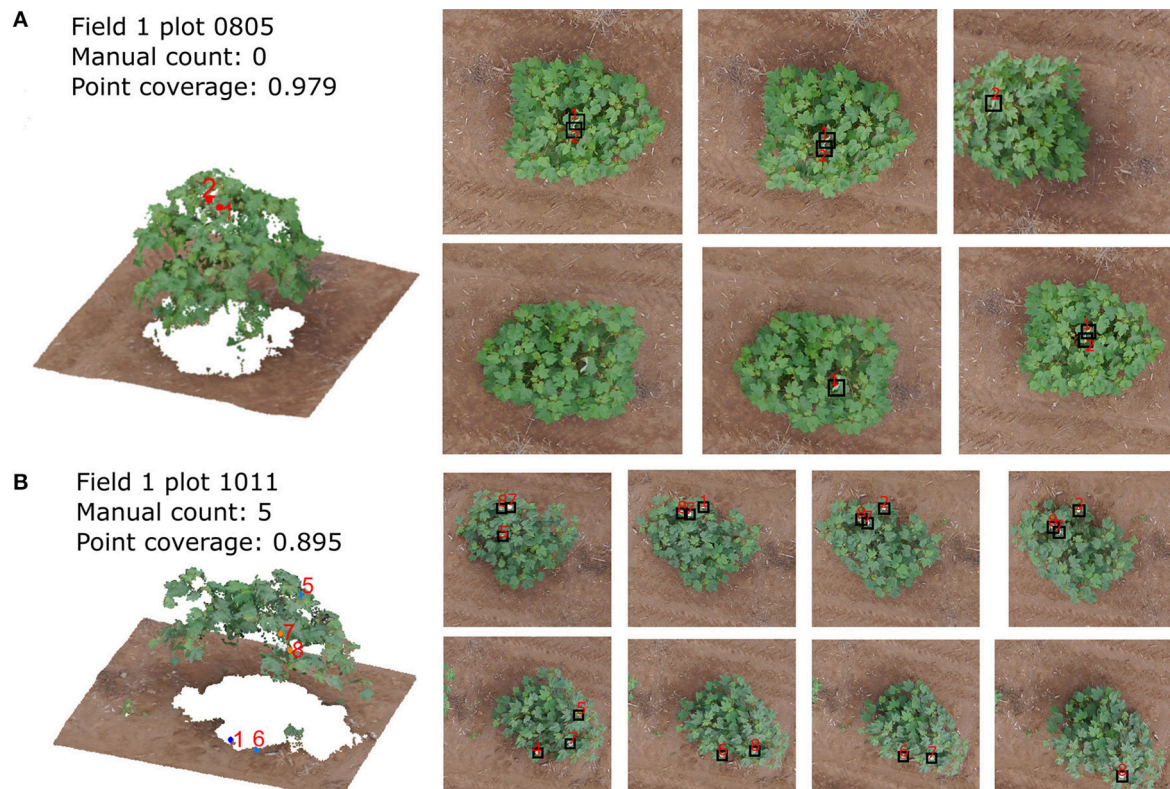


FIGURE 12 | The bloom detection results for plot 0110 (A) and plot 1011 (B) in field 1 on 8/12/2016 dataset. The left images show the point cloud and detected blooms and right images show the corresponding blooms in the raw images.

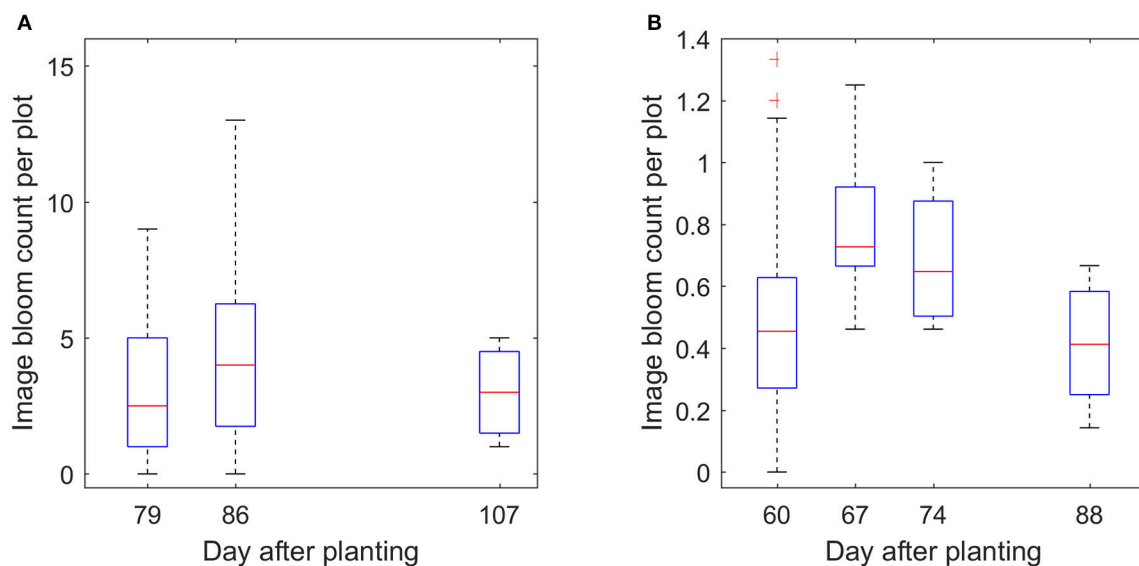


FIGURE 13 | Boxplot of the image count per plot over time for field 1 (A) and field 2 (B).

used to monitor cotton flowering over the season for cotton production management and yield estimation. The method generally underestimated the bloom count due to the inability to

count hidden flowers, but the bloom count for the single plant layout was less likely to be underestimated than the multiple plant layout. The accuracy and completeness of the dense point

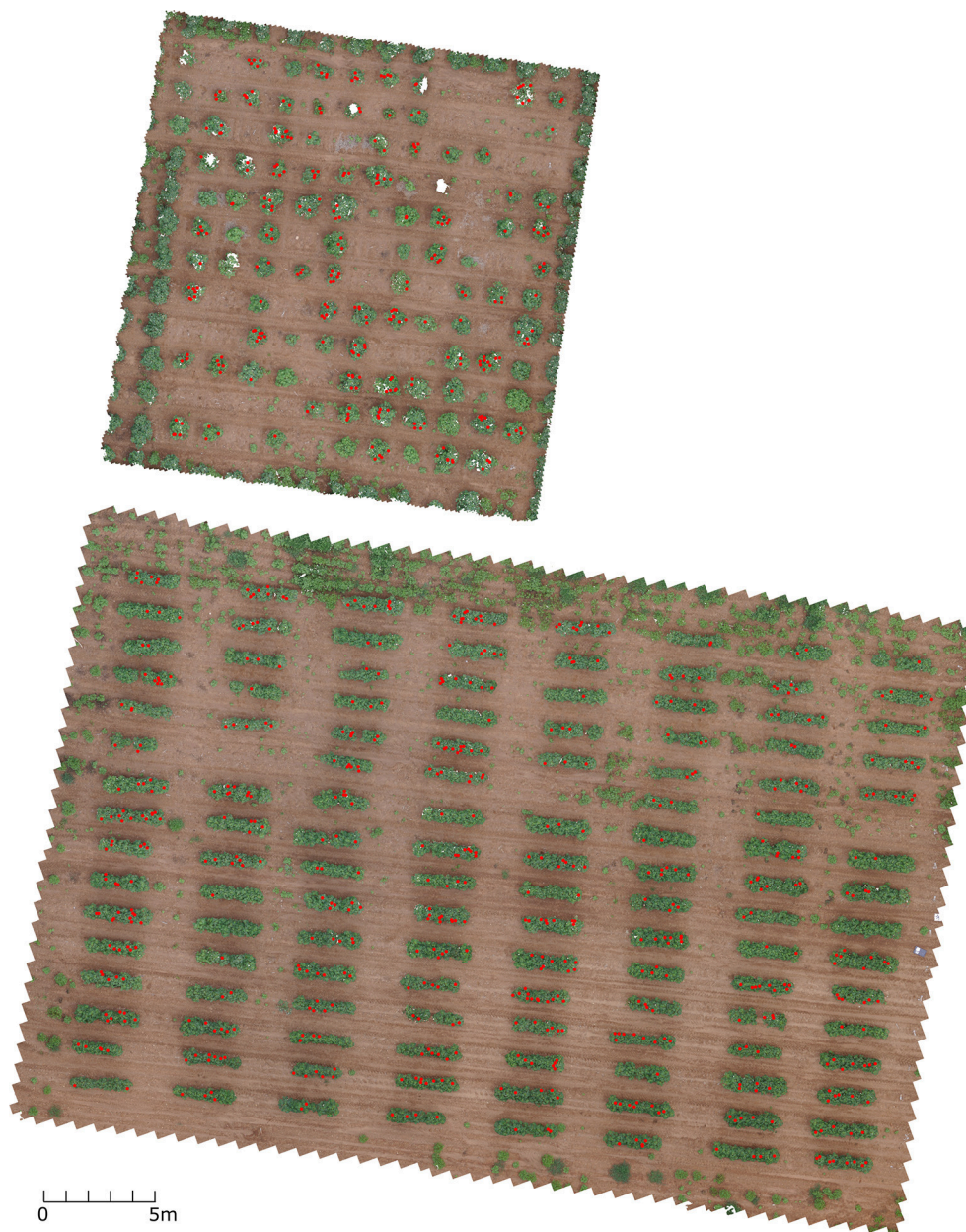


FIGURE 14 | Top view of the two test fields with red dots indicating the flower locations using 8/12/2016 dataset. The image was rendered from dense point cloud.

cloud has an impact on the bloom count result, so generating a good dense point cloud can improve the results significantly. The bloom registration algorithm developed in this study was efficient in terms of runtime and was more prone to under-clustering but less prone to over-clustering. The trained CNN correctly classified more than 97% of the training and test images, and more than 90% of the potential flowers extracted from individual datasets. Since the false classification from the CNN can result in false bloom count, designing a robust CNN that can handle images taken under different field illumination conditions and cotton growth stages will be included in future studies. In

addition, oblique imagery will be explored to improve the quality of the dense point cloud.

AUTHOR CONTRIBUTIONS

RX: performed the data analysis and wrote the main manuscript text; CL and AP: designed the experiments and wrote the manuscript; YJ: designed plot layout and contributed to the data collection; SS and JR: contributed to the data collection and field management; All of the authors have read and approved the final manuscript.

FUNDING

This study was funded jointly by the Agricultural Sensing and Robotics Initiative of the College of Engineering, and the College of Agricultural and Environmental Sciences of the University of Georgia. The project was also partially supported by the National Robotics Initiative grant (NIFA grant No. 2017-67021-25928) and Georgia Cotton Commission.

ACKNOWLEDGMENTS

The authors gratefully thank Mr. Jeevan Adhikari, Dr. Tariq Shehzad, Mr. Zikai Wei, Mr. Jamal Hunter

and Mr. Shuxiang Fan for their great assistance in planting, data collection and harvesting. The authors also want to thank Mr. Joshua Griffin for his great assistance in field management. The authors thank Dr. Peng Wah Chee for selecting cotton lines and providing cottonseeds.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.02235/full#supplementary-material>

REFERENCES

- Adamsen, F., Coffelt, T., Nelson, J. M., Barnes, E. M., and Rice, R. C. (2000). Method for using images from a color digital camera to estimate flower number. *Crop Sci.* 40, 704–709. doi: 10.2135/cropsci2000.403704x
- Biradar, B. V., and Shrikhande, S. P. (2015). Flower detection and counting using morphological and segmentation technique. *Int. J. Comput. Sci. Inform. Technol.* 6, 2498–2501.
- Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., et al. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. *Remote Sens* 8:416. doi: 10.3390/rs8050416
- Heitholt, J. J. (1993). Cotton boll retention and its relationship to lint yield. *Crop Sci.* 33, 486–490. doi: 10.2135/cropsci1993.0011183X003300030014x
- Heitholt, J. J. (1995). Cotton flowering and boll retention in different planting configurations and leaf shapes. *Agron. J.* 87, 994–998. doi: 10.2134/agronj1995.00021962008700050037x
- Horton, R., Cano, E., Bulanon, D., and Fallahi, E. (2017). Peach flower monitoring using aerial multispectral imaging. *J. Imaging* 3:2. doi: 10.3390/jimaging3010002
- Hsu, T.-H., Lee, C.-H., and Chen, L.-H. (2011). An interactive flower image recognition system. *Multimed. Tools Appl.* 53, 53–73. doi: 10.1007/s1042-010-0490-6
- Liu, Y., Tang, F., Zhou, D., Meng, Y., and Dong, W. (2016). “Flower classification via convolutional neural network” in *International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA)* (Qingdao).
- Pettigrew, W. T. (1994). Source-to-sink manipulation effects on cotton lint yield and yield components. *Agron. J.* 86, 731–735. doi: 10.2134/agronj1994.00021962008600040027x
- Seeland, M., Rzanny, M., Alaqraa, N., Thuille, A., Boho, D., Wäldchen, J., et al. (2016). “Description of flower colors for image based plant species classification,” in *Paper Presented at the Proceedings of the 22nd German Color Workshop (FWS)* (Ilmenau: Zentrum für Bild-und Signalverarbeitung eV).
- Siraj, F., Salahuddin, M. A., and Yusof, S. A. M. (2010). “Digital image classification for Malaysian blooming flower,” in *2010 Second International Conference on Computational Intelligence, Modelling and Simulation* (Bali).
- Thorp, K., Wang, G., Badaruddin, M., and Bronson, K. (2016). Lesquerella seed yield estimation using color image segmentation to track flowering dynamics in response to variable water and nitrogen management. *Ind. Crops Prod.* 86, 186–195. doi: 10.1016/j.indcrop.2016.03.035
- Torr, P. H., and Zisserman, A. (2000). MLESAC: a new robust estimator with application to estimating image geometry. *Comput. Vis. Image Understand.* 78, 138–156. doi: 10.1006/cviu.1999.0832
- Wells, R., and Meredith, W. R. (1984). Comparative growth of obsolete and modern cotton cultivars. III. Relationship of yield to observed growth characteristics. *Crop Sci.* 24, 868–872.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Xu, Li, Paterson, Jiang, Sun and Robertson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Fast High Resolution Volume Carving for 3D Plant Shoot Reconstruction

Hanno Scharr*, Christoph Brieze, Patrick Embgenbroich, Andreas Fischbach, Fabio Fiorani and Mark Müller-Linow

Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich GmbH, Jülich, Germany

OPEN ACCESS

Edited by:

Alexander Bucksch,
University of Georgia, United States

Reviewed by:

Suxing Liu,
University of Georgia, United States
Dimitrios Fanourakis,
Institute of Olive Tree, Subtropical
Plants and Viticulture, Hellenic
Agricultural Organization "Demeter"
(NAGREF), Greece

*Correspondence:

Hanno Scharr
h.scharr@fz-juelich.de

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 18 August 2017

Accepted: 12 September 2017

Published: 28 September 2017

Citation:

Scharr H, Brieze C, Embgenbroich P,
Fischbach A, Fiorani F and
Müller-Linow M (2017) Fast High
Resolution Volume Carving for 3D
Plant Shoot Reconstruction.
Front. Plant Sci. 8:1680.
doi: 10.3389/fpls.2017.01680

Volume carving is a well established method for visual hull reconstruction and has been successfully applied in plant phenotyping, especially for 3d reconstruction of small plants and seeds. When imaging larger plants at still relatively high spatial resolution (≤ 1 mm), well known implementations become slow or have prohibitively large memory needs. Here we present and evaluate a computationally efficient algorithm for volume carving, allowing e.g., 3D reconstruction of plant shoots. It combines a well-known multi-grid representation called "Octree" with an efficient image region integration scheme called "Integral image." Speedup with respect to less efficient octree implementations is about 2 orders of magnitude, due to the introduced refinement strategy "Mark and refine." Speedup is about a factor 1.6 compared to a highly optimized GPU implementation using equidistant voxel grids, even without using any parallelization. We demonstrate the application of this method for trait derivation of banana and maize plants.

Keywords: image processing, 3D from silhouettes, visual hull, octree, integral image, refinement strategy, performance analysis

1. INTRODUCTION

Complementary to genomics, the quantitative description of plant phenotypes is at the core of basic research for the analysis of plant development and physiological responses to abiotic and biotic challenges as well as for applications in plant genetic improvement and precision agriculture. An increasing amount of phenotypic data are generated using digital images and time series experiments using a variety of methods and sensors both in controlled environment and in the field (reviewed e.g., in Furbank and Tester, 2011; Fiorani and Schurr, 2013; Mulla, 2013; Araus and Cairns, 2014). Most of these methodologies, ranging from RGB to spectral imaging, are based in high-throughput phenotyping pipelines primarily on 2D spatial analyses for the estimation from image analysis of plant traits such as total leaf area, crop coverage and leaf biomass (Homolova et al., 2013; Chen et al., 2014), leaf color (Hu et al., 2013), plant height measurements using light curtains (Fanourakis et al., 2014) or root morphological features (Das et al., 2015). The use of optical imaging with RGB cameras for estimation of shoot area and above-ground biomass is an established method deployed in indoor phenotyping platforms imaging stations for large-scale studies (Al-Tamimi et al., 2016). These methodologies are based on estimation of leaf area and leaf biomass by acquisition and image analysis of a series of RGB projections from multiple view-angles (Golzarian et al., 2011). Issues that may arise from this approach that generally use a limited number of view-angles are related to the likely underestimation of total leaf area for relatively complex shoot architecture and for large plants at advanced developmental stages. In many cases and for the same

reasons, quantifying the growth rates and leaf angles of individual leaves from global shoot images remains challenging. Optical 3D reconstruction of plant shoots using a variety of methods assists in alleviating these issues. For example, studies were conducted in *Arabidopsis* using light-field cameras for depth reconstruction of rosettes (Apelt et al., 2015), in tobacco using high-resolution 3D imaging and a mesh approach (Paprocki et al., 2012), and also in the field using stereo imaging with consumer cameras (Müller-Linow et al., 2015). Progress in this field is still limited by the required computational power and time investment for image analysis (Minervini et al., 2015). In this respect, improvements are required both for methods using 3D reconstruction from silhouettes and 3D imaging.

1.1. Related Work

Measuring plant geometry from single view-point 2D images often suffers from insufficient information, especially when plant organs occlude each other (self-occlusion). In order to achieve more detailed information and recover the plants 3D geometric structure volume carving is a well established method to generate 3D point clouds of plant shoots (Koenderink et al., 2009; Golbach et al., 2015; Klodt and Cremers, 2015), seeds (Roussel et al., 2015, 2016; Jahnke et al., 2016), and roots (Clark et al., 2011; Zheng et al., 2011; Topp et al., 2013). Volume carving can be applied in high-throughput scenarios (Golbach et al., 2015): For the reconstruction of relatively simple plant structures like tomato seedlings image reconstruction takes ~25–60 ms, based on a well thought out camera geometry using 10 cameras and a suitably low voxel resolution $240 \times 240 \times 300$ voxels at 0.25 mm voxel width. Short reconstruction times are achieved by precomputing voxel to pixel projections for each of the fully calibrated cameras. However, precomputing lookup-tables is not feasible for high voxel resolutions due to storage restrictions (Ladikos et al., 2008). Current implementations popular in plant sciences suffer from high computational complexity, when voxel resolutions are high. We therefore implemented and tested a fast and reliable volume carving algorithm based on octrees (cmp. Klodt and Cremers, 2015) and integral images (cmp. Veksler, 2003), and investigate different refinement strategies. This work summarizes and extends our findings presented in Embgenbroich (2015).

Visual hull reconstruction via volume carving is a well-known shape-from-silhouette technique (Martin and Aggarwal, 1983; Potmesil, 1987; Laurentini, 1994) and found many applications. Also octree as multigrid approach and integral image for reliable and fast foreground testing have been used successfully with volume carving in medical applications (Ladikos et al., 2008) and human pose reconstruction (Kanauija et al., 2013). Realtime applications at 512^3 voxel resolution have been achieved where suitable caching strategies on GPUs can be applied e.g., for video conferencing (Waizenegger et al., 2009). Here we demonstrate that even higher spatial resolutions are achievable on consumer computer hardware without prohibitively large computational cost. Subsequent octree-voxel-based processing allows extraction of plant structural features suitable for plant phenotypic trait extraction.

2. RECONSTRUCTING SHAPES FROM SILHOUETTES

2.1. Voxelbased Volume Carving

Here, we revisit voxel-based volume carving, closely following the description found in Roussel et al. (2016). In the subsequent sections we will then extend this formulation using octrees as multigrid extension.

Consider an imaging setup with a set of fully calibrated cameras, and a turn-table allowing to rotate our plant of interest around its vertical axis. Golbach et al. (2015) use 10 cameras and no turn-table, Roussel et al. (2016) a single camera and a robot instead of a turn-table; here we use different setups combining up to three cameras and a turntable (see section 3.1).

For each camera c we obtain the 3×3 intrinsic camera matrix \mathbf{K}_c , 3×3 rotation matrix \mathbf{R}_c and translation vector \vec{t}_c with respect to the reference camera (cmp. Hartley and Zisserman, 2004), and the distance between the origin of our working volume and the reference camera center from calibration (cmp. section 3.2). The origin of the working volume is selected to be the rotation center of the turn-table (cmp. section 3.1).

We acquire N images, showing the plant of interest under (equidistantly spaced) rotation angles α_i where $i \in \{1, \dots, N\}$. We segment each image into a binary mask \mathbf{M}_i being one at the foreground, i.e., plant, and zero at background locations. For segmentation, we either use HSV color thresholding (Walter et al., 2007) or a support vector machine (SVM) based learning algorithm (cmp. e.g., Hearst et al., 1998; Wang et al., 2011; Li et al., 2013). Both methods are parameterized offline, where SVM parameters are learned using suitable training data. HSV parameters are hand-tuned by setting 6 threshold values on the 3 color channels. Computational effort as a preprocessing step for carving is negligible, as it typically takes fractions of a second per image for both methods. Subsequently, small objects like noise are removed and small holes filled. Suitable filling sizes depend on the imaging setup and need to be determined empirically.

For each image and thus segmentation mask we calculate the homogeneous camera projection matrix \mathbf{P}_i , from the rotation angle α_i by

$$\mathbf{P}_i = \mathbf{K}_c(\mathbf{R}_c|\vec{t}_c + \vec{t}_0) \begin{pmatrix} \mathbf{R}_i & 0 \\ 0 & 1 \end{pmatrix} \quad (1)$$

where \mathbf{R}_i is the rotation matrix corresponding to the given angle α_i , and translation vector \vec{t}_0 is calculated using the distance of the world origin to the reference camera center, also known from calibration (see e.g., Hartley and Zisserman, 2004). By this, the world coordinate frame rotates with the object, i.e., the plant.

We define an equidistantly spaced, cubic voxel grid around the world origin, being large enough to contain the plant. Such a working volume depends on the plant size in order to keep voxel number and thus complexity low. We can relax this requirement later, when using octrees.

Each voxel center with homogeneous world coordinates \vec{X} is projected to a point \vec{x}_i in each mask \mathbf{M}_i by

$$\vec{x}_i = \mathbf{P}_i\vec{X} \quad (2)$$

If \vec{X} is projected to the background region of at least one of the N masks \mathbf{M}_i , then this voxel does not belong to the foreground object and its value $V(\vec{X})$ is set to 0, i.e.,

$$V(\vec{X}) = \prod_{i=1}^N \mathbf{M}_i(\vec{x}_i) \quad (3)$$

Thus, if a voxel belongs to the foreground object, its value $V(\vec{X})$ is set to 1.

When high voxel resolution is desired, and thus runtimes increase, parallelization of the carving algorithm (Ladikos et al., 2008; Brenscheidt, 2014) is feasible. However, as complexity increases linearly with the voxel number and thus cubically with voxel resolution, equidistant voxel discretization quickly reaches its limits for any computer hardware. High resolutions become available on current desktop computer hardware, when hierarchically representing the voxel grid, e.g., as an octree as described next (Szeliski, 1993; Ladikos et al., 2008; Klodt and Cremers, 2015).

2.2. Octrees

An *Octree* (Meagher, 1980, 1995; Szeliski, 1993) is a hierarchical tree data-structure. Each node corresponds to a cube in 3d space (i.e., a large or small “voxel”) where each node not being an end-node or “leaf” is subdivided in eight child nodes. The child nodes have each half the size, i.e., edge length, of their parent node and are non-overlapping such that they fill the same volume as their parent. (cmp. **Figure 1**). Octrees are designed to efficiently store voxel grids, where large, unstructured volumes can be stored using only few nodes (i.e., large voxels), whereas small structures or surfaces in space can be finely resolved using nodes further up in the tree (i.e., small voxels), i.e., higher “level” nodes. To this end, when building up a representation of an object, an octree node is kept as an end-node, when the corresponding 3d volume is either empty, or completely filled by the object. When the corresponding volume is only partly filled by the object, the node is split into its eight children. This is iterated until the finest resolution is reached. Then leaf nodes at finest resolution can be considered as partly filled and represent surface voxels of the object.

Consequently, memory investment for storing a working volume only marginally depends on its size, but rather on the size and complexity of the object of interest. For example storing an equidistantly sampled working volume of 1 m^3 using voxels of 0.25 mm edge length requires $6.4 \cdot 10^{10}$ voxels or $\sim 64 \text{ GB}$ when representing each voxel by a single byte without any compression. Storing the same working volume takes only a few bytes, when it is empty, and when occupied by an object, still much less than storing an equidistantly sampled version of the occupied volume only (see section 4.1 for an example). We can therefore select our initial working volume such that it well contains the overall visible volume of our setup, without compromising on runtime.

2.3. Integral Image

The occupancy test in Equation (3) assumes, that a each voxel \vec{X} corresponds to a single pixel or point \vec{x}_i in a mask image \mathbf{M}_i . This can be sufficiently well fulfilled when image resolution is low,

such that pixel size at working distance is well above voxel size. For large voxels as in the lower levels of an octree, this assumption is not sufficiently fulfilled. We need to adapt the occupancy test, such that we can decide if a voxel is empty, completely or only partly filled. Consider a large voxel being projected to a mask image (see **Figure 2**, left). In general, the 2d central projection of a cube is a hexagon and we needed to test every pixel within the region of this hexagon. When foreground is 1 and background 0, this corresponds to integrating the mask image over the region of the hexagon and compare the result R_{hex} with the area of the hexagon A_{hex} . Then only when $R_{\text{hex}}/A_{\text{hex}}$ is exactly 1 or 0, the octree node would not be split. However, integrating over a hexagon for every voxel is time consuming. Instead of this, we integrate over the bounding-box of the hexagon, yielding R_{bbx} , as this can be done very efficiently using the integral image approach. A bounding-box is a rectangle with sides parallel to the coordinate axes, being spanned by \vec{x}_{\min} as the upper left corner and \vec{x}_{\max} as the lower right corner¹. Here $\vec{x}_{\min} = (x_{\min}, y_{\min})$ contains the minimum coordinate values² of all corner points of the voxel projected to the mask image, and \vec{x}_{\max} their maximum³, accordingly. In case the bounding-box is completely filled by the object, i.e., $R_{\text{bbx}} = A_{\text{bbx}}$ or completely empty, $R_{\text{bbx}} = 0$, it will be the same for the hexagon.

$$A_{\text{bbx}} = (x_{\max} - x_{\min}) \cdot (y_{\max} - y_{\min}) \quad (4)$$

is the area covered by the bounding box. For a partly filled bounding-box we will split the corresponding octree node, in order to be on the safe side, even though there may be cases when the bounding box is partly filled, but the hexagon actually is not.

The integral over the bounding box can be evaluated very efficiently using the summed area table (Crow, 1984) or *Integral Image* \mathbf{I}_i of the mask \mathbf{M}_i (see **Figure 2**, right, and Veksler, 2003; Viola and Jones, 2004). Every pixel of \mathbf{I}_i contains the integral of \mathbf{M}_i in the rectangle being spanned by $\vec{x}_{\min} = \vec{0}$ and \vec{x}_{\max} being the current pixel.

$$\mathbf{I}_i(x, y) = \sum_{x' \leq x, y' \leq y} \mathbf{M}_i(x', y') \quad (5)$$

\mathbf{I}_i can be derived efficiently by a recursive convolution scheme starting at the upper left corner and proceeding row- or column-wise, where we calculate the current pixel value from the value of its upper and left neighbors, i.e.,

$$\mathbf{I}_i(x, y) = \mathbf{M}_i(x, y) + \mathbf{I}_i(x-1, y) + \mathbf{I}_i(x, y-1) - \mathbf{I}_i(x-1, y-1) \quad (6)$$

with considering $\mathbf{I} = 0$ outside the mask image. From the integral image the sum over any rectangular region can be calculated using four values

$$R_{\text{bbx},i} = \mathbf{I}_i(x_{\min}, y_{\min}) - \mathbf{I}_i(x_{\min}, y_{\max}) - \mathbf{I}_i(x_{\max}, y_{\min}) + \mathbf{I}_i(x_{\max}, y_{\max}) \quad (7)$$

¹The image coordinate origin is assumed to be in the upper left corner and x - and y -axes pointing to the right and downwards, as usual for images.

²Rounded to lower values, using the C++ operation *floor*.

³Rounded to higher values, using the C++ operation *ceil*.

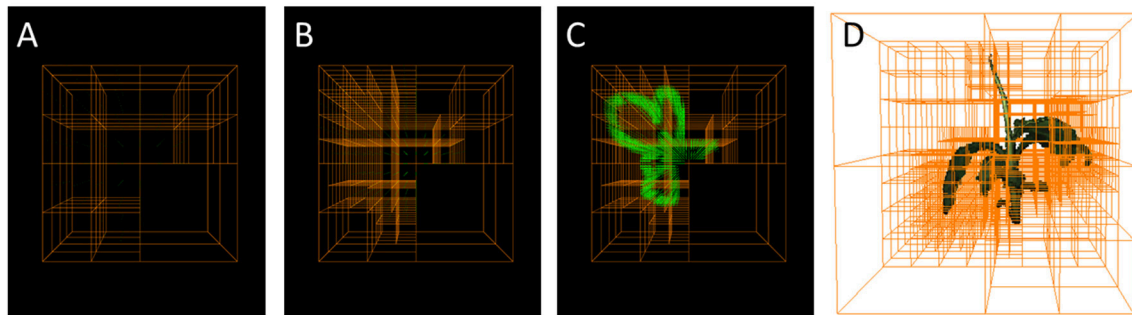


FIGURE 1 | Octrees: A well-established multigrid data structure reducing complexity wrt. an equidistantly spaced voxel grid. **(A–C)** Selected refinement steps in an Octree-based reconstruction. Foreground corner points are green. **(D)** Reconstructed plant with corresponding Octree.

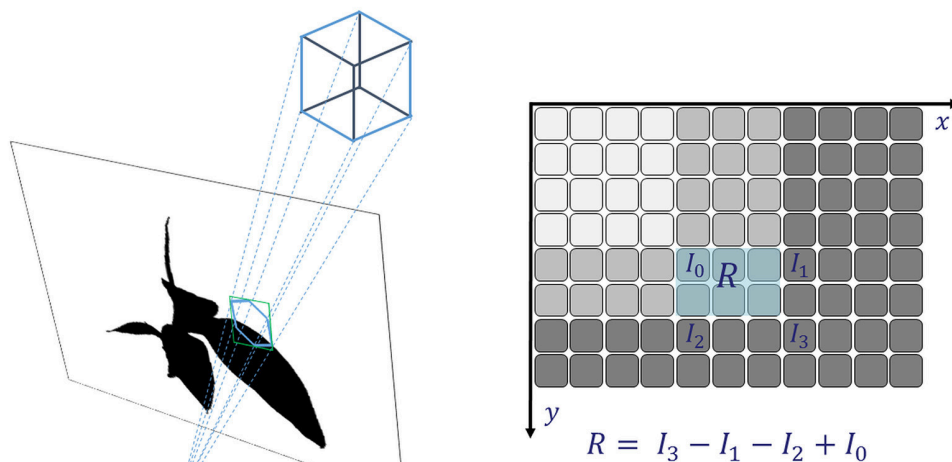


FIGURE 2 | Integral images: fast and reliable checking if a (potentially large) voxel covers fore- and background and thus needs to be refined.

For border handling, bounding-box corners outside the current field of view, are moved to the nearest pixel in I_i .

2.4. Volume Carving with Octrees

When using octrees instead of equidistantly spaced, cubic voxel grids, the occupancy test from Equation (3) needs to be exchanged. To do so, we first calculate for every mask image M_i the corresponding integral image I_i using the scheme from Equation (6).

We start with an octree containing only a single coarse voxel covering the whole working space, i.e., the trunk node of the octree. This node is initially marked as “object,” i.e., it is assumed it contains an object. Then, for all images, all nodes are iteratively updated, when a node N is marked as “object” and

$$R_{\text{bbx},i}/A_{\text{bbx},i} = \begin{cases} 1 & \text{then } N \text{ is marked as “object”} \\ 0 & \text{then } N \text{ is marked as “empty”} \\ \text{else } N \text{ is marked as “refine.”} \end{cases} \quad (8)$$

Voxels marked as “refine” are split, marked as “empty” and their Children are marked as “object.” This update scheme is iterated

until the desired finest resolution is reached. In the following, “one iteration” means projecting all current octree voxels on one image once. We have several options, how to organize the iterations:

1. Depth first: On the current image, we iterate until all nodes are split to the finest resolution, where we visit each image only once, or
2. Breadth first: We make a single pass over all nodes for the current image, split the ones marked as “refine,” and go to the next image for the next pass, where we visit each image multiple times.
3. Refining resolution: Set the maximum resolution to a coarse level and iterate breadth first to convergence. Then set the resolution such that finest voxels are allowed to be split once and run breadth first once. Repeat the last step until the desired resolution is reached.
4. Mark and refine: Make a single pass over all nodes for all images, only marking nodes, not actually splitting them. Then split all nodes marked as “refine.” Iterate until the desired resolution is reached.

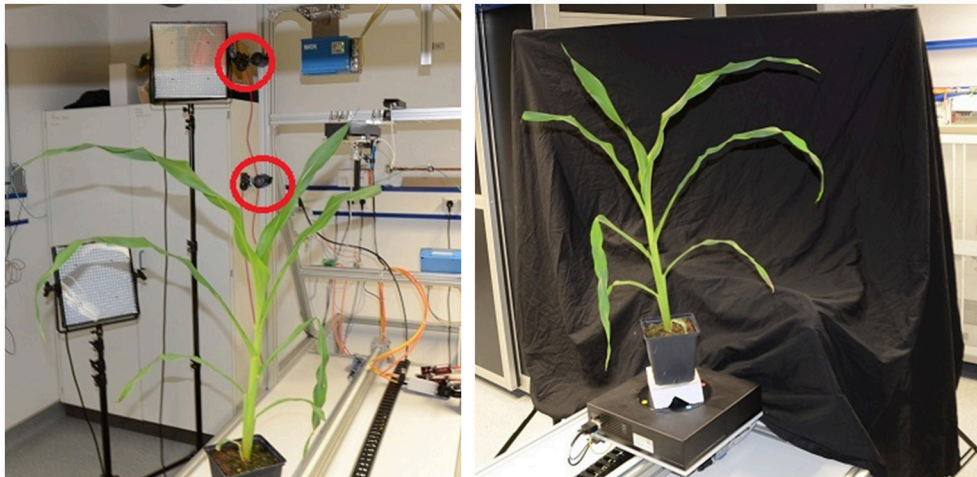


FIGURE 3 | Semi-automatic imaging setup. **(Left)** View from behind the plant with cameras indicated by red circles. **(Right)** Setup as seen from the lower camera.

We investigate memory usage and runtimes for these different refinement strategies in our experiments (see section 4.1).

For our experiments with the banana plants we used a finest resolution of 0.244 mm, being a bit smaller than the pixel size at working distance of $\sim 0.3\text{--}0.4$ mm depending on the camera distance to the voxel of interest. The volume of the visual hull of a plant is estimated by summing over the volumes of all nodes marked as “object” where nodes at the finest resolution, i.e., with unknown filling percentage, are accounted for half their volume in order to get an unbiased estimate. Clearly, as the visual hull includes convex volumes or other occluded volumes, its volume only coarsely reflects the volume of a complex object such as a plant.

For visualization (cmp. e.g., **Figures 7, 8**) we assigned to each “object” voxel a single RGB color value for simplicity. To this end, we calculated for each voxel its nearest camera⁴ and selected the color of the center pixel of the voxel’s bounding box. No visibility or plausibility tests have been performed. Please note, that this approach is good for visualization purposes, only, and is not well suitable for color-based trait evaluation.

3. MATERIALS AND METHODS

3.1. Imaging

Two setups for image acquisition were used for this study, a semi-automatic setup combining a turntable and up to two cameras, as well as a fully automated system including a turntable and three industrial cameras.

The semi-automatic setup (see **Figure 3**) needs manual initial positioning of a plant, which is then imaged fully automatic. This setup includes an automated turntable (Steinmeyer Mechatronic, Dresden Germany, DT180-SM, 0.1° accuracy) with color markers for image-based rotation angle estimation, and two industrial cameras (Allied Vision Technologies, Stadtroda

Germany, AVT MANTA G-504, 5MP color camera). Images can be taken either in a stop-and-go fashion, where the plant is turned to the next pose, stopped, and then images are acquired, or alternatively, the turntable can rotate and images are taken in equidistant time intervals corresponding to equidistant rotation angles.

The fully automated setup called *Screenhouse* at IBG-2 is used for screening of the shoot structure and function of different mono- and dicotyledonous plant species in a greenhouse environment (cmp. Nakhforoosh et al., 2016). Plants are automatically fed to the system being equipped with an imaging station for automated imaging (see **Figure 4**, top and left). Imaging is routinely performed with three cameras (Point Gray Grasshopper2, 5MP color camera, by FLIR Integrated Imaging Solutions Inc., Richmond, British Columbia, Canada) which are located at different positions to efficiently assess the plant structural properties for diverse shoot architectures. A turntable is used for stop-and-go imaging at 4 positions at 0° , 90° , 180° , and 270° . For the top camera only the image acquired at 0° is used. This means that 9 images per plant are used for reconstruction in our experiments.

3.2. Camera Calibration

One of the main drawbacks of visual hull computation with fixed calibrated cameras is its sensitivity to imprecise external camera calibration and lens aberrations. When a mask M_i is misaligned and thus does not well overlap with the “true” object volume, the non-overlapping parts are deleted from the volume without further testing or corrections. Geometric precision of the calibration needs to be in the same range as the desired precision in object space, i.e., when reconstructing relatively large objects sufficient precision is reachable using well established methods. When dealing with small objects like plant seeds, an online image-based camera pose calibration step may be needed (see e.g., Roussel et al., 2016). In our case, plants are in the size range of tens of centimeters and precision needed for calibration

⁴We used as camera position the camera center, as usual.

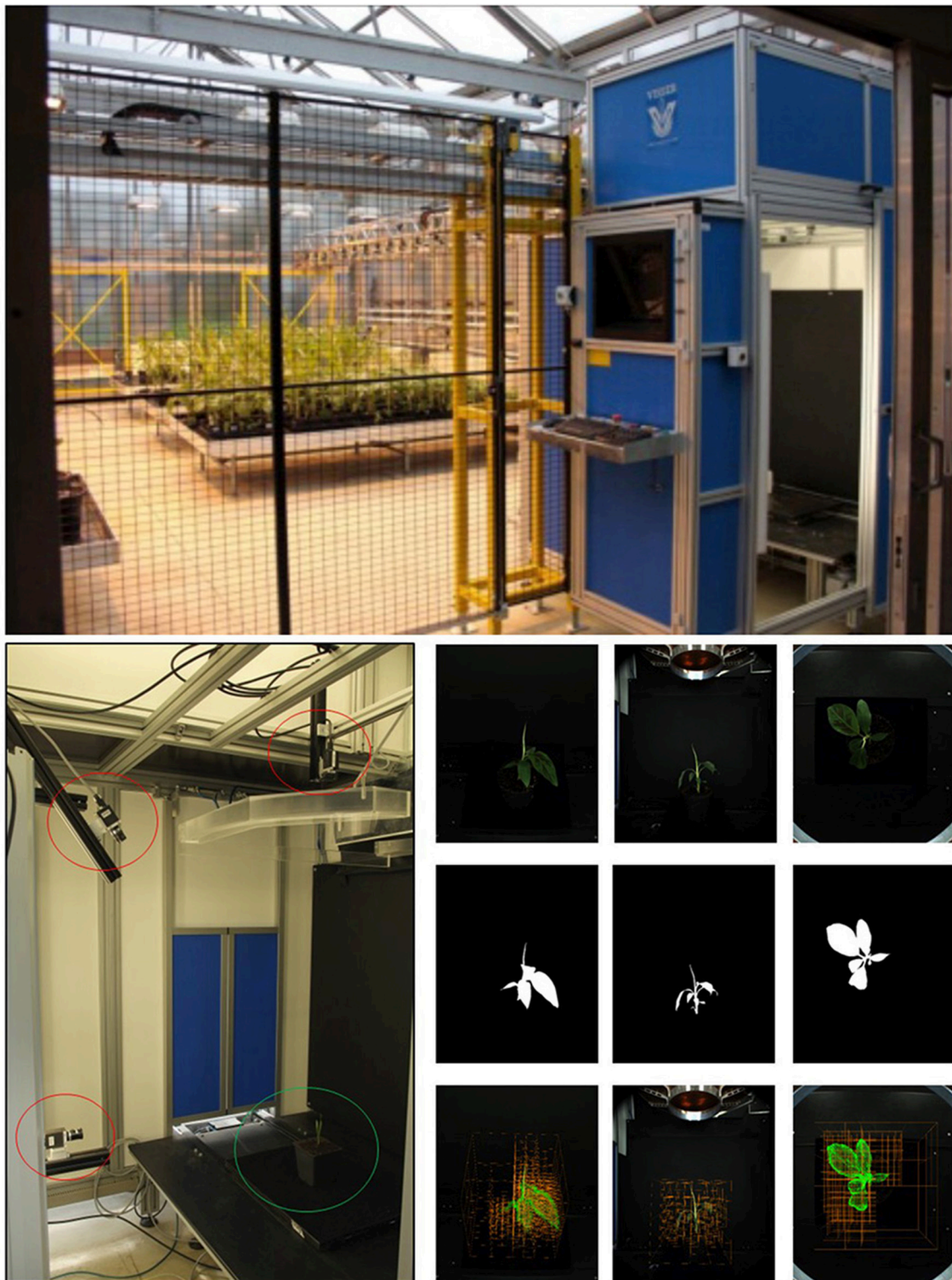


FIGURE 4 | Imaging. **(Top)** Screenhouse setup. **(Bottom left)** Camera setup in Screenhouse measurement chamber. Three 5 MP RGB cameras (red circles) with different view angles and rotating table (green circle). **(Bottom right, top)** Original RGB images taken from 3 different view angles; **(Middle)** Binary masked images; **(Bottom)** Intermediate carving step overlaid on images.

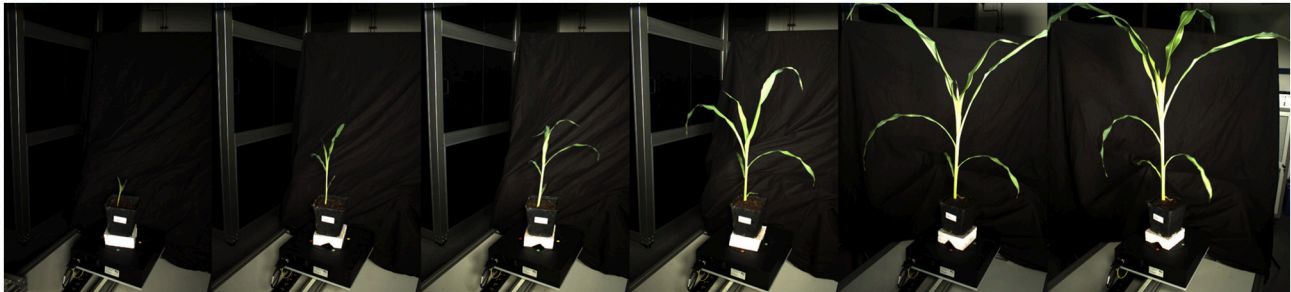


FIGURE 5 | Images of the Maize experiment. The images show the same maize plant over the course of the experiment at roughly the same rotation angle. Imaging was done with continuous rotation and equidistant time intervals for image acquisition using the semi-automatic setup (see section 3.1).

is in the range of ≈ 0.3 mm, which is approximately the lower limit of pixel size at working distance. We use the calibration approach introduced by Bouguet (1999), where multiple, here 20, images of arbitrarily posed flat calibration targets are used. Laser-printed calibration targets glued to flat material are sufficient for our requirements. We use the OpenCV ver. 2.4.9 (Bradski, 2000) implementation *calibrateCamera* of the method presented in Zhang (2000).

3.3. Software Implementation

The software framework was implemented in C++ on a Windows 7 operating system with Microsoft Visual Studio 2013.

3.4. Datasets Acquired

3.4.1. Maize Plants on the Semi-automatic Setup

We imaged 5 maize plants (Zea Maize, Genotype “Badischer Gelber”) at 6 different time points over 17 days from seedling stage with 2–8 leaves stage (see **Figure 5** for an example). Using both cameras of the semi-automatic setup this resulted in 30 datasets of 60 images each, where we had to remove 7 sets due to imaging problems, e.g., plants being already too large for our setup.

3.4.2. Banana Seedlings on the Automatic Setup

Seven groups of five banana seedlings, cultivar Khai Thong Ruang, were imaged and harvested for fresh weight determination immediately after imaging. Imaging and harvesting took place 6, 10, 14, 17, 20, 28, and 34 days after transplantation of the initial four leaf plantlets. This resulted in 35 datasets of 9 images each (see **Figure 6**). In addition some larger banana seedlings have been imaged in the same way, but not been harvested (see **Figure 7**).

4. EXPERIMENTS

4.1. Performance of Different Refinement Strategies

We investigate the runtimes needed for the four different refinement strategies *Depth first*, *Breadth first*, *Refining resolution*, and *Mark and refine* described in Section 2.4. We performed these test on a computer with 48GB RAM and two Intel Xeon E5540 @ 2.53 GHz as some refinement strategies use a lot of

memory. The *Mark and refine* strategy can also be run on a current laptop (8GB RAM, Intel Core i5-5300U @ 2.30 GHz) at comparable speeds. All runtime tests were done without any parallelization using a single CPU only, unless stated differently.

In **Figure 7A** an image from the dataset used for this investigation is shown, one of the larger banana plants being 31 cm high and 29 cm wide. **Figure 7B** shows the corresponding carving result. Its final 3d grid contains $N_{oct} = 5.4 \cdot 10^6$ foreground octree voxels from 0.244 to 15.6 mm. We observe that finest structures, like the tip of the top leaf or the tip of the large leaf on the left are lost. Missing of finest structures in volume carving typically, and also here, has two main reasons: segmentation errors and calibration inaccuracies. The main effect in the data set used here is lens imperfections being not fully compensated by radial distortion correction. In addition, the brownish part of the tip of the top leaf is missing in one segmentation mask and therefore carved away. For the purpose of this publication, i.e., performance evaluation in terms of runtime and resolution, such inaccuracies are of no interest. However for practical applications needing such detail, one may want to invest in better calibration procedures and more precise segmentation algorithms.

A uniform 0.244 mm grid requires $6.9 \cdot 10^{10}$ voxels to represent the same 1 m^3 working volume (corresponding to 64 GB, when using 1 byte per voxel), where $N_{equi} = 3.3 \cdot 10^7$ voxels are foreground i.e., plant. We see that while using an equidistant grid during carving requires a large amount of memory, this is not necessarily the case for the final result. In our case the difference is a factor of $N_{equi}/N_{oct} = 6.1$, allowing to perform subsequent processing using equidistantly sampled foreground voxels, if needed.

Complexity for deriving this carving result heavily depends on the refinement strategy used. **Figure 7C** shows the number of octree voxels generated when iterating through the images of the dataset. As before, “one iteration” means projecting all current octree voxels on one image once.

For the classic equidistantly sampled voxel grid approach, we would need 9 iterations as we have 9 images in the banana dataset. We derived the foreground voxel number plotted in **Figure 7C** from the *Depth first* results using octrees.

For the octree-based method, in order to reach the finest resolution of 0.244 mm, the initial 1 m^3 voxel needs to be refined

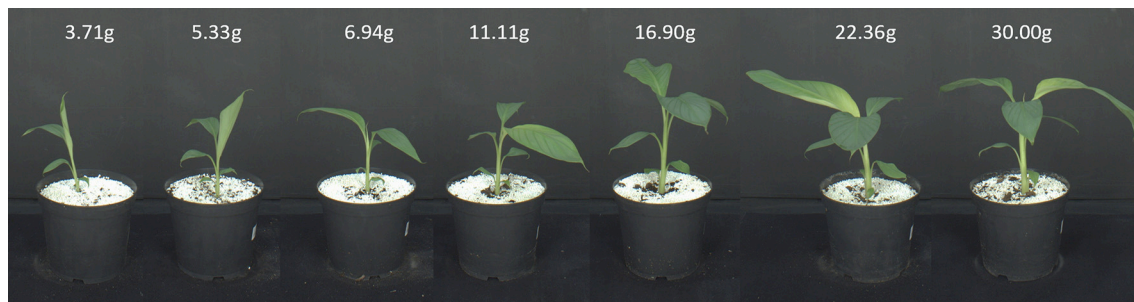


FIGURE 6 | Images of the Banana experiment. The images show one banana seedling of each of the seven groups. Imaging was done using the fully automatic Screenhouse facility (see section 3.1).

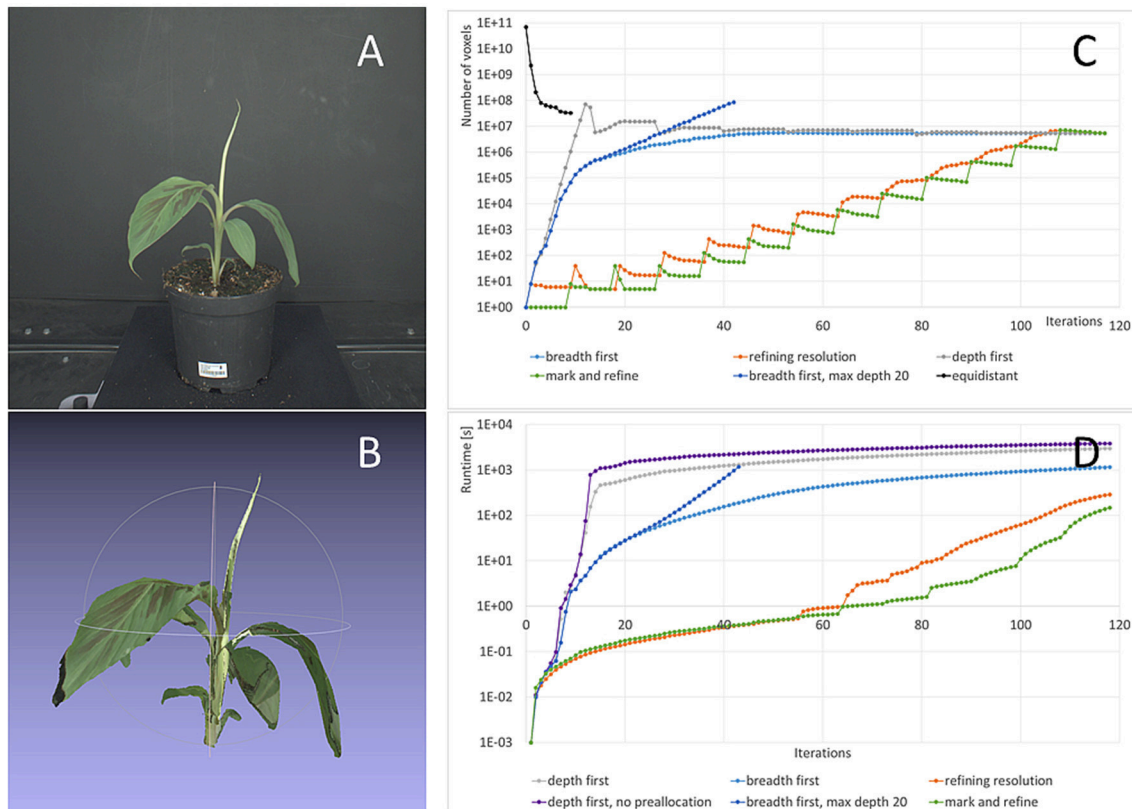


FIGURE 7 | Refinement strategies: Storage complexity and runtimes. **(A)** One original image from the banana dataset used for this test. **(B)** 3d colored point cloud at final resolution (0.244 mm). **(C)** Number of voxels used with our method and the different refinement strategies from Section 2.4 and the classic approach with equidistant sampling. **(D)** Runtimes for different refinement strategies.

12 times. As some final surface voxels may only be seen from a single image, we have to project the voxels to each image 12 times to make sure to reach the final resolution; and one more time to remove the unneeded (background) voxels at final resolution. I.e., 117 iterations are needed overall.

The first recommendation for speed-up, or to be able to finish at a usable result at all, is that the final octree resolution needs to be restricted to the desired resolution. If this restriction is relaxed, many octree voxels are refined unnecessarily. This can be seen in **Figure 7C**, case “breadth first, max. depth 20,” where the number

of voxels keeps increasing exponentially after 18 iterations, in contrast to the case “breadth first,” where we restricted resolution. The unrestricted calculation led to memory overflow and ended before reaching the final result.

The second recommendation for speed-up is to preallocate memory wherever possible. The effect of allocation on the fly⁵ is demonstrated in **Figure 7D**, cases “depth first” and “depth first,

⁵Such allocation is as offered e.g., in C++ using the `vector`-class from the standard template library.

no preallocation” where a factor of 1.3 longer final runtimes can be observed.

Generally the *Depth first* strategy shows by far longest runtimes. The reason is that in the first iterations the number of voxels generated increases exponentially, in our case to 75 M voxels after 12 iterations. This is when the finest resolution is reached for the first image. Then, with every new image voxel number drops with the first iteration, as many voxels are carved away, and then again increases with the next iterations. Due to the large number of voxels to process, runtime is very high (2,997 s).

The *Breadth first* strategy is about a factor 2.6 faster needing 1,155 s to finish. Restricting the reachable resolution with every loop through all images, i.e., using the Refining resolution strategy again yields a speed-up of a factor 4.0, i.e., a factor of 10.4 compared to *Depth first*.

Mark and refine is the fastest refinement strategy, as it produces fewest voxels during processing. It reaches a resolution of 0.98 mm after 7.7 s ($\hat{=}$ 1,024³ voxel grid), 0.49 mm after 32.7 s ($\hat{=}$ 2,048³ voxel grid), and the final 0.244 mm after 148 s ($\hat{=}$ 4,096³ voxel grid). Using *Mark and refine* on the laptop described above the final result is reached after 85.7 s, i.e., a factor 1.7 faster. In Roussel et al. (2016) 218 s are reported for the usual equidistant grid method using a 1,024³ voxel grid and a CPU implementation, 12.56 s using a well optimized GPU implementation. We even outperform this GPU optimized version by a factor 1.6, using no parallelization at all.

We also tested a simple parallelization method for *Mark and refine*, by distributing the voxel-to-image projections over 8 CPU threads, but not threading the splitting step. The final result at 0.244 mm resolution is then reached after 37.7 s. Compared to

the slowest (but still completing) method we observe an overall speedup of a factor 102.

4.2. Runtime for Different Plant Sizes

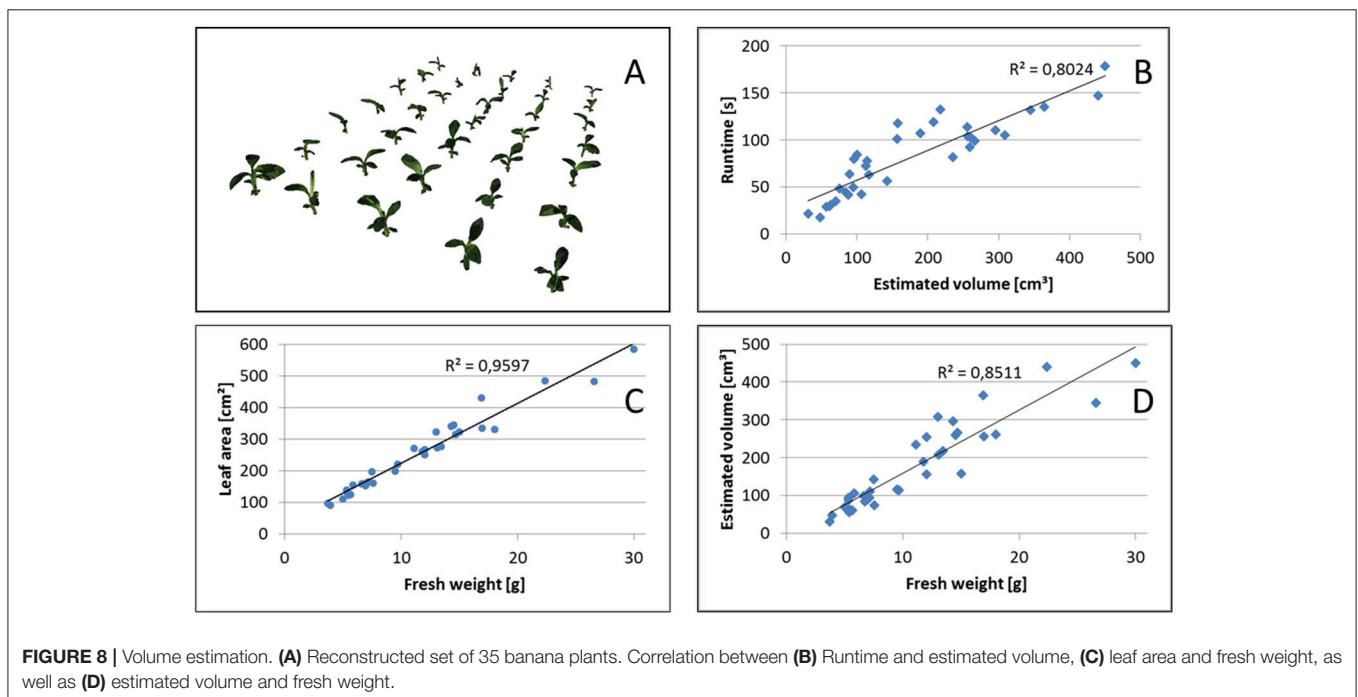
The overall runtime of octree-based volume carving strongly depends on the plant size and complexity, more precisely, the surface area of its visual hull, as the finest voxels occur there. Consequently, we observe that the number of pixels roughly increases by a factor of 4 after each pass through all images in *Mark and refine* (cmp. **Figure 7C**). This is consistent with the assumption that on average 4 of 8 of the new voxels are foreground, when the old voxel was at the surface and thus needed splitting.

Overall runtimes for the reconstruction of the 35 plants of the Banana datasets are shown in **Figure 8B**. We observe, that runtimes are faster for smaller plants, approximately linearly increasing with the estimated volume or, for banana plants, with leaf area. This is the case as the structure of banana plants is dominated by big leaves, i.e., flat objects with area being well correlated with volume or, equivalently, fresh weight (see **Figure 8C**).

4.3. Examples for Visual-Hull Based Plant Traits

We evaluated the general usability of this visual hull reconstruction approach with two scenarios using banana and maize plants (see section 3.4).

In **Figure 8A** reconstructions of the banana plants are visualized and in **Figure 8D** their fresh weight is plotted vs. the estimated volume. We observe that the estimated volume of the visual hull correlates with the fresh weight ($R^2 = 0.8511$), however differences in estimated volume for plants with



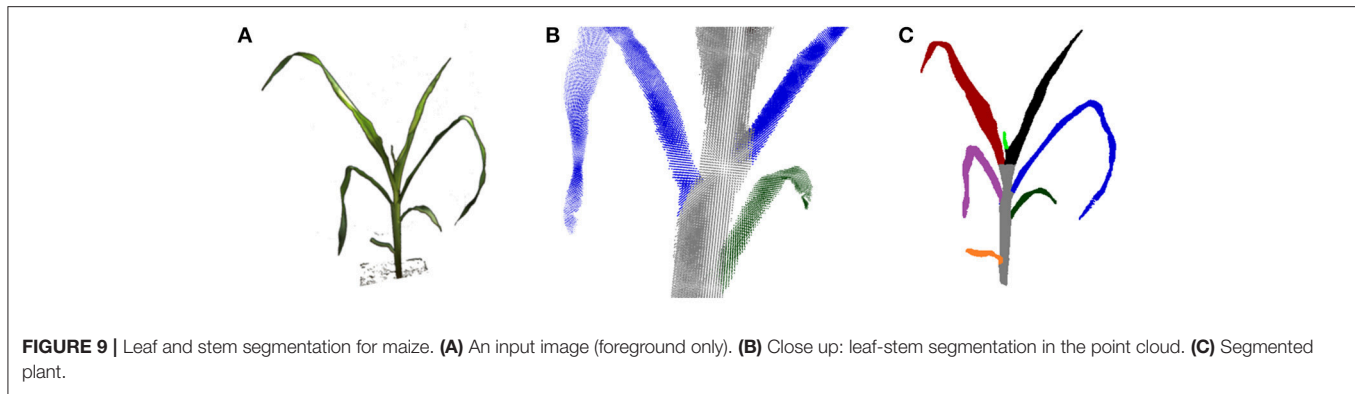


FIGURE 9 | Leaf and stem segmentation for maize. **(A)** An input image (foreground only). **(B)** Close up: leaf-stem segmentation in the point cloud. **(C)** Segmented plant.

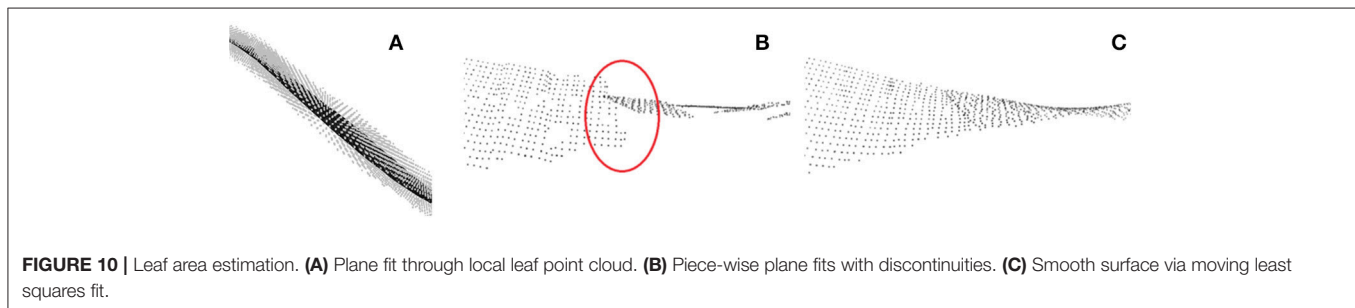


FIGURE 10 | Leaf area estimation. **(A)** Plane fit through local leaf point cloud. **(B)** Piece-wise plane fits with discontinuities. **(C)** Smooth surface via moving least squares fit.

approximately the same weight are up to about a factor of 2. This makes the estimated volume useful as non-invasive surrogate measure for fresh weight in statistical analyses with replication, but not very reliable when comparing individual plants. This is not unexpected, as the visual hull may contain non-negligible non-plant volumes due to occlusions. Such non-plant volumes largely depend on the actual configuration of leaves, i.e., leaf poses.

The visual hull can be used more reliably to derive structural plant properties, like leaf numbers or leaf areas, when applying further processing. In **Figure 9** a leaf segmentation result is depicted for a maize plant. Here a reliable leaf-wise segmentation has been reached by a sequential cluster connecting algorithm and subsequent refinement steps. A detailed description of the algorithm is beyond the scope of this publication. In short, leaf segmentation is done by searching clusters in horizontal one final voxel thick slices, and connecting the found clusters over slices scanning from top to bottom. By this, cluster labels are transferred from top to bottom, forming plant parts. When two clusters merge in one slice, the smaller label is transferred to this slice cluster. By this strategy leaves are segmented, where the stem region still belongs to the top-most leaf. They are separated in a subsequent step using simple but heuristic geometry rules. This leaf segmentation allowed e.g., to count leaves reliably. For 19 of the 23 maize plant datasets leaf numbers were correctly estimated. For the other 4 plants the smallest top leaf was missed, comparable to the small bright green leaf in **Figure 9C**. For further details we refer to Embgenbroich (2015).

When individual leaves are segmented as point clouds, leaf properties can be derived. The length of a leaf may be determined

as the shortest path through the point cloud connecting leaf base and tip. To do so, usual skeletonization algorithms may be used. In order to derive leaf area, the point cloud needs to be flattened. This can be done by piece-wise fitting planes to the leaves as depicted in **Figure 10**. In order to remove steps between fitted plane pieces (**Figure 10B**), a moving least squares method may be applied (**Figure 10C**). For further details we refer to Embgenbroich (2015).

5. CONCLUSION AND OUTLOOK

The presented volume carving using octrees and integral images allows fast and accurate visual hull computation on standard PC hardware with relatively low memory and CPU requirements. A key to efficient computations is the choice of a suitable refinement strategy, as unrestricted refinement leads to exponential “explosion” of voxel numbers and thus complexity. Speedup between efficient and non-efficient implementations can be more than 2 orders of magnitude and differences in memory requirements comparably dramatic. Efficient CPU implementations allow fast execution on usual laptop or desktop compute hardware, outperforming even GPU-optimized “brute force” implementations using equidistantly sampled voxel grids. Even faster implementations may be possible, when porting the presented multi-grid approach to GPUs. Usability of volume carving results represented in octree voxel grids is equivalent to using fine equidistant voxel grids, but at much higher reachable resolutions. Some examples how to derive traits of banana and maize plants have been presented to demonstrate

possible applications in plant phenotyping for deriving leaf level traits.

AUTHOR CONTRIBUTIONS

FF, MM, and HS designed the study. PE, CB, AF, and HS implemented and tested the software. CB and PE performed all lab experiments. FF, MM, and HS drafted the manuscript. All authors contributed to text and figures of the manuscript and approved the final manuscript.

REFERENCES

- Al-Tamimi, N., Brien, C., Oakey, H., Berger, B., Saade, S., Ho, Y. S., et al. (2016). Salinity tolerance loci revealed in rice using high-throughput non-invasive phenotyping. *Nat. Commun.* 7:13342. doi: 10.1038/ncomms13342
- Apelt, F., Breuer, D., Nikoloski, Z., Stitt, M., and Kragler, F. (2015). Phytotyping4d: a light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. *Plant J.* 82, 693–706. doi: 10.1111/tj.12833
- Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008
- Bouguet, J.-Y. (1999). *Visual Methods for Three-dimensional Modeling*. Ph.D. thesis, California Institute of Technology, Pasadena, CA.
- Bradski, G. (2000). The OpenCV library. *Doctor Dobbs J.* 25, 122–125.
- Brenscheidt, M. (2014). Rekonstruktion der visuellen Hülle von Pflanzensamen mithilfe der OpenGL. Bachelor's thesis, Fachhochschule Aachen Campus Jülich.
- Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., et al. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26, 4636–4655. doi: 10.1105/tpc.114.129601
- Clark, R. T., MacCurdy, R. B., Jung, J. K., Shaff, J. E., McCouch, S. R., Aneshansley, D. J., et al. (2011). Three-dimensional root phenotyping with a novel imaging and software platform. *Plant Physiol.* 2, 455–465. doi: 10.1104/pp.110.169102
- Crow, F. C. (1984). Summed-area tables for texture mapping. *SIGGRAPH Comput. Graph.* 18, 207–212. doi: 10.1145/964965.808600
- Das, A., Schneider, H., Burridge, J., Ascanio, A. K., Wojciechowski, T., Topp, C. N., et al. (2015). Digital imaging of root traits (dirt): a high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods* 11:51. doi: 10.1186/s13007-015-0093-3
- Embgrenbroich, P. (2015). *Bildbasierte Entwicklung eines dreidimensionalen Pflanzenmodells am Beispiel von Zea Mays*. Master's thesis, FH Aachen.
- Fanourakis, D., Briese, C., Max, J., Kleinen, S., Putz, A., Fiorani, F., et al. (2014). Rapid determination of leaf area and plant height by using light curtain arrays in four species with contrasting shoot architecture. *Plant Methods* 10:9. doi: 10.1186/1746-4811-10-9
- Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 17.1–17.25. doi: 10.1146/annurev-arplant-050312-120137
- Furbank, R. T., and Tester, M. (2011). Phenomics technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005
- Golbach, F., Kootstra, G., Damjanovic, S., Otten, G., and Zedde, R. (2015). Validation of plant part measurements using a 3d reconstruction method suitable for high-throughput seedling phenotyping. *Mach. Vis. Appl.* 27, 663–680. doi: 10.1007/s00138-015-0727-5
- Golzarian, M. R., Frick, R. A., Rajendran, K., Berger, B., Roy, S., Tester, M., et al. (2011). Accurate inference of shoot biomass from high-throughput images of cereal plants. *Plant Methods* 7:2. doi: 10.1186/1746-4811-7-2
- Hartley, R. I., and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision, 2nd Edn*. Cambridge, UK: Cambridge University Press.
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Schölkopf, B. (1998). Support vector machines. *IEEE Intell. Syst. Appl.* 13, 18–28. doi: 10.1109/5254.708428

ACKNOWLEDGMENTS

The banana dataset has been kindly provided by Eric Opoku Mensah, Thomas Bodewein, and Tobias Wojciechowski and has been acquired in the BMBF funded GlobE BiomassWeb project. Parts of this work was performed within the German-Plant-Phenotyping Network which is funded by the German Federal Ministry of Education and Research (project identification number: 031A053).

- Homolova, L., Maenovskiy, Z., Clevers, J. G. P. W., Garcia-Santos, G., and Schaepman, M. E. (2013). Review of optical-based remote sensing for plant trait mapping. *Ecol. Complex.* 15, 1–16. doi: 10.1016/j.ecocom.2013.06.003
- Hu, H., Zhang, J. Z., Sun, X. Y., and Zhang, X. M. (2013). Estimation of leaf chlorophyll content of rice using image color analysis. *Can. J. Remote Sens.* 39, 185–190. doi: 10.5589/m13-026
- Jahnke, S., Roussel, J., Hombach, T., Kochs, J., Fischbach, A., Huber, G., et al. (2016). phenoseeder - a robot system for automated handling and phenotyping of individual seeds. *Plant Physiol.* 172, 1358–1370. doi: 10.1104/pp.16.01122
- Kanaujia, A., Kittens, N., and Ramanathan, N. (2013). “Part segmentation of visual hull for 3d human pose estimation,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops* (Portland, OR), 542–549.
- Klodt, M., and Cremers, D. (2015). “High-resolution plant shape measurements from multi-view stereo reconstruction,” in *Computer Vision - ECCV 2014 Workshops*, volume 8928 of *Lecture Notes in Computer Science*, eds L. Agapito, M. M. Bronstein, and C. Rother (Zurich: Springer International Publishing), 174–184.
- Koenderink, N. J. J. P., Wigham, M., Golbach, F., Otten, G., Gerlich, R., and van de Zedde, H. J. (2009). “MARVIN: high speed 3d imaging for seedling classification,” in *Seventh European Conference on Precision Agriculture 2009* (Wageningen), 279–286.
- Ladikos, A., Benhimane, S., and Navab, N. (2008). “Efficient visual hull computation for real-time 3d reconstruction using cuda,” in *Proceedings of the 2008 Conference on Computer Vision and Pattern Recognition Workshops* (Anchorage, AK), 1–8.
- Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Trans. Patt. Anal. Mach. Intell.* 16, 150–162. doi: 10.1109/34.273735
- Li, L., Shi, D. Y., and Xu, J. (2013). “Color image segmentation based-on svm using mixed features and combined kernel,” in *Intelligence Science and Big Data Engineering: 4th International Conference, ISIDE 2013, Beijing, China, July 31 – August 2, 2013, Revised Selected Papers*, eds C. Sun, F. Fang, Z.-H. Zhou, W. Yang, and Z.-Y. Liu (Berlin; Heidelberg: Springer Berlin Heidelberg), 401–409.
- Martin, W. N., and Aggarwal, J. K. (1983). Volumetric descriptions of objects from multiple views. *IEEE Trans. Pattern Anal. Mach. Intell.* 5, 150–158. doi: 10.1109/TPAMI.1983.4767367
- Meagher, D. (1980). *Octree Encoding: A New Technique for the Representation, Manipulation and Display of Arbitrary 3-D Objects by Computer*. Troy, NY: Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Image Processing Laboratory.
- Meagher, D. (1995). High-speed image generation of complex solid objects using octree encoding. EP Patent 0,152,741.
- Minervini, M., Scharf, H., and Tsafaris, S. A. (2015). Image analysis: the new bottleneck in plant phenotyping [applications corner]. *IEEE Signal Process. Mag.* 32, 126–131. doi: 10.1109/MSP.2015.2405111
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. *Biosyst. Eng.* 114, 358–371. doi: 10.1016/j.biosystemseng.2012.08.009
- Müller-Linow, M., Pinto-Espinosa, F., Scharf, H., and Rascher, U. (2015). The leaf angle distribution of natural plant populations: assessing the canopy with a novel software tool. *Plant Methods* 11:11. doi: 10.1186/s13007-015-0052-z
- Nakhforoosh, A., Bodewein, T., Fiorani, F., and Bodner, G. (2016). Identification of water use strategies at early growth stages in durum wheat from shoot

- phenotyping and physiological measurements. *Front. Plant Sci.* 7:1155. doi: 10.3389/fpls.2016.01155
- Paprocki, A., Sirault, X., Berry, S., Furbank, R., and Frapp, J. (2012). A novel mesh processing based technique for 3d plant analysis. *BMC Plant Biol.* 12:63. doi: 10.1186/1471-2229-12-63
- Potmesil, M. (1987). Generating octree models of 3d objects from their silhouettes in a sequence of images. *Comput. Vis. Graph. Image Process.* 40, 1–29. doi: 10.1016/0734-189X(87)90053-3
- Roussel, J., Fischbach, A., Jahnke, S., and Scharr, H. (2015). “3D surface reconstruction of plant seeds by volume carving,” in *Computer Vision Problems in Plant Phenotyping 2015* (Swansea), 13.
- Roussel, J., Geiger, F., Fischbach, A., Jahnke, S., and Scharr, H. (2016). 3d surface reconstruction of plant seeds by volume carving: performance and accuracies. *Front. Plant Sci.* 7:745. doi: 10.3389/fpls.2016.00745
- Szeliski, R. (1993). Rapid octree construction from image sequences. *CVGIP Image Underst.* 58, 23–32. doi: 10.1006/ciun.1993.1029
- Topp, C. N., Iyer-Pascuzzi, A. S., Anderson, J. T., Lee, C.-R., Zurek, P. R., et al. (2013). 3d phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1695–E1704. doi: 10.1073/pnas.1304354110
- Veksler, O. (2003). “Fast variable window for stereo correspondence using integral images,” in *Computer Vision and Pattern Recognition* (Madison, WI), 1–6.
- Viola, P., and Jones, M. J. (2004). Robust real-time face detection. *Int. J. Comput. Vis.* 57, 137–154.
- Waizenegger, W., Feldmann, I., Eisert, P., and Kauff, P. (2009). “Parallel high resolution real-time visual hull on gpu,” in *Proceedings of the 16th IEEE International Conference on Image Processing, ICIP'09* (Cairo: IEEE Press), 4245–4248. doi: 10.1109/ICIP.2009.5413661
- Walter, A., Scharr, H., Gilmer, F., Zierer, R., Nagel, K. A., Ernst, M., et al. (2007). Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. *New Phytol.* 174, 447–455. doi: 10.1111/j.1469-8137.2007.02002.x
- Wang, X.-Y., Wang, T., and Bu, J. (2011). Color image segmentation using pixel wise support vector machine classification. *Patt. Recogn.* 44, 777–787. doi: 10.1016/j.patcog.2010.08.008
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Trans. Patt. Anal. Mach. Intell.* 22, 1330–1334. doi: 10.1109/34.888718
- Zheng, Y., Gu, S., Edelsbrunner, H., Tomasi, C., and Benfey, P. (2011). “Detailed reconstruction of 3d plant root shape,” in *Proceedings of the 2011 International Conference on Computer Vision, ICCV '11* (Washington, DC: IEEE Computer Society), 2026–2033.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer SL and handling Editor declared their shared affiliation.

Copyright © 2017 Scharr, Briese, Embgenbroich, Fischbach, Fiorani and Müller-Linow. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Chlorophyll Fluorescence Imaging Uncovers Photosynthetic Fingerprint of Citrus Huanglongbing

Haiyan Cen¹, Haiyong Weng¹, Jieni Yao¹, Mubin He¹, Jingwen Lv², Shijia Hua¹, Hongye Li² and Yong He^{1*}

¹ College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China, ² College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China

OPEN ACCESS

Edited by:

Eetu Puttonen,
National Land Survey of Finland,
Finland

Reviewed by:

Alireza Pourreza,
University of California System,
United States
Luis Gustavo Marcassa,
University of São Paulo, Brazil

*Correspondence:

Yong He
yhe@zju.edu.cn

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 10 May 2017

Accepted: 15 August 2017

Published: 29 August 2017

Citation:

Cen H, Weng H, Yao J, He M, Lv J,
Hua S, Li H and He Y (2017)
Chlorophyll Fluorescence Imaging
Uncovers Photosynthetic Fingerprint
of Citrus Huanglongbing.
Front. Plant Sci. 8:1509.
doi: 10.3389/fpls.2017.01509

Huanglongbing (HLB) is one of the most destructive diseases of citrus, which has posed a serious threat to the global citrus production. This research was aimed to explore the use of chlorophyll fluorescence imaging combined with feature selection to characterize and detect the HLB disease. Chlorophyll fluorescence images of citrus leaf samples were measured by an in-house chlorophyll fluorescence imaging system. The commonly used chlorophyll fluorescence parameters provided the first screening of HLB disease. To further explore the photosynthetic fingerprint of HLB infected leaves, three feature selection methods combined with the supervised classifiers were employed to identify the unique fluorescence signature of HLB and perform the three-class classification (i.e., healthy, HLB infected, and nutrient deficient leaves). Unlike the commonly used fluorescence parameters, this novel data-driven approach by using the combination of the mean fluorescence parameters and image features gave the best classification performance with the accuracy of 97%, and presented a better interpretation for the spatial heterogeneity of photochemical and non-photochemical components in HLB infected citrus leaves. These results imply the potential of the proposed approach for the citrus HLB disease diagnosis, and also provide a valuable insight for the photosynthetic response to the HLB disease.

Keywords: citrus, Huanglongbing, chlorophyll fluorescence imaging, photosynthesis, feature selection, classification

INTRODUCTION

Huanglongbing (HLB), also known as greening disease, is one of the most destructive diseases to citrus industry. It is caused by a non-cultured phloem-restricted bacterium *Candidatus Liberibacter* with three different species, *Liberibacter africanus* (Laf), *Liberibacter asiaticus* (Las), and *Liberibacter americanus* (Lam), and is transmitted by the Asian citrus psyllid *Diaphorina citri* (Bové, 2006). HLB has threatened citrus production worldwide, including Asia, Africa, the United States, and Brazil. The infected trees are asymptomatic at early stages and become symptomatic anywhere from months to years varied with the cultivars. It could serve as a source of inoculums for psyllids with 60 days from the initial infection until it is infective (Lee et al., 2015), resulting in a rapid spread of infection in a grove of citrus trees. The leaf yellowing with blotchy mottle is considered as the most typical diagnostic symptom of HLB disease, especially on sweet oranges. The severely infected trees have smaller leaves with mottled patterns that are often similar

to some nutrient deficient symptoms. In addition, the production of the HLB infected trees would dramatically decline with small and lopsided fruits, leading to the enormous economic loss.

So far, no effective control and cure measures are available for the HLB disease. There is an urgent need to develop an efficient method for the rapid and early detection of HLB disease in both field and laboratory conditions, and thereby in need of removal of the infected trees to prevent the further spread. The current commonly used HLB diagnosis involves experience-based inspection in the orchard and laboratory diagnostic tests. The former is usually performed by professional experts to recognize the typical symptoms of the disease in the field survey. Reported studies have shown that the accuracy of identifying an HLB infected tree by visual inspection is lower than 59% (Pourreza et al., 2015). Polymerase chain reaction (PCR) method that amplifies a DNA target sequence of the pathogen is usually employed as a standard laboratory-based diagnostic method. Although PCR can achieve a relative high accuracy for the HLB detection, it requires a time-consuming and labor-intensive sample preparation, and is not suitable for the continuous onsite detection and early disease warning. Therefore, a fast and non-destructive method is needed for growers to monitor their groves and control the spread of the disease.

Over the years, reflectance spectroscopy and imaging methods based on the selected or broad visible and near-infrared (VIS-NIR) electromagnetic wavelengths were investigated to detect plant diseases on the leaf and/or canopy levels (Inoue et al., 2012; Gnyp et al., 2014; Kusnierek and Korsaeath, 2015), and several studies have been reported on detection of citrus diseases. Qin et al. (2011) used two-band ratio images selected from the hyperspectral images to detect the canker disease of citrus fruit with the overall classification accuracy of 95.7%. Mishra et al. (2012) employed a VIS-NIR spectroscopy technique combined with three different classifiers for identifying HLB infected citrus trees. Although a reasonable classification accuracy was achieved by multiple spectral measurements of the tree canopy, the variation of sunlight and other environmental factors under real field conditions can add additional noises and reduce the detection accuracy. Later, they developed a vision sensor based on the polarization planar rotation of light by the starch at 591 nm, which improved the HLB detection accuracy in both zinc-deficient and non-zinc-deficient classes (Pourreza et al., 2014). However, the applicability of this method largely depends on the starch accumulation that varies in citrus cultivars and growing conditions (i.e., different orchards and seasonal factors). Sankaran et al. (2013) demonstrated the applicability of VIS-NIR and thermal imaging for detection of HLB disease in citrus trees, and they obtained an average overall accuracy of 87% for trees with symptomatic leaves. Li X. et al. (2012) applied multispectral and hyperspectral airborne imaging to detect HLB infected trees in citrus groves. However, the classification accuracies are relative low ranging from 28.7 to 90.2% by using different image classification methods due to the large positioning error of the ground truth. Garcia-Ruiz et al. (2013) compared two aerial imaging platforms for identifying HLB infected citrus trees with the classification accuracies in the range of 67–85% from

unmanned aerial vehicle-based data, and 61–74% with aircraft-based data. In general, these results were not better than those from the ground-based remote sensing methods.

Fluorescence spectroscopy is also considered as a promising method for the rapid and early detection of biotic and abiotic stress response in plants. Fluorescence emissions from molecules of certain compounds such as plant pigments in leaves can be captured after the natural and artificial ultraviolet (UV) light excitation on a small sampling point (Pereira et al., 2011; Harbinson, 2013). Sankaran and Ehsani (2013) used a commercial handheld fluorescence sensor to collect fluorescence signal from healthy and HLB infected leaves of different citrus cultivars, and 97% classification accuracy in the discrimination of healthy and symptomatic HLB infected samples was achieved based on the bagged decision tree (BDT) classifier. However, the classification accuracy was reduced to about 81% when using the validation dataset from Hamlin and Valencia samples, and the overall accuracy of differentiating healthy from asymptomatic HLB infected samples was also very low with the best result of 48.2% when using the support vector machine. Due to the limitation of the sampling area, the spatial heterogeneity in a leaf caused by plant diseases or other physiological disorders may not be accurately described. Researchers proposed that the fluorescence imaging technique could be more useful in detecting the plant physiological response since it can obtain the fluorescence signal as well as the within-plant variation (Pinto et al., 2016; Song et al., 2016). Different fluorescence imaging techniques with different excitation modes have been developed. Steady-state UV light-induced fluorescence imaging has been used to estimate anthocyanin in strawberries and visualize systemic viral infections in *Nicotiana benthamiana* plants by monitoring the signals of the chlorophyll fluorescence and blue-green fluorescence, respectively (Pineda et al., 2008a; Yoshioka et al., 2013). Fluorescence imaging spectroscopy with 530 nm excitation that can obtain a fluorescence image at 690 nm, was also employed to discriminate symptomatic HLB infected leaves from healthy ones sampling from two different orchards (Wetterich et al., 2013). Results showed a good accuracy (90%) for Brazil samples but low (61%) for United States samples. Later, they applied another two excitation sources (405 and 470 nm) to detect citrus HLB disease from nutrient deficient leaves, as well as healthy ones (Wetterich et al., 2016, 2017). The combination of fluorescence bands from two excitations improved the classification accuracy in the range of 92–95% when discriminating HLB from zinc-deficiency by using different machine learning methods. However, more detailed knowledge about energy partition in photosystem II (PSII) should be understood in HLB infected leaves. Kinetic chlorophyll fluorescence imaging provides an efficient way to trace energy partitioning in the photosystem II (PSII) and monitor the electron transport pattern in photosynthesis. It plays an important role in understanding the fundamental mechanism of photosynthesis, pathology, and phenotypic plasticity to genetic variations and plant environmental changes (Murchie and Lawson, 2013; Choi et al., 2016; Sui et al., 2017). Kinetic chlorophyll fluorescence imaging has been widely used to evaluate freeze-thaw and drought tolerance in *Arabidopsis* (Ehlert

and Hinch, 2008; Bresson et al., 2015), salt stress response in wheat (Mehta et al., 2010), chronic ozone damage to soybean leaves (Chen et al., 2009), and plant virus infection (Pineda et al., 2011; Lei et al., 2017). However, few studies have used kinetic chlorophyll fluorescence imaging to study the HLB disease effects on plant photosynthetic functions and complex pathogenesis. Furthermore, the spatial heterogeneity during the disease development implies additional challenges for the rapid diagnosis of HLB disease.

Therefore, this research was aimed to characterize the photosynthetic function of the HLB infected leaves by measuring the fluorescence signals using kinetic chlorophyll fluorescence imaging, and extract the photosynthetic fingerprints that can uniquely identify the HLB disease by performing the advanced machine learning and statistical analysis. The findings in this research provide an important insight to understand the citrus HLB disease infection related to changes in plant photosynthetic activities, and demonstrate that kinetic chlorophyll fluorescence imaging could offer a rapid and non-invasive means for detecting HLB in citrus trees.

MATERIALS AND METHODS

Leaf Sample Collection

Citrus leaf sampling was carried out at a commercial orchard in Nanping, Fujian Province, China in March 2016. The branches were detached from the citrus trees from three categories including healthy, HLB infected (symptomatic and asymptomatic), and nutrient deficient (zinc and magnesium deficient) samples marked by the HLB experts, and immediately wrapped with wet cottons and placed inside a cooler to avoid desiccation. Leaves were then detached from the branches in the laboratory (Supplementary Figure S1), and chlorophyll fluorescence image acquisition was performed immediately after detachment.

Kinetic Chlorophyll Fluorescence Imaging

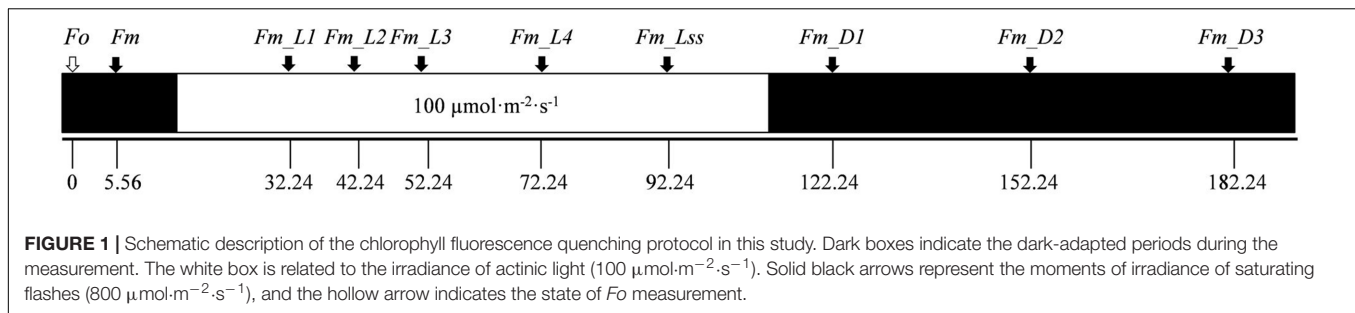
Chlorophyll fluorescence images of leaf samples were measured by an in-house chlorophyll fluorescence imaging system (FluorCam FC800, Photon Systems Instruments, Brno, Czechia) (Supplementary Figure S2) after dark adaptation. A CCD camera with a prime lens (SV-H1.4/6, VS Technology, Tokyo, Japan) was used to capture chlorophyll fluorescence transients in a batch of images with the spatial resolution of 696×520 . Four light-emitting diodes (LEDs) panels with the incident angle of 45° were installed as the light source, which include two red-orange LEDs (620 nm) panels ($<0.1 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) for flashes and actinic light 1 ($0\text{--}250 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), and two cool white LEDs ($>8000\text{ K}$) panels ($0\text{--}1600 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) for actinic light 2 and saturating flashes ($0\text{--}3000 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$). A leaf sample holder with a manually adjustable vertical stage was used for positioning samples to an imaging distance of 20 cm from the lens.

Before the experiment, a preliminary study was performed to determine the optimal protocol for the kinetic chlorophyll fluorescence image acquisition of citrus leaves. The dark-adapted

time was determined by checking the value of the maximum PSII quantum yield (F_v/F_m) of citrus leaves at different dark-adapted time as shown in Supplementary Figure S3. It was observed that F_v/F_m was stable with the value of 0.80 after 20 min dark adaptation, which was considered as the optimal time. Although this 20 min period may be a problem for practical applications, it is appropriate for achieving this research goal. In addition, the intensity of the saturating pulse was described in percentage in image acquisition software. 50% was considered as the optimal intensity of the saturating pulse when the value of F_v/F_m reached 0.81 (Supplementary Figure S4), which corresponded with the absolute intensity value of $800 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ when measured at the position of 20 cm distance from the lens using a quantum meter (Model MQ-200, Apogee Instruments, Inc. United States). The detailed chlorophyll fluorescence quenching protocol is described in **Figure 1**. The minimum fluorescence in dark-adapted state (F_o) was measured after 20 min dark adaptation, and followed by a strong flash of light of $800 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ at 5.56 s that transiently reduces the plastoquinone pool and the primary quinone acceptor Q_A so that the maximum fluorescence in dark-adapted state (F_m) was recorded. The leaf sample was then exposed to an actinic light with the intensity of $100 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ for 70 s. Saturating flashes ($800 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) were applied to measure the maximum fluorescence during light adaptation ($F_m\text{Ln}$) at 32.24, 42.24, 52.24, and 72.24 s, respectively, and the maximum fluorescence at the light-adapted steady state ($F_m\text{Lss}$) was obtained during the saturating flash ($t = 92.24\text{ s}$) at the end of the actinic light period. The instantaneous maximum fluorescence signals during dark relaxation ($F_m\text{Dn}$) with the saturating flashes at 122.24, 152.24, and 182.24 s were also measured. Based on the measured fluorescence signals, the variable fluorescence in dark-adapted state ($F_v = F_m - F_o$), instantaneous non-photochemical quenching during light adaptation ($\text{NPQLn} = (F_m - F_m\text{Ln})/F_m\text{Ln}$), steady-state non-photochemical quenching ($\text{NPQLss} = (F_m - F_m\text{Lss})/F_m\text{Lss}$), and instantaneous non-photochemical quenching during dark relaxation ($\text{NPQDn} = (F_m - F_m\text{Dn})/F_m\text{Dn}$) can be also obtained.

PCR Analysis

Polymerase chain reaction analysis was conducted to confirm the HLB status of leave samples at Plant Pathology Laboratory of Zhejiang University. Two sets of primers designed from conserved regions of 16S ribosomal DNA were synthesized for HLB detection at TSINGKE Biological Technology, China. DNA was extracted from the healthy, nutrient deficient, and HLB infected leaf samples using CTAB (Cetyltrimethylammonium Bromide) methods with the following protocol, respectively: 200 mg midrib tissue of each leaf was transferred to a 2 mL centrifuge tube and ground to the fine powder in liquid nitrogen using an automatic grinding miller (Tissuelyser-64, Shanghai Jingxin Industrial Development Co., Ltd.). 800 μL CTAB was then added into the tube, and heated at 65°C in water bath for 30 min with intermittent agitation. 800 μL mixture (phenol: chloroform: isopropanol = 25: 24: 1) was added and the tube was centrifuged at 12,000 rpm for 10 min at room temperature.



500 μL supernatant was then transferred into a new 2 mL tube, and 1000 μL 95% alcohol and 150 μL NaAC were added. The homogenate was placed on the ice for 12 min and then centrifuged at 12,000 rpm for 2 min. DNA was re-precipitated with 400 μL of 75% alcohol at 12,000 rpm for 2 min. 100 μL ddH₂O was added to the precipitation and DNA was stored at -20°C at last. PCR amplification was conducted in a 20 μL mixtures reaction using 7.5 μL ddH₂O, 1 μL Primer-F, 1 μL Primer-R, 10 μL green taq mix (Vazyme) and 0.5 μL DNA template. The detailed PCR amplification protocol can be found in the literature (Gouda et al., 2006).

Data Analysis

A total of 26 images for each leaf sample related to the fluorescence quenching process were obtained from kinetic chlorophyll fluorescence imaging, which provide detailed information of dynamics of PSII activities about plant status. Commonly used fluorescence parameters such as F_v/F_o , F_v/F_m , and NPQ_{Lss} can be easily calculated by averaging the intensity of the region of interest (ROI) for the corresponding fluorescence image. They are considered as the important parameters in the analysis of photosynthesis associated with plant physiological changes, and their capabilities of identifying the HLB infected leaves from the healthy or nutrient deficient leaves were investigated by advanced machine learning and statistical analysis. Feature selection methods were performed on all the fluorescence parameters to extract the unique fluorescence features that could develop a photosynthetic fingerprint of the HLB disease. Three feature selection methods including random forest (RF), sequential forward selection (SFS), and Monte Carlo uninformative variable elimination (MC-UVE) were used. RF algorithm is based on the reversible jump Markov Chain, and its output provides the selection probability of each feature that can be used as a measure of feature importance. The detailed information of RF can be found in the literature (Li H. et al., 2012). The SFS algorithm is a bottom-up process that starts with an empty feature subset and repeatedly adds the most significant feature selected by an objective function. The feature cannot be discarded at a later stage once it is retained. The fisher criterion is usually used as an objective function (Geng, 2014). MC-UVE develops multiple models with randomly selected calibration sample set produced by the MC method, and the variable is then evaluated with the stability of the corresponding coefficient of the model (Han et al., 2008).

The development of the HLB disease and the spread of the symptoms vary spatially in the leaf, which result in a large variation of fluorescence intensities in the image. Therefore, image features were also investigated for the further classification. Principal component analysis (PCA) was applied to reduce the dimension of the chlorophyll fluorescence image cubes and obtain an uncorrelated orthogonal basis set from the original image set. Scale-invariant feature transform (SIFT) was then used to detect the local features in principal component images. The feature descriptors were first computed as orientation histograms and then transformed to a 128 dimensional SIFT feature vector, which is invariant to the image scale and rotation and robust to local geometric distortions.

Both the mean fluorescence parameter-based and image-based classification models were developed for detecting the HLB disease, and the inputs of the models were the features extracted from the mean fluorescence parameters and the principal component images, respectively. Two classifiers including partial least squares discriminant analysis (PLS-DA) and support vector machine (SVM) were employed to discriminate the HLB infected leaves from the healthy and the nutrient deficient ones. The PLS-DA algorithm is an extension of the PLS model, where the dependent variable is a vector that represents the class label values for each class (Aliakbarzadeh et al., 2016). The SVM classifier is developed based on the statistical learning theory to find a hyperplane that gives the largest distance between the margins of the training data set, and it can be achieved by solving a convex quadratic programming problem using a kernel function (Chapelle et al., 2002; Cen et al., 2016). The three-class classification was performed by labeling “1” for healthy samples, “2” for HLB infected samples and “3” for nutrient deficient samples. All the samples were divided into two groups with 60 for the training set and 30 for the validation set by using Kennard-Stone algorithm (Macho et al., 2001). The machine learning and the statistical analysis were performed using Matlab 2011a (The Mathworks, Inc., Natick, MA, United States) and IBM SPSS Statistics (version 20.0, IBM Corporation, Armonk, NY, United States).

RESULTS

PCR Result

The status of leaves was verified though amplifying target DNA template obtained from leaf midrib tissue. The sensitivity of

PCR analysis with different primer sets is shown in **Figure 2**. Two different sizes of amplicons (1160 and 161 bp) from the 16S ribosomal DNA represented the HLB bacterium DNA template amplified with two specific primers (OI1/OI2c and rplLAS-F/rplLAS-R), respectively. It is indicated that lanes 1 and 4 were PCR positive, and lanes 2, 3, 5, and 6 were PCR negative. Finally, citrus samples were divided into three classes with 30 healthy, 30 HLB infected, and 30 nutrient deficient samples.

Effect of HLB Infection on Commonly Used Chlorophyll Fluorescence Parameters

Figure 3 shows a statistical summary (i.e., mean, range, median, outliers, and quartiles) of four commonly used chlorophyll fluorescence parameters, F_o , F_v/F_o , F_v/F_m , and NPQ_Lss , in healthy, HLB infected, and nutrient deficient leaves. There was no statistically significant difference in F_o among the three classes (**Figure 3A**); while the variation of F_o in the HLB infected leaves was larger than those in healthy and nutrient deficient samples. Compared with healthy leaves, F_v/F_o decreased by 48.9 and 51.1% in HLB infected and nutrient deficient leaves, respectively, indicating the decreased photosynthetic rate in unhealthy leaves. However, no difference in F_v/F_o between HLB infected and nutrient deficient classes was observed (**Figure 3B**). The parameters F_v/F_m and NPQ_Lss showed significant differences among healthy, HLB infected and nutrient deficient leaves (**Figures 3C,D**). F_v/F_m of healthy samples was higher than those of other two classes, and F_v/F_m of HLB infected samples decreased less than those in nutrient deficient ones (**Figure 3C**). The values of NPQ_Lss in HLB infected and nutrient deficient leaves increased compared with those in healthy leaves. This result provided the first screening of HLB disease using the commonly used chlorophyll fluorescence parameters.

Furthermore, a set of chlorophyll fluorescence quenching parameters (F_v/F_o_Ln , F_v/F_m_Ln , NPQ_Ln , F_v/F_o_Dn , F_v/F_m_Dn , and NPQ_Dn) related to the stress-induced changes in the photosynthetic process were analyzed using the spider plots as shown in **Figure 4**. It includes the parameters measured at the dark-adapted state, light adaptation, and dark relaxation. The distance from the center of the spider plot indicates the relative change of the fluorescence parameter with different leaf conditions. Generally, fluorescence quenching parameters of HLB infected and nutrient deficient leaves exhibited a strong contrast compared with those of healthy samples, which is consistent with the reported studies (Pineda et al., 2008b; Spoustová et al., 2013; Mishra et al., 2016). The instantaneous variable-to-initial fluorescence and the maximal PSII quantum yield during light adaptation (F_v/F_o_Ln and F_v/F_m_Ln) and dark relaxation (F_v/F_o_Dn and F_v/F_m_Dn) dramatically decreased in HLB infected and nutrient deficient leaves. The difference in instantaneous maximal PSII quantum yield at multiple phases between the HLB infected and nutrient deficient samples was also observed. The value of instantaneous non-photochemical quenching in HLB infected and nutrient deficient leaves during light adaptation increased significantly, particularly, a large difference

between the HLB infected and the nutrient deficient leaves was observed during the third (NPQ_L3) and fourth (NPQ_L4) of the saturating flashes and the light-adapted steady-state (NPQ_Lss). Generally, the results shown in **Figures 3, 4** reveal that commonly used fluorescence parameters have the potential for identifying the unhealthy leaves from the healthy ones, but it might be a challenge to differentiate the HLB infected leaves from the nutrient deficient ones by only using these parameters.

HLB Detection Based on Optimal Features Extracted from Mean Fluorescence Parameters

Commonly used chlorophyll fluorescence parameters have demonstrated that both HLB infection and nutrient deficiency could cause the change of fluorescence emission as presented in **Figures 3, 4**. However, they may not be the best parameters that would provide an accurate detection of HLB disease. Feature selection methods were applied here to extract the optimal features from all the mean fluorescence parameters that could characterize HLB infected leaves, and their performances are shown in **Figure 5**. It was observed that the classification accuracies varied greatly with the number of selected features, and exhibited the increasing tendency with the increased feature number. In general, the SFS selected features clearly exceeded those obtained from RF and MC-UVE methods when using both SVM and PLS-DA classifiers for three-class classification; while the performance of the RF and MC-UVE was more influenced by the classifier. As shown in **Figure 5B**, SFS-SVM achieved the best classification result of 90% with the feature subset size of 11, denoting that these 11 features might involve the most important fluorescence signatures related to the photosynthesis process affected by the HLB disease and nutrient deficiency, resulting in a better classification among healthy, HLB infected, and nutrient deficient leaves. It was also noted that the classification performance with selected features by SFS was better than that with the full dataset.

The best classification results using three feature selection methods with two classifiers based on mean fluorescence parameters are presented in **Table 1**. The SVM classifier significantly outperformed PLS-DA with the classification accuracy of 90% when using the features from all three feature selection methods. Among the combinations of the two classifiers and three feature selection methods, SFS with each of the two classifiers tended to produce the smallest feature subsets of 6 for PLS-DA and 11 for SVM with the high classification accuracies. The result also showed that the non-photochemical quenching related parameters measured during light adaptation were all selected by three methods, which was consistent with the previous analysis in **Figure 4C**.

Chlorophyll Fluorescence Image-Based Classification

Chlorophyll fluorescence images related to the first three optimal parameters selected by RF, SFS, and MC-UVE methods were shown in **Figure 6**. The variations among healthy, HLB infected

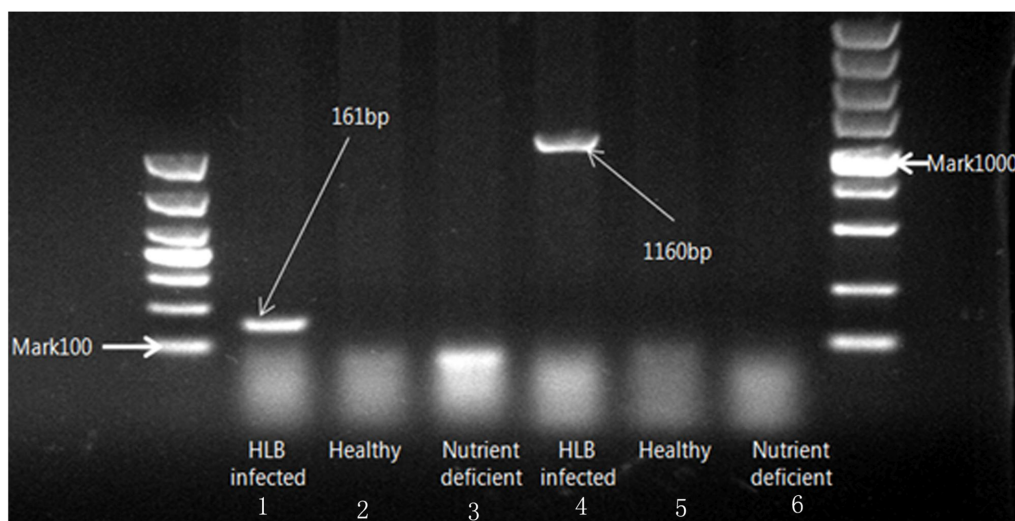


FIGURE 2 | Electrophoresis on 1% agarose gel of DNA amplified with two specific primers (OI1/OI2c and rplLAS-F/rplLAS-R).

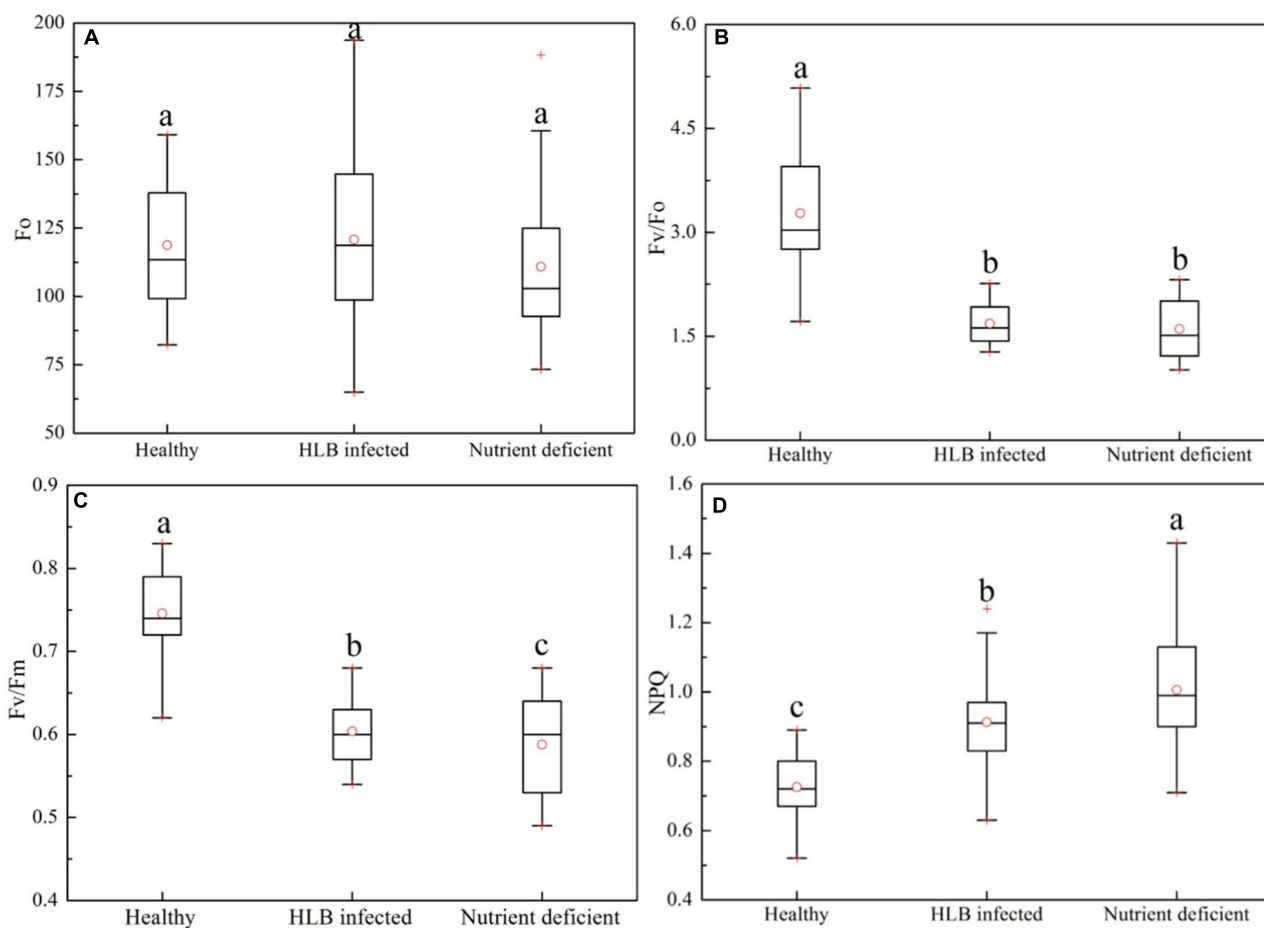


FIGURE 3 | Box-and-whisker plots of commonly used chlorophyll fluorescence parameters including (A) F_o , (B) F_v/F_o , (C) F_v/F_m , and (D) NPQ of healthy, HLB infected and nutrient deficient leaves. Different letters indicated significant differences ($p < 0.05$) by using Duncan test. Values in the figure were mean \pm standard deviation (SD).

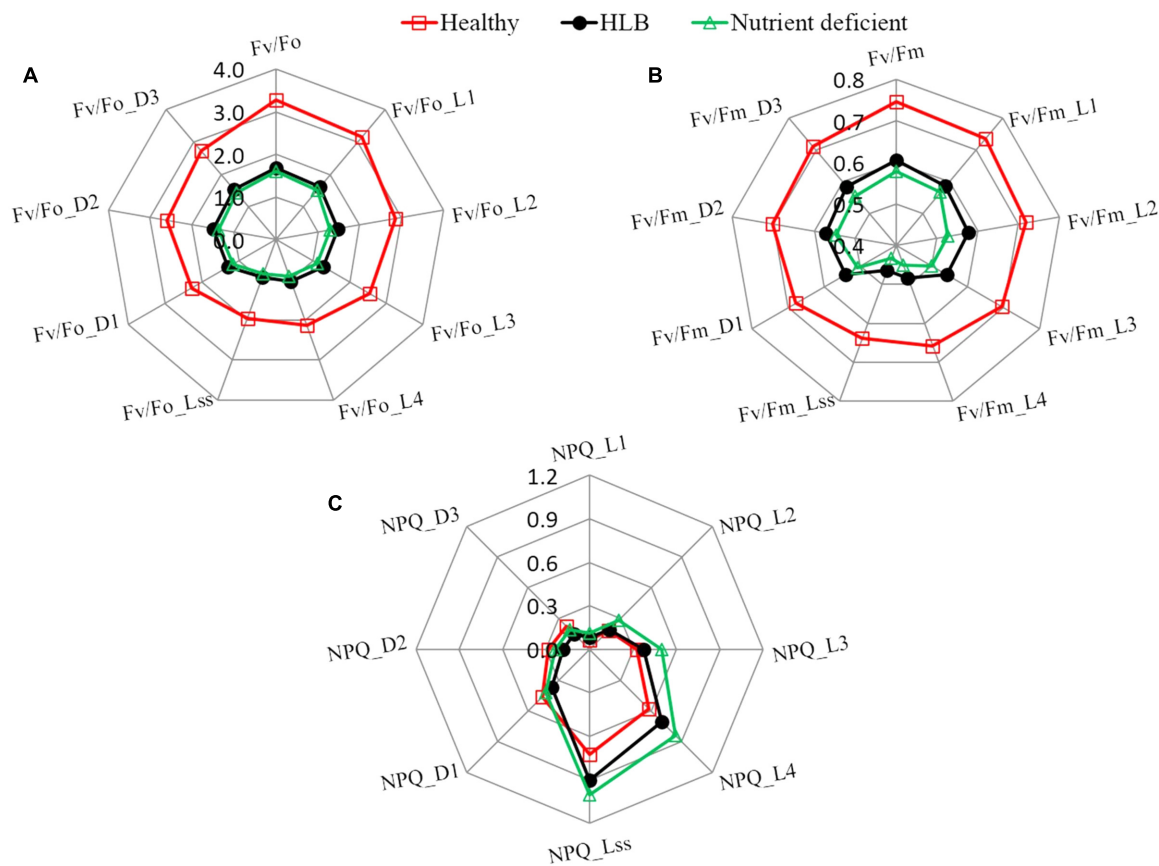


FIGURE 4 | Spider plots of selected fluorescence parameters including (A) F_v/F_o , (B) F_v/F_m , and (C) NPQ at different states of healthy, HLB infected and nutrient deficient citrus leaves measured during fluorescence quenching process.

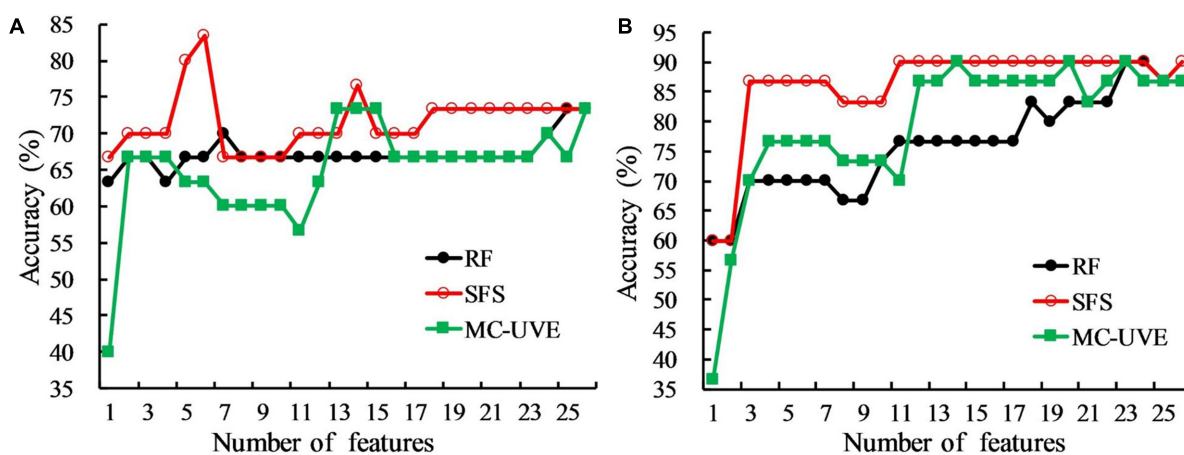


FIGURE 5 | Comparison of classification accuracies of random frog (RF), sequential forward selection (SFS), and Monte Carlo of uninformative variables elimination method (MC-UVE) using (A) partial least squares discriminant analysis (PLS-DA) and (B) support vector machine (SVM) classifiers for the classification of healthy, HLB infected, and nutrient deficient leaves.

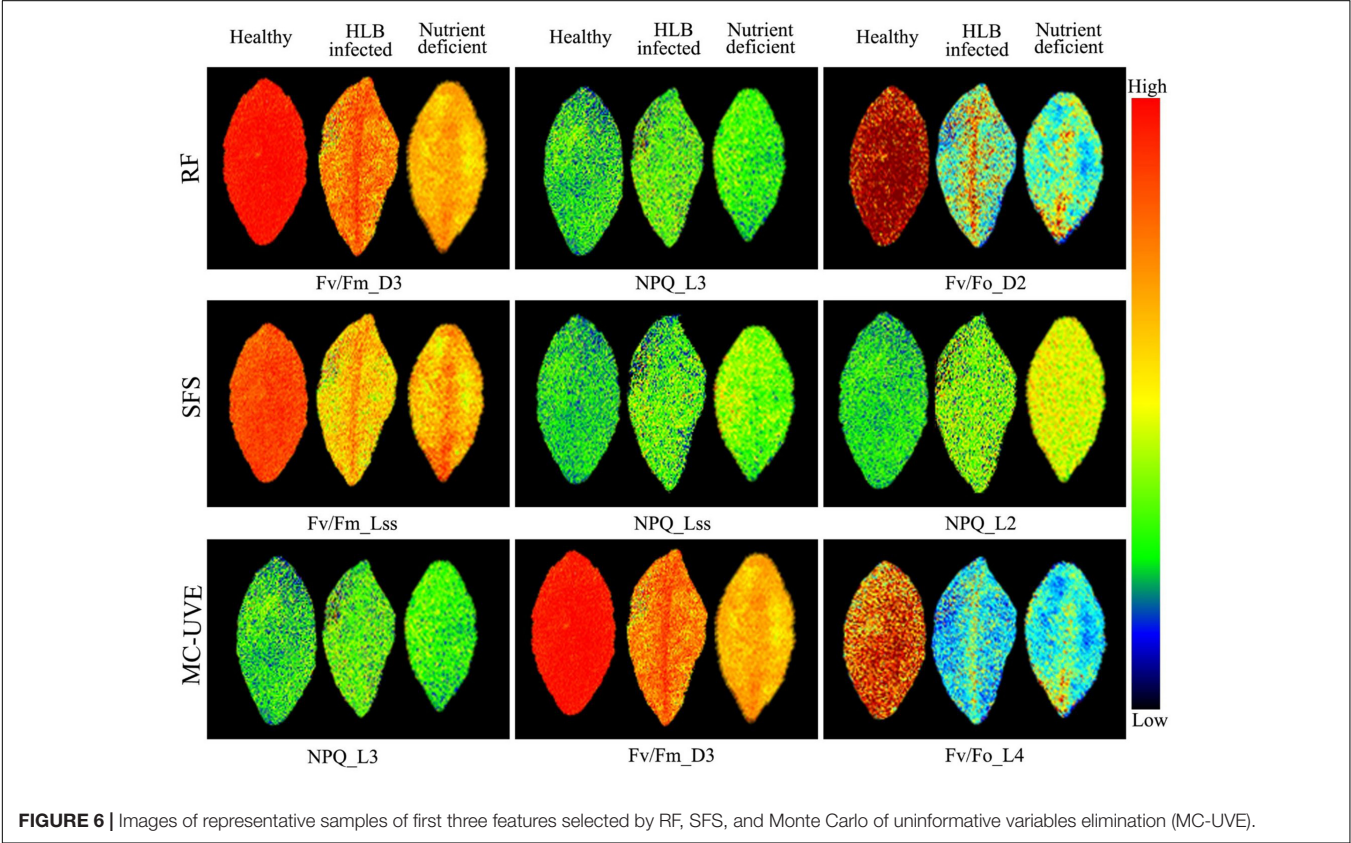
and nutrient deficient leaves were clearly observed. Some parameter images also showed the heterogeneity within the leaf due to the HLB infection or nutrient deficiency. This is agreement

with the reported study on a localized decrease in photosynthesis during the disease infection or other physiological disorder processes (Kenny et al., 2011; Calatayud et al., 2013).

TABLE 1 | Classification results based on chlorophyll fluorescence parameter analysis with the optimal features.

Feature selection	First three selected features	Classifier	Feature number	Healthy (%)	HLB infected (%)	Nutrient deficient (%)	Overall accuracy (%)
RF	<i>Fv/Fm_D3</i> , <i>NPQ_L3</i> , <i>Fv/Fo_D2</i>	PLS-DA	7	80	80	50	70
		SVM	23	80	90	100	90
SFS	<i>Fv/Fm_Lss</i> , <i>NPQ_Lss</i> , <i>NPQ_L2</i>	PLS-DA	6	90	80	80	83
		SVM	11	80	90	100	90
MC-UVE	<i>NPQ_L3</i> , <i>Fv/Fm_D3</i> , <i>Fv/Fo_L4</i>	PLS-DA	13	90	80	50	73
		SVM	23	80	90	100	90

RF, random frog; SFS, sequential forward selection; MC-UVE, Monte Carlo of uninformative variables elimination method; PLS-DA, partial least squares discriminant analysis; SVM, support vector machine.



The spatial heterogeneity of the HLB symptoms caused by the HLB bacterium presents more challenges for the disease identification. Hence, chlorophyll fluorescence image-based analysis was also performed to improve the HLB detection accuracy. The first six principal component (PC) images representing over 95% variables were used for image feature extraction. Features extracted by SIFT from PC images were used as the SVM classifier input. **Table 2** summarized the three-class classification results of using SVM based on image features and image features combined with the mean fluorescence parameters, respectively. The overall accuracy based on image features was 77%, which was not as good as that using mean fluorescence parameters with the classification accuracy of 90%. However, the combination of the image features and the mean fluorescence parameters significantly improved the classification performance

with the overall accuracy of 97%. This suggests that contrast information among these three classes was enhanced when using data fusion.

DISCUSSION

Various studies have reported that chlorophyll fluorescence is a promising technique for non-invasive measurement of PSII activities, and could achieve a rapid screening of photosynthetic processes related to the plant status (Ehlert and Hinch, 2008; Rolfe and Scholes, 2010; Ivanov and Bernards, 2016). Commonly used chlorophyll fluorescence parameters, such as *Fo*, *Fv/Fo*, *Fv/Fm*, and *NPQ_Lss*, obtained by exposing the leaf to a combination of darkness, actinic lights and a series of saturating

TABLE 2 | Support vector machine (SVM) classification results based on the principal component (PC) image features and the combination of image features and mean fluorescence parameters.

Sample status	PC images features			Combination of PC image features and mean fluorescence parameters		
	Healthy	HLB infected	Nutrient deficient	Healthy	HLB infected	Nutrient deficient
Healthy	8 (80%)	2	0	10 (100%)	0	0
HLB infected	0	7 (70%)	3	0	9 (90%)	1
Nutrient deficient	0	2	8 (80%)	0	0	10 (100%)
Overall accuracy		77%			97%	

flashes as shown in **Figure 1**, are considered as the most useful parameters to interpret energy dissipation of chlorophyll in the thylakoid membrane. In our study, these four fluorescence parameters measured in healthy, HLB infected, and nutrient deficient leaves showed the potential of differentiating the unhealthy leaves from the healthy ones except *Fo* (**Figure 3**). Large variations in *Fo* might be one of the reasons that it is difficult to provide a straightforward explanation about the impacts of HLB infection and nutrient deficiency on the PSII reaction center and energy transfer. The parameter *Fv/Fo* is considered as an indicator of the number and the size of active photosynthetic reaction centers (Dan et al., 2000). The dramatic declines of *Fv/Fo* in HLB infected and nutrient deficient leaves indicate the change in the rate of electron transport from PSII to the primary electron acceptors with the reduction of the number and the size, which have also been reported in different plants exposed to the disease and the environmental stresses (Zhou et al., 2009; Martinazzo et al., 2012; Janka et al., 2013). The decrease of *Fv/Fm* in HLB infected and nutrient deficient leaves could be a result of the major damage to the photosynthetic apparatus in response to the HLB infection and nutrient deficiency. The value of *NPQ_Lss* in HLB infected and nutrient deficient leaves might be related to the different degrees of tissue damage, which can also be reflected from the severity of the HLB and nutrient deficient symptoms. Reported studies have revealed that the increase of the *NPQ_Lss* depends on the stimulated electron flow as a protection mechanism (Tatagiba et al., 2016). What's more, it indicates the excess excitation energy dissipated as heat in order to reduce the photooxidative damage, which is also coupled to a gradual loss of chloroplastic pigments in response to the HLB infection and nutrient deficiency, eventually resulting in blotchy yellowing symptoms in leaves. Meanwhile, the significant differences observed in *Fv/Fm* and *NPQ_Lss* values between the HLB infected and nutrient deficient leaves demonstrated that it was possible to identify HLB infected leaves from the nutrient deficient ones, although the interpretation is not straightforward for a number of reasons.

Although previous research has reported a desired discrimination accuracy (92–95%) with adding fluorescence at 690 nm by using fluorescence imaging spectroscopy suggesting the significant contribution of emission from chlorophyll *a* molecule for HLB detection (Wetterich et al., 2016, 2017), more fundamental understanding about energy partition in PSII

in the HLB infected leaves should be gained. The chlorophyll fluorescence quenching process represents the complex dynamics of plant photosynthetic reaction during a transition from a dark-adapted to a light-adapted state (Govindjee, 1995). Some parameters measured at different time-courses during quenching process could offer photosynthetic signatures related to the plant disease (Berger et al., 2007). Further analysis of chlorophyll fluorescence parameters obtained during fluorescence quenching process as shown in **Figure 4** provided additional feature information related to the HLB disease, especially the consistent decrease in instantaneous non-photochemical quenching at several states in response to the HLB infection indicated a direct effect of HLB disease on PSII due to the irreversible damage of the leaf tissue.

Advanced machine learning and statistical analysis of finding the photosynthetic fingerprints of the HLB disease detection was shown to be superior to the commonly used fluorescence parameters. Designing an effective criterion to select an optimal subset of features is a challenging problem for data classification. For the mean fluorescence parameters analysis, SFS-SVM reached a best overall accuracy of 90% suggesting that features selected by SFS can better describe the pattern of the fluorescence data than the other two methods. Compared with the commonly used fluorescence parameters, feature selection methods provided a new insight to achieve a better interpretation of the original data by removing the redundant information and enhancing the most useful information related to the HLB infected characteristics. Although image features could provide the spatial information about the photochemical and non-photochemical components, the performance of the three-class classification based on image features extracted from PC images combined with the SVM classifier, was not better than those from mean fluorescence parameters (**Tables 1, 2**). This suggests that nutrient deficiency in leaves would induce similar image texture to HLB infected ones as shown in **Figure 6**. This is probably a reason why local image features had a poorer classification performance. The best classification accuracy of 97% using the combination of image features and mean fluorescence parameters revealed its capability to discriminate among healthy, HLB infected, and nutrient deficient leaves. In general, the results obtained in this study was comparable to those obtained using other spectroscopy and imaging techniques (Garcia-Ruiz et al., 2013; Sankaran et al., 2013; Wetterich et al., 2016, 2017). However, one should notice the differences

of experimental conditions and samples when comparing our classification accuracy with those from other techniques. It is true that the conventional method depending on the commonly used chlorophyll fluorescence parameters provided a better plant physiological interpretation, but it is challenging to perform a quantitative analysis of decision making for the HLB diagnosis in various conditions. Recent studies have approved that feature selection methods as well as data fusion have exploited a new way to interpret and extract the hidden information from the original data, such as image and spectral data (Li et al., 2015; Liu and Li, 2016; Fu et al., 2017). Better classification results by employing proper feature selection method and classifier (Table 2) demonstrated the capability of this data-driven strategy for the chlorophyll fluorescence image analysis.

In this study, we presented a comprehensive investigation on the photosynthetic characteristics of HLB infected citrus leaves and the possibility of the HLB disease detection using chlorophyll fluorescence imaging combined with feature selection. The combination of mean fluorescence parameters and images features significantly improved the classification performance with the accuracy of 97%, which indicated a better interpretation for the spatial heterogeneity of photochemical and non-photochemical components in HLB infected citrus leaves. The proposed approach cannot only achieve a better detection

accuracy of the HLB disease, but also be developed as a new means for the plant photosynthetic analysis.

AUTHOR CONTRIBUTIONS

HC and HW designed the experiment. HW, JY, MH, JL, and SH performed the experiment. HC and HW contributed to the data analysis and wrote the manuscript. HL and YH provided suggestions on the experiment design and discussion sections.

ACKNOWLEDGMENTS

This work was supported by the Key Research and Development Program from the Science Technology Department of Zhejiang Province (2015C02007) and China Agriculture Research System (CARS-27).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01509/full#supplementary-material>

REFERENCES

- Aliakbarzadeh, G., Parastar, H., and Sereshti, H. (2016). Classification of gas chromatographic fingerprints of saffron using partial least squares discriminant analysis together with different variable selection methods. *Chemometr. Intell. Lab. Syst.* 158, 165–173. doi: 10.1016/j.chemolab.2016.09.002
- Berger, S., Benediktyova, Z., Matous, K., Bonfig, K., Mueller, M. J., Nedbal, L., et al. (2007). Visualization of dynamics of plant-pathogen interaction by novel combination of chlorophyll fluorescence imaging and statistical analysis: differential effects of virulent and avirulent strains of *P. syringae* and of oxylipins on *A. thaliana*. *J. Exp. Bot.* 58, 797–806. doi: 10.1093/jxb/erl208
- Bové, J. M. (2006). Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *J. Plant Pathol.* 88, 7–37. doi: 10.4454/jpp.v88i1.828
- Bresson, J., Vasseur, F., Dauzat, M., Koch, G., Granier, C., and Vile, D. (2015). Quantifying spatial heterogeneity of chlorophyll fluorescence during plant growth and in response to water stress. *Plant Methods* 11:23. doi: 10.1186/s13007-015-0067-5
- Calatayud, Á., San Bautista, A., Pascual, B., Maroto, J. V., and López-Galarza, S. (2013). Use of chlorophyll fluorescence imaging as diagnostic technique to predict compatibility in melon graft. *Sci. Hortic.* 149, 13–18. doi: 10.1016/j.scienta.2012.04.019
- Cen, H., Lu, R., Zhu, Q., and Mendoza, F. (2016). Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biol. Technol.* 111, 352–361. doi: 10.1016/j.postharvbio.2015.09.027
- Chapelle, O., Vapnik, V., Bousquet, O., and Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Mach. Learn.* 46, 131–159. doi: 10.1023/A:1012450327387
- Chen, C. P., Frank, T. D., and Long, S. P. (2009). Is a short, sharp shock equivalent to long-term punishment? Contrasting the spatial pattern of acute and chronic ozone damage to soybean leaves via chlorophyll fluorescence imaging. *Plant Cell Environ.* 32, 327–335. doi: 10.1111/j.1365-3040.2008.01923.x
- Choi, H. G., Moon, B. Y., and Kang, N. J. (2016). Correlation between strawberry (*Fragaria ananassa* Duch.) productivity and photosynthesis-related parameters under various growth conditions. *Front. Plant Sci.* 7:1607. doi: 10.3389/fpls.2016.01607
- Dan, T. V., Krishnaraj, S., and Saxena, P. K. (2000). Metal tolerance of scented geranium (*Pelargonium* sp. 'Frensham'): effects of cadmium and nickel on chlorophyll fluorescence kinetics. *Int. J. Phytoremediat.* 2, 91–104. doi: 10.1080/15226510008500032
- Ehlert, B., and Hinch, D. K. (2008). Chlorophyll fluorescence imaging accurately quantifies freezing damage and cold acclimation responses in *Arabidopsis* leaves. *Plant Methods* 4:12. doi: 10.1186/1746-4844-4-12
- Fu, Y., Zhao, C., Wang, J., Jia, X., Yang, G., Song, X., et al. (2017). An improved combination of spectral and spatial features for vegetation classification in hyperspectral images. *Remote Sens.* 9:261. doi: 10.3390/rs9030261
- García-Ruiz, F., Sankaran, S., Maja, J. M., Lee, W. S., Rasmussen, J., and Ehsani, R. (2013). Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees. *Comput. Electron. Agric.* 91, 106–115. doi: 10.1016/j.compag.2012.12.002
- Geng, X. (2014). An efficient unsupervised band selection method based on an autocorrelation matrix for a hyperspectral image. *Int. J. Remote Sens.* 35, 7458–7476. doi: 10.1080/01431161.2014.968686
- Gnyp, M. L., Miao, Y., Yuan, F., Ustin, S. L., Yu, K., Yao, Y., et al. (2014). Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages. *Field Crop Res.* 155, 42–55. doi: 10.1016/j.fcr.2013.09.023
- Gouda, K. A., Baranwal, V. K., and Ahlawat, Y. S. (2006). Simplified DNA extraction and improved PCR based detection of greening bacterium in citrus. *J. Plant Biochem. Biotechnol.* 15, 117–121. doi: 10.1007/BF03321914
- Govindjee, R. (1995). Sixty-three years since Kautsky: chlorophyll *a* fluorescence. *Aust. J. Plant Physiol.* 22, 131–160. doi: 10.1071/PP9950131
- Han, Q., Wu, H., Cai, C., Xu, L., and Yu, R. (2008). An ensemble of Monte Carlo uninformative variable elimination for wavelength selection. *Anal. Chim. Acta* 612, 121–125. doi: 10.1016/j.aca.2008.02.032
- Harbinson, J. (2013). Improving the accuracy of chlorophyll fluorescence measurements. *Plant Cell Environ.* 36, 1751–1754. doi: 10.1111/pce.12111
- Inoue, Y., Sakaiya, E., Zhu, Y., and Takahashi, W. (2012). Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sens. Environ.* 126, 210–221. doi: 10.1016/j.rse.2012.08.026
- Ivanov, D. A., and Bernards, M. A. (2016). Chlorophyll fluorescence imaging as a tool to monitor the progress of a root pathogen in a perennial plant. *Planta* 243, 263–279. doi: 10.1007/s00425-015-2427-9

- Janka, E., Körner, O., Rosenqvist, E., and Ottosen, C. O. (2013). High temperature stress monitoring and detection using chlorophyll a fluorescence and infrared thermography in chrysanthemum (*Dendranthema grandiflora*). *Plant Physiol. Biochem.* 67, 87–94. doi: 10.1016/j.plaphy.2013.02.025
- Kenny, P., Deák, Z., Csász, M., and Purnhauser, L. (2011). Characterization and early detection of tan spot disease in wheat in vivo with chlorophyll fluorescence imaging. *Acta Biol. Szeged.* 55, 87–90.
- Kusnier, K., and Korsath, A. (2015). Simultaneous identification of spring wheat nitrogen and water status using visible and near infrared spectra and powered partial least squares regression. *Comput. Electron. Agric.* 117, 200–213. doi: 10.1016/j.compag.2015.08.001
- Lee, J. A., Halbert, S. E., Dawson, W. O., Robertson, C. J., Keesling, J. E., and Singer, B. H. (2015). Asymptomatic spread of huanglongbing and implications for disease control. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7605–7610. doi: 10.1073/pnas.1508253112
- Lei, R., Jiang, H., Hu, F., Yan, J., and Zhu, S. (2017). Chlorophyll fluorescence lifetime imaging provides new insight into the chlorosis induced by plant virus infection. *Plant Cell Rep.* 36, 327–341. doi: 10.1007/s00299-016-2083-y
- Li, H., Xu, Q., and Liang, Y. (2012). Random frog: an efficient reversible jump Markov Chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. *Anal. Chim. Acta* 740, 20–26. doi: 10.1016/j.aca.2012.06.031
- Li, X., Lee, W. S., Li, M., Ehsani, R., Mishra, A. R., Yang, C., et al. (2012). Spectral difference analysis and airborne imaging classification for citrus greening infected trees. *Comput. Electron. Agric.* 83, 32–46. doi: 10.1016/j.compag.2012.01.010
- Li, X., Sun, C., Zhou, B., and He, Y. (2015). Determination of hemicellulose, cellulose and lignin in moso bamboo by near infrared spectroscopy. *Sci. Rep.* 5, 17210–17220. doi: 10.1038/srep17210
- Liu, D., and Li, J. (2016). Data field modeling and spectral-spatial feature fusion for hyperspectral data classification. *Sensors* 16:2146. doi: 10.3390/s16122146
- Macho, S., Rius, A., Callao, M. P., and Larrech, M. S. (2001). Monitoring ethylene content in heterophase copolymers by near-infrared spectroscopy: standardisation of the calibration model. *Anal. Chim. Acta* 445, 213–220. doi: 10.1016/S0003-2670(01)01281-8
- Martinazzo, E. G., Ramm, A., and Bacarin, M. A. (2012). The chlorophyll a fluorescence as an indicator of the temperature stress in the leaves of *Prunus persica*. *Braz. J. Plant Physiol.* 24, 237–246. doi: 10.1590/S1677-04202013005000001
- Mehta, P., Jajoo, A., Mathur, S., and Bharti, S. (2010). Chlorophyll a fluorescence study revealing effects of high salt stress on Photosystem II in wheat leaves. *Plant Physiol. Biochem.* 48, 16–20. doi: 10.1016/j.plaphy.2009.10.006
- Mishra, A. R., Karimi, D., Ehsani, R., and Lee, W. S. (2012). Identification of citrus greening HLB using a VIS-NIR spectroscopy technique. *Trans. ASABE* 55, 711–720.
- Mishra, K. B., Mishra, A., Novotna, K., Rapantova, B., Hodanova, P., Urban, O., et al. (2016). Chlorophyll a fluorescence, under half of the adaptive growth-irradiance, for high-throughput sensing of leaf-water deficit in *Arabidopsis thaliana* accessions. *Plant Methods* 12:46. doi: 10.1186/s13007-016-0145-3
- Murchie, E. H., and Lawson, T. (2013). Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *J. Exp. Bot.* 64, 3983–3998. doi: 10.1093/jxb/ert08
- Pereira, F. M., Milori, D. M. B. P., Pereira-Filho, E. R., Venâncio, A. L., Russo, M. D. S. T., Cardinali, M. C. D. B., et al. (2011). Laser-induced fluorescence imaging method to monitor citrus greening disease. *Comput. Electron. Agric.* 79, 90–93. doi: 10.1016/j.compag.2011.08.002
- Pineda, M., Gaspar, L., Morales, F., Szigeti, Z., and Baron, M. (2008a). Multicolor fluorescence imaging of leaves—a useful tool for visualizing systemic viral infections in plants. *Photochem. Photobiol.* 84, 1048–1060. doi: 10.1111/j.1751-1097.2008.00357.x
- Pineda, M., Olejnickova, J., Csefalvay, L., and Baron, M. (2011). Tracking viral movement in plants by means of chlorophyll fluorescence imaging. *J. Plant Physiol.* 168, 2035–2040. doi: 10.1016/j.jplph.2011.06.013
- Pineda, M., Soukupová, J., Matouš, K., Nedbal, L., and Barón, M. (2008b). Conventional and combinatorial chlorophyll fluorescence imaging of tobamovirus-infected plants. *Photosynthetica* 46, 441–451. doi: 10.1007/s11099-008-0076-y
- Pinto, F., Damm, A., Schickling, A., Panigada, C., Cogliati, S., Muller-Linow, M., et al. (2016). Sun-induced chlorophyll fluorescence from high-resolution imaging spectroscopy data to quantify spatio-temporal patterns of photosynthetic function in crop canopies. *Plant Cell Environ.* 39, 1500–1512. doi: 10.1111/pce.12710
- Pourreza, A., Lee, W. S., Etxeberria, E., and Banerjee, A. (2015). An evaluation of a vision-based sensor performance in huanglongbing disease identification. *Biosyst. Eng.* 130, 13–22. doi: 10.1016/j.biosystemseng.2014.11.013
- Pourreza, A., Lee, W. S., Raveh, E., Ehsani, R., and Etxeberria, E. (2014). Citrus Huanglongbing detection using narrow-band imaging and polarized illumination. *Trans. ASABE* 57, 259–272.
- Qin, J., Burks, T. F., Zhao, X., Niphadkar, N., and Ritenour, M. A. (2011). Multispectral detection of citrus canker using hyperspectral band selection. *Trans. ASABE* 54, 2331–2341.
- Rolfe, S. A., and Scholes, J. D. (2010). Chlorophyll fluorescence imaging of plant-pathogen interactions. *Protoplasma* 247, 163–175. doi: 10.1007/s00709-010-0203-z
- Sankaran, S., and Ehsani, R. (2013). Detection of huanglongbing-infected citrus leaves using statistical models with a fluorescence sensor. *Appl. Spectrosc.* 67, 463–469. doi: 10.1366/12-06790
- Sankaran, S., Maja, J. M., Buchanon, S., and Ehsani, R. (2013). Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques. *Sensors* 13, 2117–2130. doi: 10.3390/s130202117
- Song, X., Zhou, G., Xu, Z., Lv, X., and Wang, Y. (2016). Detection of photosynthetic performance of *Stipa bungeana* seedlings under climatic change using chlorophyll fluorescence imaging. *Front. Plant Sci.* 6:1254. doi: 10.3389/fpls.2015.01254
- Spoustová, P., Synková, H., Valcke, R., and Čerovská, N. (2013). Chlorophyll a fluorescence as a tool for a study of the *Potato virus Y* effects on photosynthesis of nontransgenic and transgenic *pssu-ipt* tobacco. *Photosynthetica* 51, 191–201. doi: 10.1007/s11099-013-0023-4
- Sui, X., Shan, N., Hu, L., Zhang, C., Yu, C., Ren, H., et al. (2017). The complex character of photosynthesis in cucumber fruit. *J. Exp. Bot.* 68, 1625–1637. doi: 10.1093/jxb/erx034
- Tatagiba, S. D., DaMatta, F. M., and Rodrigues, F. A. (2016). Silicon partially preserves the photosynthetic performance of rice plants infected by *Monographella albescens*. *Ann. Appl. Biol.* 168, 111–121. doi: 10.1111/aab.12246
- Wetterich, C. B., Felipe, D. O. N. R., Belasque, J., Ehsani, R., and Marcassa, L. G. (2017). Detection of huanglongbing in Florida using fluorescence imaging spectroscopy and machine-learning methods. *Appl. Opt.* 56, 15–23. doi: 10.1364/AO.56.000015
- Wetterich, C. B., Felipe, D. O. N. R., Belasque, J., and Marcassa, L. G. (2016). Detection of citrus canker and huanglongbing using fluorescence imaging spectroscopy and support vector machine technique. *Appl. Opt.* 55, 400–407. doi: 10.1364/AO.55.000400
- Wetterich, C. B., Kumar, R., Sankaran, S., Junior, J. B., Ehsani, R., and Marcassa, L. G. (2013). A comparative study on application of computer vision and fluorescence imaging spectroscopy for detection of Huanglongbing citrus disease in the USA and Brazil. *J. Spectrosc.* 2013, 841738. doi: 10.1155/2013/841738
- Yoshioka, Y., Nakayama, M., Noguchi, Y., and Horie, H. (2013). Use of image analysis to estimate anthocyanin and UV-excited fluorescent phenolic compound levels in strawberry fruit. *Breed Sci.* 63, 211–217. doi: 10.1270/jsbbs.63.211
- Zhou, J., Yang, L., Hao, F., and You, Y. (2009). Photosynthesis and chlorophyll fluorescence of *Magnolia grandiflora* seedlings under low temperature stress. *Acta Bot. Boreali Occident. Sin.* 29, 136–142.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Cen, Weng, Yao, He, Lv, Hua, Li and He. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Can Leaf Water Content Be Estimated Using Multispectral Terrestrial Laser Scanning? A Case Study With Norway Spruce Seedlings

Samuli Junttila^{1,2*}, Junko Sugano¹, Mikko Vastaranta^{1,2,3}, Riikka Linnakoski^{1,4}, Harri Kaartinen^{5,6}, Antero Kukko^{2,5}, Markus Holopainen^{1,2}, Hannu Hyypä^{2,7} and Juha Hyypä^{2,5}

OPEN ACCESS

Edited by:

Alexander Bucksch,
University of Georgia, United States

Reviewed by:

Asish Kumar Parida,
CSIR-Central Salt & Marine Chemicals
Research Institute, India
Ioannis Konstantinos Christodoulakis,
National and Kapodistrian University
of Athens, Greece
András Zlinszky,
Institute of Ecology Research Center
(MTA), Hungary

*Correspondence:

Samuli Junttila
samuli.junttila@helsinki.fi

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 30 September 2017

Accepted: 21 February 2018

Published: 08 March 2018

Citation:

Junttila S, Sugano J, Vastaranta M,
Linnakoski R, Kaartinen H, Kukko A,
Holopainen M, Hyypä H and
Hyypä J (2018) Can Leaf Water
Content Be Estimated Using
Multispectral Terrestrial Laser
Scanning? A Case Study With Norway
Spruce Seedlings.
Front. Plant Sci. 9:299.
doi: 10.3389/fpls.2018.00299

¹ Department of Forest Sciences, University of Helsinki, Helsinki, Finland, ² Centre of Excellence in Laser Scanning Research, Finnish Geospatial Research Institute (FGI), Masala, Finland, ³ School of Forest Sciences, University of Eastern Finland, Joensuu, Finland, ⁴ Natural Resources Institute Finland (Luke), Helsinki, Finland, ⁵ Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute (FGI), Masala, Finland, ⁶ Department of Geography and Geology, University of Turku, Turku, Finland, ⁷ Department of Built Environment, Aalto University, Aalto, Finland

Changing climate is increasing the amount and intensity of forest stress agents, such as drought, pest insects, and pathogens. Leaf water content, measured here in terms of equivalent water thickness (EWT), is an early indicator of tree stress that provides timely information about the health status of forests. Multispectral terrestrial laser scanning (MS-TLS) measures target geometry and reflectance simultaneously, providing spatially explicit reflectance information at several wavelengths. EWT and leaf internal structure affect leaf reflectance in the shortwave infrared region that can be used to predict EWT with MS-TLS. A second wavelength that is sensitive to leaf internal structure but not affected by EWT can be used to normalize leaf internal effects on the shortwave infrared region and improve the prediction of EWT. Here we investigated the relationship between EWT and laser intensity features using multisensor MS-TLS at 690, 905, and 1,550 nm wavelengths with both drought-treated and *Endoconidiophora polonica* inoculated Norway spruce seedlings to better understand how MS-TLS measurements can explain variation in EWT. In our study, a normalized ratio of two wavelengths at 905 and 1,550 nm and length of seedling explained 91% of the variation (R^2) in EWT as the respective prediction accuracy for EWT was 0.003 g/cm² in greenhouse conditions. The relation between EWT and the normalized ratio of 905 and 1,550 nm wavelengths did not seem sensitive to a decreased point density of the MS-TLS data. Based on our results, different EWTs in Norway spruce seedlings show different spectral responses when measured using MS-TLS. These results can be further used when developing EWT monitoring for improving forest health assessments.

Keywords: terrestrial laser scanning, tree health, drought stress, multispectral laser scanning, leaf water content, forest damage, *Endoconidiophora polonica*

INTRODUCTION

Measuring tree health is an increasingly important topic as the world's climate is warming and the human population is growing, putting the environment and forests under increasing stress (Williams et al., 2013). Stress and different disturbance events caused by, for example, insects are natural processes in forest ecosystems that can help maintain healthy and heterogeneous forests (Raffa et al., 2009). However, forests are facing new stresses and disturbance events such as drought, invasive pests, and air pollution, which are more intense than those they have previously adapted to (Trumbore et al., 2015). Forests provide many ecosystem services, such as carbon storage, recreational activities, timber, and non-timber forest products, which are jeopardized by declining forest health, increased forest mortality, and increased susceptibility to forest fires (Hanewinkel et al., 2013; McDowell and Allen, 2015; Morris et al., 2016). Managing declining forests efficiently in the face of climate change requires more information on the condition of forests to be able to monitor and control forest damage and fires (Allen et al., 2010); thus, new methods for mapping and monitoring forest health are needed.

Existing forest health mapping methods can be divided according to spatial, spectral, and temporal resolutions (Senf et al., 2017). Large-scale mapping of insect infestation has been conducted using single and multi-date Landsat data at 30 m resolution (Goodwin et al., 2008; Meddens et al., 2013). These methods are capable of capturing forest health information on a large scale and are suitable for countries where extensive forest management is applied, but spatially, more detailed approaches are necessary for capturing small-scale dynamics of forest decline in smaller countries with a more intensive approach to forest management. Spatially more accurate hyperspectral remote sensing data with 5 m resolution have been used to investigate the mapping of bark beetle induced tree mortality resulting in high overall accuracy (84–96%) in classifying dead trees (Fassnacht et al., 2014). Multispectral satellite imagery with 4 m resolution has been used to detect the red stage in mountain pine beetle infestation, resulting in high overall accuracy of 71–92% (White et al., 2005). However, none of the methods studied so far have been able to accurately detect early stages of bark beetle infestation and early tree stress when the trees have exhibited few visual symptoms (Fassnacht et al., 2014).

Leaf water content, typically measured as the weight of water per leaf area or equivalent water thickness (EWT), is an indicator of tree health that is affected by several tree stressors (Carter, 1993; Chaerle and Van Der Straeten, 2000). Drought, pest insects, and pathogens affect EWT, often before other symptoms appear, such as discoloration or defoliation; thus, detecting changes in EWT provides an early-warning signal of tree stress (Skakun et al., 2003; Zarco-Tejada et al., 2003). Leaf and canopy water content derived from Landsat data have been used as a proxy to detect large-scale insect infestation (Skakun et al., 2003; Wulder et al., 2006). Early detection of tree stress is crucial in providing timely information about the condition of forests, mitigating damage, and preventing the spread of the stressor (Wermelinger, 2004). For example, the European spruce bark

beetle (*Ips typographus* L.) is a forest insect pest whose main host is Norway spruce trees. The bark beetle and a fungal pathogen have a mutual relationship causing disruptions in the flow of water and nutrients in the phloem and sapwood of the trees, resulting in decreased water content in the canopy and the sapwood (Horntvedt et al., 1983). Recently infested trees do not exhibit visual symptoms (green attack), but a reddish-brown coloring of tree crowns appears when the beetles have already left the host tree (Lausch et al., 2013). Thus, early detection of tree stress would greatly assist in mitigating and managing the damage caused by *I. typographus*. Efforts have been made to solve this issue but none have succeeded with reasonable accuracy thus far (Lausch et al., 2013; Fassnacht et al., 2014).

Improving reference data for assessing forest health has been identified as a key research area in need of development (Senf et al., 2017). Because field reference data or the ground truth is the gold standard of modeling, any errors in the data can lead to significant bias and errors in modeling. The high costs of collecting field data usually lead to a compromise between the number of plots and the spatial extent of the investigation. National forest inventories may collect forest health data, but forest inventory data are not available in all countries (Gschwantner et al., 2016). Visual assessment of forest health in field data collection is prone to error and bias due to the subjective nature of the estimation, and it requires substantial expertise. The early stages of tree stress can be visually difficult to determine because trees may not exhibit visual symptoms at this stage. Thus, objective and reliable methods for collecting forest health related reference data in the field are needed.

Equivalent water thickness is closely related to another metric for calculating vegetation water content—fuel moisture content (FMC)—that is more frequently used in forest fire danger literature (Chuvieco et al., 2004). Fuel moisture content is expressed as the ratio between the weight of water and sample dry weight and is an important parameter in fire ignition (Dimitrakopoulos and Papaioannou, 2001). A multitemporal analysis of NOAA-AVHRR images has been used to derive FMC for Mediterranean grasslands and shrub species with R^2 -values over 0.8 (Chuvieco et al., 2004). Generally, studies aiming to estimate FMC use spatially coarse resolution satellite data that can provide frequent imagery, important in forest fire danger estimation (Sow et al., 2013).

Multispectral terrestrial laser scanning (MS-TLS) measures the range and reflectance of the target at several wavelengths simultaneously, providing three-dimensional estimates of reflectance. Several MS-TLS instruments have been developed recently mainly for laboratory use (Douglas et al., 2012; Hakala et al., 2012; Wallace et al., 2012; Wei et al., 2012; Niu et al., 2015; Li et al., 2016). These instruments utilize a supercontinuum laser source that is powerful enough to not be eye-safe, limiting their operation in the field, with the exception of a terrestrial laser scanner operating at two wavelengths, which is also portable (Douglas et al., 2012). Recently, a commercial multispectral laser scanner—the Optech Titan (Teledyne Optech, Canada)—with three channels has become available for airborne measurements (Matikainen et al., 2016). The development of

multispectral laser scanning instruments widens the possibilities of using laser scanning data as more information about the target's reflectance becomes available in a three-dimensional format.

Alterations in EWT affect leaf reflectance in the shortwave infrared region (Ceccato et al., 2001), which can be measured using MS-TLS. A number of studies have investigated the use of MS-TLS and terrestrial laser scanning (TLS) in the estimation of EWT. Gaulton et al. (2013) studied the potential of estimating leaf EWT with a dual-wavelength terrestrial laser scanner at 1,063 and 1,545 nm wavelengths, resulting in a significant correlation ($R^2 = 0.8$, root mean square error [RMSE] = 0.0069 g/cm²) between leaf EWT and a normalized ratio of the two wavelengths, but with a low number of samples and species. The estimation of leaf EWT has been recently studied using MS-TLS at 690 and 1,550 nm wavelengths with single leaves, resulting in a strong correlation ($R^2 = 0.93$, RMSE = 0.004 g/cm²) between EWT and a normalized ratio of the two wavelengths, but the selection of wavelengths was concluded to be suboptimal (Junttila et al., 2016). Zhu et al. (2015) showed that TLS intensity data at 1,550 nm were able to explain 76% of the variation in EWT at leaf level after radiometric correction of the TLS intensity data. Canopy EWT has been estimated with full-waveform single wavelength TLS at 1,550 nm wavelength, resulting in a significant correlation between canopy EWT and backscatter coefficient ($R^2 = 0.66$, RMSE = 0.001 g/cm²) (Zhu et al., 2017). However, single wavelength data seem to be sensitive to confounding factors, such as incidence angle, requiring complicated correction of the intensity data, and they are also affected by structural variables of the leaf other than EWT (Zhu et al., 2015, 2017). The requirement of an incidence angle correction limits the utilization of the single wavelength method to deciduous species with distinguishable leaves. Coniferous forests comprise a large part of the world's forest biomes, but thus far only Junttila et al. (2016) have looked into the estimation of EWT with coniferous species using MS-TLS. The use of spectral ratios calculated from dual-wavelength intensity data has been shown to be insensitive to incidence angle effects thus enabling higher accuracy in the estimation of EWT (Hancock et al., 2017). A second wavelength in the near infrared region coupled with a wavelength in the shortwave infrared region is also capable of normalizing leaf structural effects on the estimation of EWT (Ceccato et al., 2001).

The aim of this study was to investigate the capability of multisensor MS-TLS in monitoring leaf EWT, with Norway spruce (*Picea abies* L.) seedlings. First, the dependencies between EWT and laser intensity features at different wavelengths (and calculated spectral ratios) of segmented point clouds are examined to find explanatory variables for predicting EWT. Secondly, we compare the combination of 690, 905, and 1,550 nm wavelengths in predicting EWT using different sensors. Thirdly, the importance of point density in estimating EWT is evaluated using random sampling of the point clouds to decrease the point density. This research contributes to the first steps in developing a method for the accurate estimation of leaf EWT in forests using MS-TLS.

MATERIALS AND METHODS

Experiment Design

Two-year-old commercial Norway spruce seedlings ($n = 145$) were grown for 12 weeks in a greenhouse between May and August 2016. The seedlings were subjected to different treatments to induce drought and variation in EWT. The seedlings were divided into five groups for different treatments. Three of the groups received 75% (28 ml), 50% (20 ml), and 25% (12 ml) of "normal" watering (watered as required to maintain moist soil) three times a week (groups D75, D50, and D25, respectively) (Table 1). The fourth group was grown with a sufficient amount of water (36 ml) for 10 weeks until irrigation was completely stopped, and the seedlings then went without water for 3 weeks (group D_total). The fifth group of seedlings was inoculated with *Endoconidiophora polonica* (isolate CBS 142283), a fungal pathogen associated with the European spruce bark beetle (*I. typographus*) that disturbs the flow of water and nutrients in the phloem and sapwood (group F). The amount of water given to the seedlings in the drought treatment groups (D75, D50, and D25) was adjusted during the experiment because the seedlings showed resistance to drought (Table 2). After 34 days from the start of the experiment, irrigation was reduced to twice a week; after 46 days, the amount of water given was cut to half; and finally after 56 days from the start of the experiment the irrigation was reduced to once a week. During the 12 weeks, 5–16 seedlings were randomly collected for TLS measurements at eight time intervals. Two or three seedlings were collected from each of the drought treatment groups (D75, D50, and D25) and eight seedlings from the fungal treatment group (F) for each MS-TLS measurement. Seedlings that showed severe symptoms during the final stages of the experiment, and which were unlikely to survive until the next measurement time were prioritized to avoid the loss of research material, resulting in uneven sample numbers toward the end of the experiment. Details of the measurements can be found in Table 3. The mean length of the seedlings was 30.3 cm (std \pm 5.34 cm) during the experiment.

Terrestrial Laser Scanners

The seedlings were scanned consecutively with three different terrestrial laser scanners from the same position from a distance of 5.2 m inside the greenhouse. The terrestrial laser scanners used were a Leica HDS6100 (Leica Geosystems AG, Heerbrugg, Switzerland), a FARO S120 (FARO Europe GmbH & Co. KG, Korntal-Münchingen, Germany), and a FARO X330, utilizing wavelengths of 690, 905, and 1,550 nm, respectively. The technical specifications of the terrestrial laser scanners can be found in Table 4. The varying beam diameter at exit and beam divergence resulted in different spot sizes at target distance, which were 4.14, 3.99, and 3.24 mm for the Leica, the S120, and the X330, respectively. The scanning was done using the highest resolution available for each instrument, resulting in a point spacing of ~ 1 mm; thus, there was significant overlap between consecutive laser measurements.

Three white spheres were used as common targets to register the scans to the same coordinate system as each other, facilitating

the segmentation of point clouds in the processing of the data. A four-grade Spectralon (Labsphere, North Sutton, NH, USA) panel was used as a reference target for intensity normalization. The size of the panel was 460 mm x 460 mm consisting of four different reflectance panels (nominal reflectances of 99, 50, 25, and 12%), each sized 115 × 460 mm. Approximately 65,000 points were received from each of the panels.

Ecophysiological Measurements

After scanning, the seedlings were transported to the laboratory for EWT measurements (Table 3). A sample of 30–40 needles was randomly collected from each seedling, constituting a single measurement of EWT. As the seedlings were growing during the experiment, only previous year needles were sampled to avoid the influence of needle development on the EWT measurements. The needles were weighed (with a precision of 0.0001 g) to measure fresh weight, scanned with an Epson V370 Photo Scanner (Epson America, Inc., CA, USA) at 800 dpi resolution to measure leaf area, and dried in an oven at 60°C for 48 h to measure dry weight. The images were analyzed with the open source software EasyLeafArea (Easlon and Bloom, 2014) to segment the needles and calculate leaf area based on the number of segmented pixels. The EWT was then calculated according to Danson et al. (1992):

$$\frac{FW - DW}{A} \left(\frac{g}{cm^2} \right), \quad (1)$$

where FW is the fresh weight of the needles (g), DW is the weight of the dried needles (g), and A is the surface area of the fresh needles (cm²).

Terrestrial Laser Scanning Data Processing

The TLS scans were registered to a common coordinate system using the Z + F LaserControl software package (Zoller + Fröhlich GmbH, Wangen im Allgäu, Germany), resulting in a mean accuracy of 2.7 mm between targets. The registration

process was based on Helmert transformation and three sphere-shaped targets with known locations (Watson, 2006). The registration of the scans was visually inspected after the registration process to ensure the alignment of the point clouds at the range of the seedlings. Because the terrestrial laser scanners utilize a phase shift measurement technique to determine the range to the target, a large number of ghost points were produced in the scans (Balduzzi et al., 2011). To reduce the number of ghost points, the point clouds were filtered with a statistical outlier algorithm based on the distance between points by using the CloudCompare v.2.8.1 (Girardeau-Montaut, 2011) software. The algorithm calculates the maximum distance for a point to be included using the following equation:

$$MaxD = MeanD + nSigma * std, \quad (2)$$

where *MaxD* is the maximum distance for a point to be included, *MeanD* is the mean distance of the neighboring points, *nSigma* is a standard deviation multiplier threshold, and *std* is the standard deviation of the distance of the neighboring points. Thus, the algorithm requires two parameters: the number of neighbors and *nSigma*. The parameters were determined through iteration and visual estimation of the result. Fifty neighbors were used for every scan and the *nSigma* was set to 1.0 and 0.65 for the Leica and FARO scanners respectively.

Then, cloud-to-cloud distances were calculated for the point clouds in the CloudCompare software package where the FARO X330 was used as a reference cloud for calculating the distances because of its expected sensitivity to EWT (Junttila et al., 2016). The tool calculates the distance of each point to the nearest point in the reference cloud. A maximum distance of 2 mm (which is also the ranging accuracy of the terrestrial laser scanners) was used to filter the point clouds and match the geometry of the point clouds produced by different terrestrial laser scanners with each other.

Points from each seedling were then manually clipped from the georeferenced point clouds for further processing. The point clouds from each seedling were further manually segmented into stem and leaf points to evaluate their effect on the estimation of EWT.

Because the point clouds were produced using a very high-resolution scanning setting, the point clouds were randomly sampled to reduce the amount of points in each point cloud and investigate how the point density affects the estimation of EWT. The point clouds were randomly sampled to four different point cloud sizes: 2,000, 1,000, 500, and 250 points per seedling. The mean number of points for each seedling was 14,702 in the unsegmented point clouds; thus, the amount of points was greatly reduced.

Intensity Calibration

Terrestrial laser scanners are primarily designed for range measuring; thus, they are not optimized for reflectivity measurements. Some terrestrial laser scanners have been shown to have a nonlinear intensity scale (Kaasalainen et al., 2011); thus,

TABLE 1 | Treatment groups and related statistics.

Group name	Irrigation amount per seedling (ml)	Number of seedlings
D75	28	25
D50	20	23
D25	12	23
D_total	36 (until stopping)	23
F	36	51

TABLE 2 | Irrigation of the drought experiment groups.

Dates	Irrigation amount per seedling (ml) for each group			Irrigation times per week
	D75	D50	D25	
June 19–June 23	28	20	12	3
June 24–July 5	28	20	12	2
July 6–July 15	14	10	6	2
July 16–August 18	14	10	6	1

TABLE 3 | EWT for each measurement and treatment group, and their respective statistics.

Measurement date	Days from starting the experiment	Group name	Measurements	Mean EWT (g/cm ²)	Min EWT (g/cm ²)	Max EWT (g/cm ²)	Standard deviation of EWT (g/cm ²)
EWT							
May 19	0	D75	3	0.035	0.033	0.038	0.0032
May 19	0	D50	2	0.035	0.031	0.039	0.0061
May 19	0	D25	3	0.031	0.027	0.038	0.0055
May 19	0	F	8	0.033	0.027	0.040	0.0040
June 3	15	D75	2	0.037	0.034	0.039	0.0035
June 3	15	D50	3	0.026	0.031	0.038	0.0060
June 3	15	D25	3	0.029	0.028	0.030	0.0007
June 3	15	F	8	0.016	0.006	0.032	0.0092
June 9	21	D75	2	0.032	0.032	0.032	0.0004
June 9	21	D50	3	0.036	0.029	0.043	0.0072
June 9	21	D25	2	0.032	0.030	0.033	0.0020
June 9	21	F	8	0.016	0.004	0.031	0.0110
June 17	29	D75	2	0.032	0.029	0.035	0.0037
June 17	29	D50	2	0.029	0.029	0.030	0.0004
June 17	29	D25	3	0.030	0.026	0.035	0.0046
June 17	29	F	9	0.025	0.007	0.032	0.0091
June 23	35	D75	2	0.034	0.030	0.038	0.0055
June 23	35	D50	3	0.032	0.029	0.034	0.0025
June 23	35	D25	3	0.032	0.028	0.037	0.0043
June 23	35	F	8	0.028	0.013	0.033	0.0073
July 5	47	D50	2	0.029	0.029	0.030	0.0009
July 5	47	D25	4	0.027	0.024	0.030	0.0025
July 15	57	D75	4	0.029	0.026	0.032	0.0030
July 15	57	D50	3	0.026	0.020	0.031	0.0056
July 15	57	D25	4	0.018	0.015	0.025	0.0049
July 15	57	F	3	0.038	0.036	0.039	0.0016
July 22	64	D75	9	0.017	0.010	0.025	0.0058
July 22	64	D50	4	0.023	0.014	0.033	0.010
July 22	64	D25	1	0.022	0.022	0.022	–
July 22	64	F	2	0.040	0.039	0.040	0.0004
August 18	81	D75	1	0.003	0.003	0.003	–
August 18	81	D50	1	0.004	0.004	0.004	–
August 18	81	D_total	28	0.012	0.003	0.032	0.0058

TABLE 4 | Technical specifications for the terrestrial laser scanners.

Scanner type	Beam divergence (mrad)	Beam diameter at exit (mm)	Wavelength (nm)	Output power (mW)	Scan rate (kHz)	Intensity recording (DN)	Ranging error (mm)
Leica HDS6100	0.22	3	690	30	508	–1,228 to 2,048	±2
FARO S120	0.19	3	905	20	488	–2,048 to 2,033	±2
FARO X330	0.19	2.25	1,550	500	488	–2,048 to 2,033	±2

DN, digital number.

the calibration of intensity data may be required. The stationary four-grade Spectralon panel was used as a reference target for the calibration procedure using the mean of ~65,000 points from each of the reflectance panels. The standard deviation

of raw intensity values of these samples varied between 4.7 and 8.9. The panel's reflectance was also measured with a FieldSpec Pro FR (Analytical Spectral Devices, Inc., Longmont, CO, USA) field spectrophotometer. The relationship between

the spectrophotometer measured reflectance and the laser scanner measured intensity was then investigated to detect any nonlinearities in the intensity response.

The relationship between measured intensity and reflectance was shown to be linear for the Leica scanner and logarithmic for the FARO scanners (**Figure 1**). The measured intensity from the FARO X330 scanner was shown to be saturated in the 99% panel because the maximum intensity of the digital number was reached; thus, the calibration of FARO X330 data was done using only the 50, 25, and 12% panels. To correct for the logarithmic effect of the FARO scanners, Equation 3 was used (Kaasalainen et al., 2009):

$$y = 10^{\frac{(x-A_1)}{A_0}}, \quad (3)$$

where A_0 and A_1 are empirical parameters determined by fitting the raw intensity measured by the FARO scanners to the spectrophotometer measured reflectance. Additionally, a linear regression model was developed to model the relationship between the raw intensity and reflectance for the Leica scanner to convert the raw intensity values to calibrated intensity. The fitted parameter values of A_1 and A_0 were 1833.8 and 446.9 for FARO S120, and 2018.7 and 379.9 for FARO X330. The linear regression model of the Leica scanner had a coefficient of 0.00119 and an intercept of -0.57186 .

The effect of ambient temperature, temperature of the scanner, and background illumination on measured intensity was then reduced by normalizing each scan with the mean intensity of $\sim 20,000$ returns from the 99% Spectralon panel in the case of the Leica and FARO S120, and from the 50% Spectralon panel for the FARO X330 due to the saturation of intensity values in the 99% panel. The intensity of the X330 points from the seedlings had values below the mean of the 50% Spectralon panel; thus, the intensity was within the calibrated range.

The result of this process is referred to as the calibrated intensity within this paper. Because the samples were scanned

within a narrow range bin of less than 1 m, no range corrections were conducted. A range correction is required when the targets are located at different distances from the scanner, i.e., scanning trees in a forest environment (Kaasalainen et al., 2011).

Laser Intensity Features

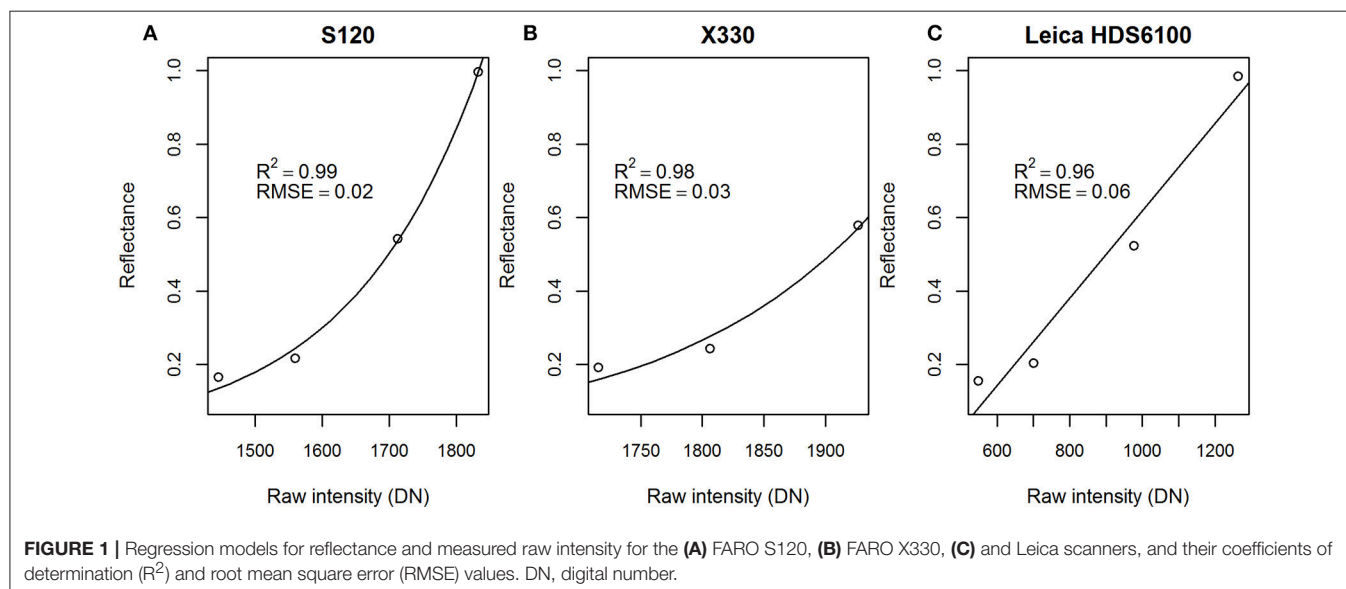
A set of statistical features was calculated from the intensity data from each point cloud representing a seedling at each wavelength. Intensity features were calculated from the calibrated intensity data of each point cloud. These features were the mean, minimum, maximum, standard deviation and percentiles 10, 20, 30, 40, 60, 70, 80, and 90 (p -value, e.g., p_{20} , p_{60}) of the intensity value distribution. Within this paper these statistics are referred to with a wavelength and a subscript describing the feature (e.g., $1,550_{\text{mean}}$, $690_{p_{30}}$). The included features are shown in **Table 5**.

Based on these features, a set of spectral indices was calculated for each point cloud (**Table 6**). We aimed to observe combinations of the 690 and 1,550 nm and 905 and 1,550 nm wavelengths based on ratio and normalization operations by calculating a simple ratio (SR) index (Jordan, 1969) and a normalized difference index (NDI) (Hancock et al., 2017) as follows:

$$\text{SR}_{i\rho 1,\rho 2} = \frac{\rho_{1i}}{\rho_{2i}}, \quad (4)$$

TABLE 5 | Intensity features.

Feature	Abbreviation	Description
Mean	mean	Mean calibrated intensity
Standard deviation	std	Standard deviation of calibrated intensity
Percentile	$p \dots i$	i th percentile of the intensity value distribution
Minimum	min	Minimum calibrated intensity
Maximum	max	Maximum calibrated intensity



$$NDI_{i\rho1,\rho2} = \frac{\rho1_i - \rho2_i}{\rho1_i + \rho2_i}, \quad (5)$$

where $\rho1_i$ and $\rho2_i$ are the i feature of wavelength $\rho1$ and $\rho2$.

The indices are referred to with the abbreviation, feature name, and a subscript describing the wavelengths used for the calculation (e.g., $NDI_{p80_{690, 1,550}}$). The total number of calculated laser intensity features was 48.

Statistical Analysis

The Student's t -test was used to determine whether there were statistically significant differences in the length of the seedlings in the different treatment groups. The relationship between EWT and each laser intensity feature was investigated using simple linear regression. The regression models were developed for all points, leaf points and stem points separately. Regression models were also developed for the randomly sampled point clouds to investigate how point density affects the estimation of EWT. Multiple regression was used to investigate whether the estimation of EWT could be improved by including length of the seedling as a structural variable in the regression model with the best explanatory laser intensity feature. We used the coefficient of determination (R^2) and RMSE with cross-validation to assess the goodness of fit between the dependent and the predictor using the following equations:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (6)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \quad (7)$$

where n is the number of observations, y_i is the observed value for the measurement i , \hat{y}_i is the predicted value for the measurement i , and \bar{y} is the mean of the observed data. All of the statistical analyses were performed using the open source software package R ver. 3.2.3 (R Core Team, 2013). Due to the large number of explanatory laser intensity features, only the 10 best explaining variables with the highest R^2 were reported.

RESULTS

Effects of Drought on Growth and Equivalent Water Thickness

The drought treatments resulted in varying lengths of the seedlings in the treatment groups (Figure 2). The D75 group had a mean length of 29.5 cm, while the D50 and D25 groups had significantly lower mean lengths of 26.0 and 23.7 cm, respectively. The drought treatments significantly affected the growth of the seedlings.

The seedlings showed resistance to drought. Despite the different amounts of water that the seedlings in the drought treatment groups received, EWT was shown to stay at relatively similar levels between treatment groups until the irrigation amount was further reduced after 35 days from starting the

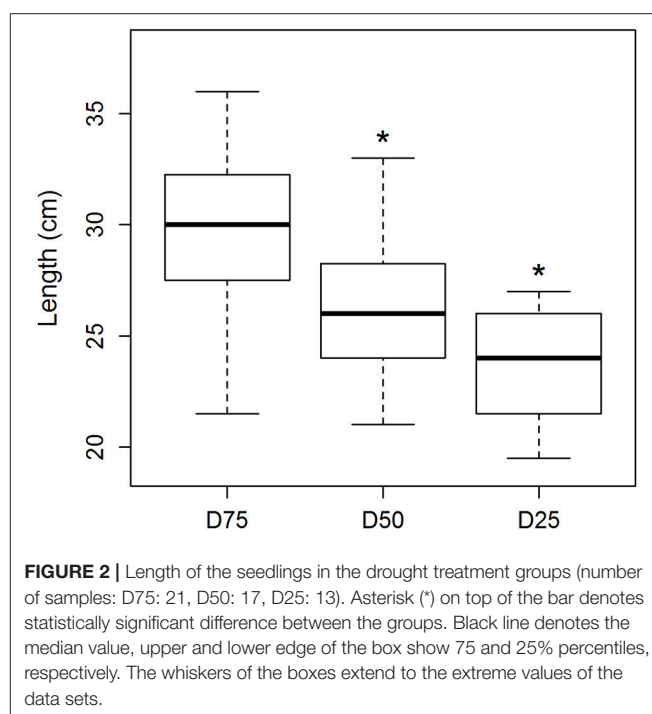


TABLE 6 | Calibrated intensity values for the mean of each wavelength and each spectral index for all points, leaf points, and stem points.

Laser intensity feature	All points				Leaf points				Stem points			
	Mean	Min	Max	Std	Mean	Min	Max	Std	Mean	Min	Max	Std
690 _{mean}	0.16	0.10	0.27	0.03	0.14	0.08	0.27	0.04	0.33	0.18	0.43	0.04
905 _{mean}	0.16	0.11	0.25	0.03	0.14	0.10	0.25	0.03	0.35	0.29	0.42	0.03
1,550 _{mean}	0.09	0.07	0.12	0.02	0.08	0.06	0.11	0.01	0.20	0.14	0.32	0.04
SR _{mean690, 1550}	1.70	0.84	2.69	0.43	1.64	0.82	2.53	0.44	1.69	0.76	2.35	0.31
SR _{mean905, 1550}	1.75	1.26	2.31	0.30	1.75	1.19	2.35	0.33	1.73	1.25	2.44	0.24
NDI _{mean690, 1550}	0.24	-0.09	0.46	0.12	0.22	-0.10	0.43	0.13	0.24	-0.14	0.40	0.10
NDI _{mean905, 1550}	0.26	0.11	0.39	0.08	0.26	0.09	0.40	0.09	0.26	0.11	0.42	0.07

NDI, normalized difference index; SR, simple ratio (index).

experiment (Figure 3). The seedlings still showed few visual symptoms due to drought, and after further reductions of irrigation the EWT decreased significantly. The lowest EWT levels in the experiment were found at the end of the experiment in the D_total group.

Effects of *E. polonica* Treatment on Equivalent Water Thickness

The fungal pathogen treatment of the seedlings resulted in decreased EWT (Figure 4). The affected seedlings showed a rapid decrease in EWT because the effects were apparent only 14 days after the inoculation. The infected seedlings were shown to respond to the treatment during the first 35 days of the experiment. After that period, no signs of infection in the remaining seedlings were observed.

Does Equivalent Water Thickness Affect the Spectral Response of Multispectral Terrestrial Laser Scanning?

The calibrated intensity at 1550 nm wavelength from the FARO X330 showed significantly increasing values with decreasing EWT (Figure 5). The linear regression model of EWT and the best explanatory variable at 1,550 nm wavelength, $1,550_{p80}$, had a coefficient of determination of 0.7. For comparison, $1,550_{mean}$ was able to explain only 47% of the variation in EWT. Spectral indices calculated from both combinations – 905 and 1,550 nm and 690 and 1,550 nm – were able to predict EWT with improved accuracy. The combination of 905 and 1,550 nm showed the best prediction accuracy of EWT with R^2

of 0.89 and 0.87 for $NDI_{p70_{905, 1,550}}$ and $NDI_{p60_{905, 1,550}}$, respectively (Table 7). Similar prediction accuracy, with an R^2 of 0.87, was also observed for $SR_{p70_{905, 1,550}}$ and $NDI_{p80_{905, 1,550}}$. The best explanatory features for the combination of 690 and 1,550 nm showed prediction accuracies of R^2 of 0.82 and 0.80 for $NDI_{p70_{690, 1,550}}$ and $SR_{p70_{690, 1,550}}$, respectively.

The best explanatory laser intensity features varied slightly for leaf points. The $1,550_{std}$ feature could explain 79%, and $1,550_{p90}$ 73%, of the variation in EWT as the best predicting features for single wavelength data. The spectral indices showed

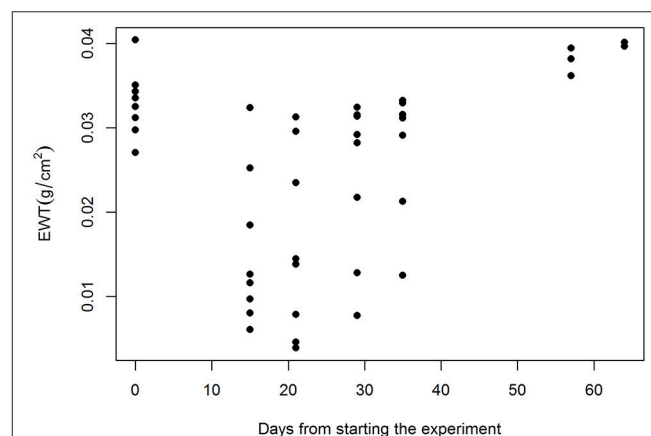


FIGURE 4 | EWT alterations during the experiment in the fungal pathogen treatment group.

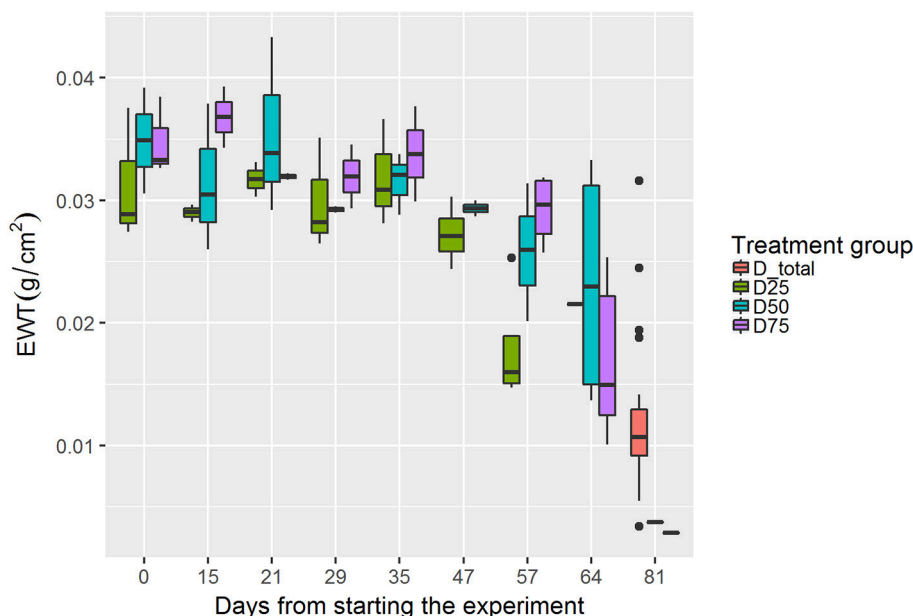


FIGURE 3 | Changes in EWT in each drought treatment group during the experiment (each bar represents approximately 4 samples). Black line denotes the median value, upper and lower edges of the boxes show 75 and 25% percentiles, respectively. The whiskers of the boxes extend to extreme values no longer than 1.5 times the interquartile range. Values further than that are plotted as outliers.

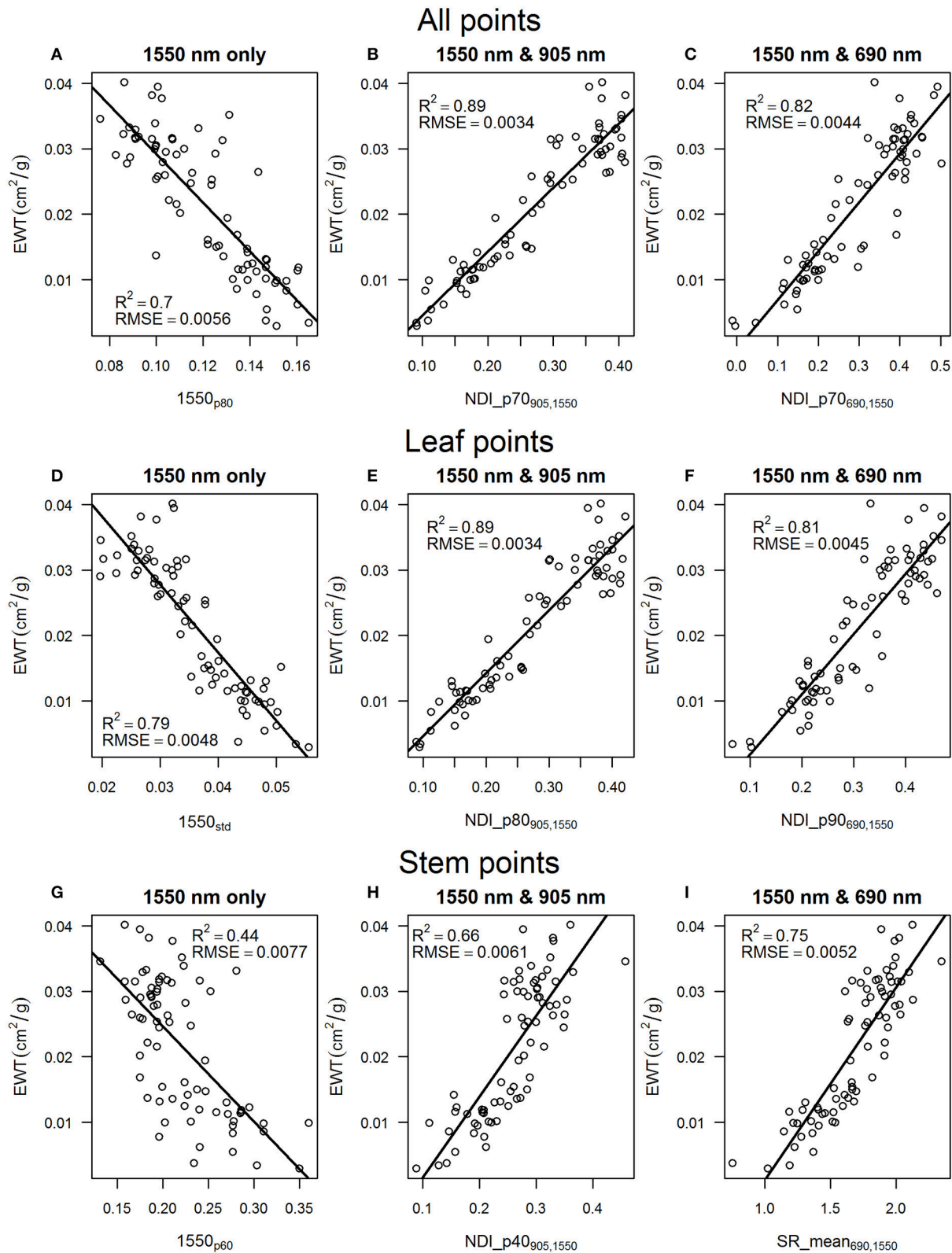


FIGURE 5 | Relationships between EWT and the best explanatory laser intensity features with their respective R^2 and RMSE values for all points for (A) 1,550 nm wavelength, (B) 1,550 and 905 nm wavelengths, (C) 1,550 and 690 nm wavelengths. The same for leaf points for (D) 1,550 nm wavelength, (E) 1,550 and 905 nm wavelengths, (F) 1,550 and 690 nm wavelengths, and the same for stem points for (G) 1,550 nm wavelength, (H) 1,550 and 905 nm wavelengths, (I) 1,550 and 690 nm wavelengths.

TABLE 7 | R^2 and RMSE values for the linear regression models of equivalent water thickness and laser intensity features for all points, leaf points, and stem points.

All points			Leaf points			Stem points		
Laser intensity feature	R^2	RMSE (g/cm ²)	Laser intensity feature	R^2	RMSE (g/cm ²)	Laser intensity feature	R^2	RMSE (g/cm ²)
NDI_p70 _{905, 1,550}	0.89	0.0034	NDI_p80 _{905, 1,550}	0.89	0.0034	SR_mean _{690, 1,550}	0.75	0.0052
NDI_p60 _{905, 1,550}	0.87	0.0036	NDI_p70 _{905, 1,550}	0.89	0.0035	SR_p40 _{690, 1,550}	0.74	0.0053
SR_p70 _{905, 1,550}	0.87	0.0037	SR_p70 _{905, 1,550}	0.87	0.0037	SR_p30 _{690, 1,550}	0.73	0.0054
NDI_p80 _{905, 1,550}	0.87	0.0037	SR_p80 _{905, 1,550}	0.87	0.0038	SR_p60 _{690, 1,550}	0.73	0.0054
SR_p60 _{905, 1,550}	0.86	0.0038	NDI_p60 _{905, 1,550}	0.86	0.0038	SR_p70 _{690, 1,550}	0.71	0.0056
NDI_mean _{905, 1,550}	0.85	0.0040	NDI_p90 _{905, 1,550}	0.86	0.0039	NDI_mean _{690, 1,550}	0.70	0.0056
SR_p80 _{905, 1,550}	0.84	0.0041	NDI_mean _{905, 1,550}	0.85	0.0040	SR_p20 _{690, 1,550}	0.70	0.0057
SR_mean _{905, 1,550}	0.83	0.0043	SR_p60 _{905, 1,550}	0.84	0.0041	NDI_p30 _{690, 1,550}	0.67	0.0059
NDI_p70 _{690, 1,550}	0.82	0.0044	SR_p90 _{905, 1,550}	0.83	0.0043	NDI_p60 _{690, 1,550}	0.67	0.0059
SR_p70 _{690, 1,550}	0.80	0.0046	SR_mean _{905, 1,550}	0.82	0.0043	NDI_p40 _{690, 1,550}	0.67	0.0059

similar explanatory power compared with using all points, but the best features for estimating EWT were higher intensity percentiles than for all points. The NDI_p80_{905, 1,550} and the NDI_p70_{905, 1,550} features explained 89% of the variation in EWT, while SR_p70_{905, 1,550} and SR_p80_{905, 1,550} explained 87% of the variation in EWT. The NDI_p90_{690, 1,550} and the SR_p90_{690, 1,550} features were the best predictors for the combination of 690 and 1,550 nm wavelengths, with R^2 of 0.81 and 0.79, respectively.

The stem points were also shown to have statistically significant power in explaining the differences in EWT but with lower percentiles of the intensity value distribution and lower coefficients of determination of 0.49, 0.66, and 0.75 for 1550 nm only, 1,550 and 905 nm, and 1,550 and 690 nm, respectively. Single wavelength data explained 49% of the variation in EWT with the 1,550_{mean} feature. The combination of 690 and 1,550 nm wavelengths showed the highest prediction accuracy, with the SR_mean_{690, 1,550} and SR_p40_{690, 1,550} features resulting in an R^2 of 0.75 and 0.74, respectively. The NDI_p40_{905, 1,550} and the NDI_p10_{905, 1,550} features were able to predict EWT with R^2 of 0.66 and 0.65, respectively.

The regression models explained 89% of the variation in EWT when 2,000 points from each seedling were randomly sampled. Thus, there was only a 0.0001 g/cm² increase in RMSE for both spectral indices compared using all points (Table 8). A slight decrease of 0.02 in R^2 and an increase of 0.0002 g/cm² was present for NDI_p70_{905, 1,550} after random sampling of 1,000 points, whereas NDI_p70_{690, 1,550} showed no changes compared with 2,000 points. Reducing the number of points to 500 resulted in no changes in the estimation accuracy of EWT for NDI_p70_{905, 1,550} but R^2 decreased to 0.79 and RMSE increased to 0.0048 g/cm² for NDI_p70_{690, 1,550}. When the number of points was reduced to 250, R^2 decreased to 0.79 and 0.78 and RMSE increased to 0.0047 g/cm² and 0.0049 g/cm² for NDI_p70_{905, 1,550} and NDI_p70_{690, 1,550}, respectively.

Multiple regression was able to improve the prediction of EWT. According to the multiple regression analysis, the length was statistically significant at $p < 0.001$. The estimation accuracy of EWT was improved compared with using simple

TABLE 8 | R^2 and RMSE values for the linear regression models of the spectral indices and equivalent water thickness after random sampling of the point clouds.

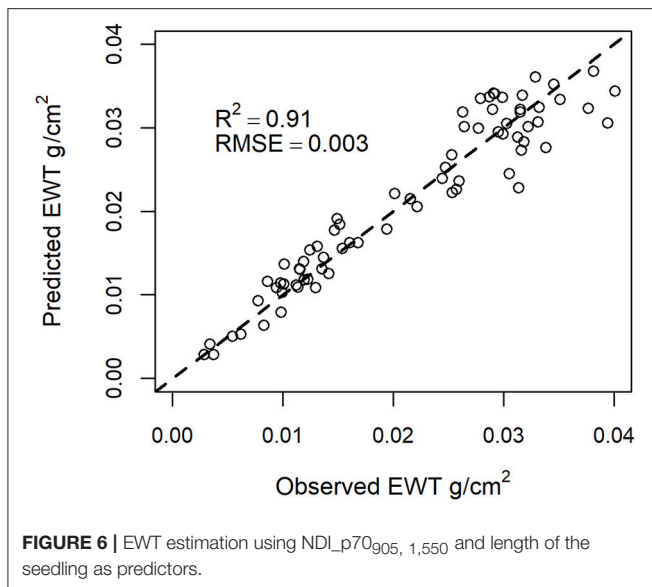
Spectral index	Number of points/seedling	R^2	RMSE
NDI_p70 _{905, 1,550}	2,000	0.89	0.0035
	1,000	0.87	0.0037
	500	0.87	0.0037
	250	0.79	0.0047
NDI_p70 _{690, 1,550}	2,000	0.81	0.0045
	1,000	0.81	0.0045
	500	0.79	0.0048
	250	0.78	0.0049

regression models, with $R^2 = 0.91$ and RMSE = 0.003 g/cm² (Figure 6).

DISCUSSION

The seedlings showed surprisingly high resistance to drought despite the low amounts of irrigation for each seedling. The drought led to a significant reduction in growth, which correlated with the intensity of the drought treatment. The seedlings seemed to cope with drought by reducing growth, which led to reduced demand for water because there was less foliage to support and less transpiring leaf area. However, significant reduction in EWT was observed after further reduction in irrigation.

The results showed that the investigated laser intensity features were able to explain most of the variation in leaf EWT and provide a relatively good prediction accuracy. Single wavelength intensity features at 1,550 nm were able to explain 70–79% of the differences in EWT using all points and leaf points only. Similar results have been obtained before for broadleaf species after applying an incidence angle correction to the intensity data at 1,550 nm (Zhu et al., 2015; Junttila et al., 2016). The correlations obtained in this study are relatively high for single wavelength data considering that no radiometric



corrections have been applied to the data. Due to the shape and size of the spruce needles, an incidence angle correction would be impossible to apply because the laser footprint illuminates several needles at a time.

The spectral indices calculated from the 690, 905, and 1,550 nm wavelengths showed higher correlations with leaf EWT than the single wavelength data, which is in accordance with previous research (Junttila et al., 2016). The regression model of EWT and NDI_p70690, 1,550 showed a coefficient of determination of 0.82 for all points, while NDI_p70905, 1,550 was able to explain 89% of the changes in EWT. The better correlation of the combination of 1,550 and 905 nm with EWT can be explained by the sensitivity of the 905 nm wavelength to internal structure and dry matter content, a property needed in the estimation of EWT according to the literature (Ceccato et al., 2001). Additionally, the scanners providing measurements at 1,550 and 905 nm wavelengths were both developed by FARO and seem to employ similar algorithms for producing the discrete point clouds. It can be seen from the relationships between raw intensity and reflectance (**Figure 1**) that FARO scanners are more similar compared with the Leica, which could have affected our results.

Leaf points were used to develop linear regression models of the relationship between EWT and laser intensity features. The models showed that the most accurate prediction of EWT was achieved with the NDI_p80905, 1,550 feature with an R^2 of 0.89. The accuracy of the prediction of EWT with segmented leaf points was the same as with all points but with a different feature. The best explanatory feature using all points was NDI_p70690, 1,550. This suggests that there are a lot of “noise” points still present in the point clouds considering the estimation of EWT in the lowest intensity percentiles because a higher percentile of the intensity distribution best explained EWT using leaf points. An intensity threshold could be applied to reduce the amount of noise points from the needles, but applying such a

threshold greatly affects the distribution of the intensity values. Thus, a robust method would be needed to apply a threshold to different scans without losing valuable information regarding the intensity distribution.

The point density in each point cloud representing a seedling was reduced using random sampling to investigate the effect of point density on the estimation of EWT. The results showed that a lower scanning resolution could be used because the spectral indices calculated from the sparser point clouds still explained EWT almost as well as with maximum point density. After reducing the number of points down to 500 per seedling, the amount of variation explained was only slightly reduced. A more severe reduction in the amount of variance explained was present after the number of points was reduced to 250 per seedling. The NDI_p70905, 1,550 index seemed to be more affected than NDI_p70690, 1,550 by the reduction in point density.

A multiple regression model was developed with a laser intensity feature and length as predictors to estimate EWT. The inclusion of the length of the seedling in the regression model increased the accuracy of the estimation of EWT, although it should be pointed out that the length of the seedling varied only from 16.5 to 42 cm. Presumably, the inclusion of the length of the seedling is able to explain a part of the variation in EWT due to physiological alterations that occur during the life span of the needles. During the growing season, the specific leaf area (i.e., dry matter content per leaf area) increases due to accumulation of dry matter resulting in a lower EWT (Jach and Ceulemans, 2000).

The measurements in this study were conducted in a greenhouse, where environmental conditions were similar between measurements. Temperature and ambient light conditions did change a little between measurements depending on whether it was a sunny or cloudy day, but otherwise no alterations in range or background of the seedlings were present. The change in temperature and light conditions is a limitation that needs to be taken into account considering the applicability of the method for the forest environment. Here, we also found that stem points responded to changes in EWT, which is likely due to the thin bark of the seedlings. Mature trees have thicker bark; thus, the estimation of EWT with multispectral TLS likely requires segmentation or classification of laser returns to leaf and stem points.

The terrestrial laser scanners had slightly different technical parameters: The FARO scanners had similar specifications except for the laser beam diameter at output, which was 2.25 and 3 mm for the X330 and S120, respectively. The different specifications resulted in a different laser footprint at target but it seems that having the same laser footprint size at target is not crucial for calculating spectral indices because the correlations were high despite this factor. However, an airborne laser scanning measurement would result in much larger laser footprints, ~5–30 cm in diameter depending on the flying altitude, and a significantly lower point density. At these resolutions, the separation of woody and leaf material is practically impossible. The results obtained in this study show that the separation is not always necessary, but mature trees with large branches are likely to show a different response in terms of laser intensity.

The effect of incidence angle on spectral indices developed from multispectral laser scanning has been recently studied (Kasalanen et al., 2016; Hancock et al., 2017). An NDI utilizing 1,063 and 1,545 nm wavelengths was concluded to not be sensitive to incidence angle effects when retrieving leaf moisture content. The combination of 905 and 1,550 nm wavelengths could be assumed to have similar properties due to the physical similarity of the wavelengths. Here, the needles had varying incidence angles and also varying receiving area because needles are not evenly spaced along the branches. Using an NDI calculated from the 905 and 1,550 nm wavelengths seems to reduce the effect of such confounding factors because using single wavelength data at 1,550 nm did not provide estimates of EWT as accurate as the spectral index did.

A commercially available multispectral laser scanner has been developed for airborne measurements, called Optech Titan (Teledyne Optech, Concord, Canada), which employs three wavelengths at 532, 1,064, and 1,550 nm. The combination of 1,064 and 1,550 nm wavelengths is interesting from the forest health mapping aspect, because these bands have been used before for EWT estimation (Gaulton et al., 2013). However, the different wavelengths of the Optech Titan have different footprints in terms of spatial distribution due to the instrument design (Morsy et al., 2016). The 1,064 nm channel is looking at nadir while the 1,550 nm channel is looking forward at 3.5° angle. This feature of the instrument presumably complicates the use of NDIs with the data, but the matter requires further investigation on the extent of the effect of varying spatial footprints at the canopy scale. A multispectral laser scanner for airborne measurements with overlapping footprints would be of great benefit in exploiting all the gains of the multispectral domain.

Based on the results, multisensor MS-TLS is capable of explaining most of the variation in leaf EWT of trees ($R^2 = 0.91$) in a controlled environment. The combination of 905 and 1,550 nm wavelengths explained variation in EWT better than the combination of 690 and 1,550 nm, which is also supported by previous research (Ceccato et al., 2001). The high coefficients of determination in the estimation of EWT with spectral indices did not seem to require a high point density like we had in the original scans, but the scanning could be conducted at a

lower resolution. Single trees measured from the sample plots provide a base for mapping and monitoring forests. Considering the small scale of our study and the larger scale of information that is required for monitoring declining forests and trees, we suggest that in the next step, tree-level EWT estimation in forest environment should be investigated using multispectral laser scanning at the 905 and 1,550 nm wavelengths and by classifying leaf and stem points.

AUTHOR CONTRIBUTIONS

SJ designed the experiment, conducted data processing and statistical analysis, and wrote the manuscript. JS contributed to the laboratory analysis and writing of the manuscript. MV assisted in designing the experiment, statistical analysis, and the writing of the manuscript. RL contributed to the design of the experiment, data collection, and commented on the manuscript. HK contributed to the design of the experiment, gave technical advice, and commented on the manuscript. AK assisted with the measurement set-up and provided technical support. MH, HH, and JH arranged all the required resources, supervised the experiment, and participated in the writing of the manuscript. All of the authors have contributed to the finalizing of the manuscript.

FUNDING

This work was supported by the Academy of Finland project Centre of Excellence in Laser Scanning Research (CoE-LaSR) under Grant 272195, Unmanned Airborne Vehicle-based 4D Remote Sensing for Mapping Rain Forest Biodiversity and Its Change in Brazil (Decision number 273806), the Finnish Cultural Foundation under Grant 2CE87B6E, the Finnish Society of Forest Science, and the Niemi Foundation.

ACKNOWLEDGMENTS

We would like to thank the Natural Resources Institute Finland (Luke) for providing greenhouse and laboratory facilities, and for assistance in maintaining the seedlings.

REFERENCES

- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., et al. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manage.* 259, 660–684. doi: 10.1016/j.foreco.2009.09.001
- Balduzzi, M. A., Van der Zande, D., Stuckens, J., Verstraeten, W. W., and Coppin, P. (2011). The properties of terrestrial laser system intensity for measuring leaf geometries: a case study with conference pear trees (*Pyrus Communis*). *Sensors* 11, 1657–1681. doi: 10.3390/s110201657
- Carter, G. A. (1993). Responses of leaf spectral reflectance to plant stress. *Am. J. Bot.* 80, 239–243. doi: 10.1002/j.1537-2197.1993.tb13796.x
- Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., and Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ.* 77, 22–33. doi: 10.1016/S0034-4257(01)00191-2
- Chaerle, L., and Van Der Straeten, D. (2000). Imaging techniques and the early detection of plant stress. *Trends Plant Sci.* 5, 495–501. doi: 10.1016/S1360-1385(00)01781-7
- Chuvieco, E., Cocero, D., Riano, D., Martin, P., Martinez-Vega, J., de la Riva, J., et al. (2004). Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. *Remote Sens. Environ.* 92, 322–331. doi: 10.1016/j.rse.2004.01.019
- Danson, F., Steven, M., Malthus, T., and Clark, J. (1992). High-spectral resolution data for determining leaf water content. *Int. J. Remote Sens.* 13, 461–470. doi: 10.1080/01431169208904049

- Dimitrakopoulos, A., and Papaioannou, K. K. (2001). Flammability assessment of Mediterranean forest fuels. *Fire Technol.* 37, 143–152. doi: 10.1023/A:1011641601076
- Douglas, E. S., Strahler, A., Martel, J., Cook, T., Mendillo, C., Marshall, R., et al. (2012). "DWEL: a dual-wavelength echidna lidar for ground-based forest scanning," in *Geoscience and Remote Sensing Symposium (IGARSS)*, 2012 IEEE International: (Munich: IEEE), 4998–5001.
- Easlon, H. M., and Bloom, A. J. (2014). Easy leaf area: automated digital image analysis for rapid and accurate measurement of leaf area. *Appl. Plant Sci.* 2:1400033. doi: 10.3732/apps.1400033
- Fassnacht, F. E., Latifi, H., Ghosh, A., Joshi, P. K., and Koch, B. (2014). Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sens. Environ.* 140, 533–548. doi: 10.1016/j.rse.2013.09.014
- Gaulton, R., Danson, F., Ramirez, F., and Gunawan, O. (2013). The potential of dual-wavelength laser scanning for estimating vegetation moisture content. *Remote Sens. Environ.* 132, 32–39. doi: 10.1016/j.rse.2013.01.001
- Girardeau-Montaut, D. (2011). *Cloudcompare-Open Source Project*. OpenSource Project.
- Goodwin, N. R., Coops, N. C., Wulder, M. A., Gillanders, S., Schroeder, T. A., and Nelson, T. (2008). Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sens. Environ.* 112, 3680–3689. doi: 10.1016/j.rse.2008.05.005
- Gschwantner, T., Lanz, A., Vidal, C., Bosela, M., Di Cosmo, L., Fridman, J., et al. (2016). Comparison of methods used in European National Forest Inventories for the estimation of volume increment: towards harmonisation. *Ann. For. Sci.* 73, 807–821. doi: 10.1007/s13595-016-0554-5
- Hakala, T., Suomalainen, J., Kaasalainen, S., and Chen, Y. (2012). Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Opt. Express* 20, 7119–7127. doi: 10.1364/OE.20.007119
- Hancock, S., Gaulton, R., and Danson, F. M. (2017). Angular reflectance of leaves with a dual-wavelength terrestrial lidar and its implications for leaf-bark separation and leaf moisture estimation. *IEEE Trans. Remote Sens.* 55, 3084–3090. doi: 10.1109/TGRS.2017.2652140
- Hanewinkel, M., Cullmann, D. A., Schelhaas, M.-J., Nabuurs, G.-J., and Zimmermann, N. E. (2013). Climate change may cause severe loss in the economic value of European forest land. *Nat. Clim. Change* 3, 203–207. doi: 10.1038/nclimate1687
- Hornqvist, R., Christiansen, E., Solheim, H., and Wang, S. (1983). Artificial inoculation with [ps typographus-associated blue-stain fungi can kill healthy Norway spruce trees. *Medd. Nor. Inst. Skogforsk.* 38, 1–20.
- Jach, M. E., and Ceulemans, R. (2000). Effects of season, needle age and elevated atmospheric CO₂ on photosynthesis in Scots pine (*Pinus sylvestris*). *Tree Physiol.* 20, 145–157. doi: 10.1093/treephys/20.3.145
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50, 663–666. doi: 10.2307/1936256
- Junttila, S., Vastaranta, M., Liang, X., Kaartinen, H., Kukko, A., Kaasalainen, S., et al. (2016). Measuring leaf water content with dual-wavelength intensity data from terrestrial laser scanners. *Remote Sensing* 9:8. doi: 10.3390/rs9010008
- Kaasalainen, S., Jaakkola, A., Kaasalainen, M., Krooks, A., and Kukko, A. (2011). Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: search for correction methods. *Remote Sens.* 3, 2207–2221. doi: 10.3390/rs3102207
- Kaasalainen, S., Krooks, A., Kukko, A., and Kaartinen, H. (2009). Radiometric calibration of terrestrial laser scanners with external reference targets. *Remote Sens.* 1, 144–158. doi: 10.3390/rs1030144
- Kaasalainen, S., Nevalainen, O., Hakala, T., and Anttila, K. (2016). Incidence angle dependency of leaf vegetation indices from hyperspectral lidar measurements. *Photogramm. Fernerkundung Geoinform.* 2016, 75–84. doi: 10.1127/pfg/2016/0287
- Lausch, A., Heurich, M., Gordalla, D., Dobner, H.-J., Gwilym-Margianto, S., and Salbach, C. (2013). Forecasting potential bark beetle outbreaks based on spruce forest vitality using hyperspectral remote-sensing techniques at different scales. *For. Ecol. Manage.* 308, 76–89. doi: 10.1016/j.foreco.2013.07.043
- Li, W., Niu, Z., Sun, G., Gao, S., and Wu, M. (2016). Deriving backscatter reflective factors from 32-channel full-waveform LiDAR data for the estimation of leaf biochemical contents. *Opt. Express* 24, 4771–4785. doi: 10.1364/OE.24.004771
- Matikainen, L., Hyyppä, J., and Litkey, P. (2016). Multispectral airborne laser scanning for automated map updating. *ISPRS Intern. Arch. Photogramm. Remote Sens. Spatial Inform. Sci.* XLI-B3, 323–330. doi: 10.5194/isprsarchives-XLI-B3-323-2016
- McDowell, N. G., and Allen, C. D. (2015). Darcy's law predicts widespread forest mortality under climate warming. *Nat. Clim. Change* 5, 669–672. doi: 10.1038/nclimate2641
- Meddens, A. J., Hicke, J. A., Vierling, L. A., and Hudak, A. T. (2013). Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sens. Environ.* 132, 49–58. doi: 10.1016/j.rse.2013.01.002
- Morris, J. L., Cottrell, S., Fettig, C. J., Hansen, W. D., Sherriff, R. L., Carter, V. A., et al. (2016). Managing bark beetle impacts on ecosystems and society: priority questions to motivate future research. *J. Appl. Ecol.* 54, 750–760. doi: 10.1111/1365-2664.12782
- Morsy, S., Shaker, A., El-Rabbany, A., and LaRocque, P. E. (2016). Airborne multispectral LiDAR data for land-cover classification and land/water mapping using different spectral indexes. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inform. Sci.* 3, 217–224. doi: 10.5194/isprsannals-III-3-217-2016
- Niu, Z., Xu, Z., Sun, G., Huang, W., Wang, L., Feng, M., et al. (2015). Design of a new multispectral waveform LiDAR instrument to monitor vegetation. *IEEE Geosci. Remote Sens. Lett.* 12, 1506–1510. doi: 10.1109/LGRS.2015.2410788
- Raffa, K. F., Aukema, B., Bentz, B. J., Carroll, A., Erbilgin, N., Herms, D. A., et al. (2009). A literal use of forest health safeguards against misuse and misapplication. *J. For.* 107, 276–277. doi: 10.1093/jof/107.5.276
- Senf, C., Seidl, R., and Hostert, P. (2017). Remote sensing of forest insect disturbances: current state and future directions. *Int. J. Appl. Earth Observ. Geoinform.* 60, 49–60. doi: 10.1016/j.jag.2017.04.004
- Skakun, R. S., Wulder, M. A., and Franklin, S. E. (2003). Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sens. Environ.* 86, 433–443. doi: 10.1016/S0034-4257(03)00112-3
- Sow, M., Mbow, C., Hély, C., Fensholt, R., and Sambou, B. (2013). Estimation of herbaceous fuel moisture content using vegetation indices and land surface temperature from MODIS data. *Remote Sens.* 5, 2617–2638. doi: 10.3390/rs5062617
- R Core Team (2013). *A language and Environment For Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Trumbore, S., Brando, P., and Hartmann, H. (2015). Forest health and global change. *Science* 349, 814–818. doi: 10.1126/science.aac6759
- Wallace, A., Nichol, C., and Woodhouse, I. (2012). Recovery of forest canopy parameters by inversion of multispectral LiDAR data. *Remote Sens.* 4, 509–531. doi: 10.3390/rs4020509
- Watson, G. A. (2006). Computing helmert transformations. *J. Comput. Appl. Math.* 197, 387–394. doi: 10.1016/j.cam.2005.06.047
- Wei, G., Shalei, S., Bo, Z., Shuo, S., Faquan, L., and Xuewu, C. (2012). Multi-wavelength canopy LiDAR for remote sensing of vegetation: design and system performance. *ISPRS J. Photogramm. Remote Sens.* 69, 1–9. doi: 10.1016/j.isprsjprs.2012.02.001
- Wermelinger, B. (2004). Ecology and management of the spruce bark beetle *Ips typographus* — a review of recent research. *For. Ecol. Manage.* 202, 67–82. doi: 10.1016/j.foreco.2004.07.018
- White, J. C., Wulder, M. A., Brooks, D., Reich, R., and Wheate, R. D. (2005). Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery. *Remote Sens. Environ.* 96, 340–351. doi: 10.1016/j.rse.2005.03.007

- Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., et al. (2013). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nat. Clim. Change* 3, 292–297. doi: 10.1038/nclimate1693
- Wulder, M. A., White, J., Bentz, B., Alvarez, M., and Coops, N. (2006). Estimating the probability of mountain pine beetle red-attack damage. *Remote Sens. Environ.* 101, 150–166. doi: 10.1016/j.rse.2005.12.010
- Zarco-Tejada, P. J., Rueda, C., and Ustin, S. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* 85, 109–124. doi: 10.1016/S0034-4257(02)00197-9
- Zhu, X., Wang, T., Darvishzadeh, R., Skidmore, A. K., and Niemann, K. O. (2015). 3D leaf water content mapping using terrestrial laser scanner backscatter intensity with radiometric correction. *ISPRS J. Photogramm. Remote Sens.* 110, 14–23. doi: 10.1016/j.isprsjprs.2015.10.001
- Zhu, X., Wang, T., Skidmore, A. K., Darvishzadeh, R., Niemann, K. O., and Liu, J. (2017). Canopy leaf water content estimated using terrestrial LiDAR. *Agric. For. Meteorol.* 232, 152–162. doi: 10.1016/j.agrformet.2016.08.016

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Junttila, Sugano, Vastaranta, Linnakoski, Kaartinen, Kukko, Holopainen, Hyyppä and Hyyppä. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Quantitative Analysis of Cotton Canopy Size in Field Conditions Using a Consumer-Grade RGB-D Camera

Yu Jiang¹, Changying Li^{1*}, Andrew H. Paterson^{2,3}, Shangpeng Sun¹, Rui Xu¹ and Jon Robertson²

¹ Bio-sensing and Instrumentation Laboratory, School of Electrical and Computer Engineering, College of Engineering, University of Georgia, Athens, GA, United States, ² Plant Genome Mapping Laboratory, College of Agricultural and Environmental Sciences, University of Georgia, Athens, GA, United States, ³ Department of Genetics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States

OPEN ACCESS

Edited by:

Norbert Pfeifer,
Technische Universität Wien, Austria

Reviewed by:

Maria Balota,
Virginia Tech, United States
Marta Silva Lopes,
International Maize and Wheat
Improvement Center, Mexico

*Correspondence:

Changying Li
cyl@uga.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 29 September 2017

Accepted: 19 December 2017

Published: 30 January 2018

Citation:

Jiang Y, Li C, Paterson AH, Sun S,
Xu R and Robertson J (2018)
Quantitative Analysis of Cotton
Canopy Size in Field Conditions Using
a Consumer-Grade RGB-D Camera.
Front. Plant Sci. 8:2233.
doi: 10.3389/fpls.2017.02233

Plant canopy structure can strongly affect crop functions such as yield and stress tolerance, and canopy size is an important aspect of canopy structure. Manual assessment of canopy size is laborious and imprecise, and cannot measure multi-dimensional traits such as projected leaf area and canopy volume. Field-based high throughput phenotyping systems with imaging capabilities can rapidly acquire data about plants in field conditions, making it possible to quantify and monitor plant canopy development. The goal of this study was to develop a 3D imaging approach to quantitatively analyze cotton canopy development in field conditions. A cotton field was planted with 128 plots, including four genotypes of 32 plots each. The field was scanned by GPhenoVision (a customized field-based high throughput phenotyping system) to acquire color and depth images with GPS information in 2016 covering two growth stages: canopy development, and flowering and boll development. A data processing pipeline was developed, consisting of three steps: plot point cloud reconstruction, plant canopy segmentation, and trait extraction. Plot point clouds were reconstructed using color and depth images with GPS information. In colorized point clouds, vegetation was segmented from the background using an excess-green (ExG) color filter, and cotton canopies were further separated from weeds based on height, size, and position information. Static morphological traits were extracted on each day, including univariate traits (maximum and mean canopy height and width, projected canopy area, and concave and convex volumes) and a multivariate trait (cumulative height profile). Growth rates were calculated for univariate static traits, quantifying canopy growth and development. Linear regressions were performed between the traits and fiber yield to identify the best traits and measurement time for yield prediction. The results showed that fiber yield was correlated with static traits after the canopy development stage ($R^2 = 0.35\text{--}0.71$) and growth rates in early canopy development stages ($R^2 = 0.29\text{--}0.52$). Multi-dimensional traits (e.g., projected canopy area and volume) outperformed

one-dimensional traits, and the multivariate trait (cumulative height profile) outperformed univariate traits. The proposed approach would be useful for identification of quantitative trait loci (QTLs) controlling canopy size in genetics/genomics studies or for fiber yield prediction in breeding programs and production environments.

Keywords: cotton, high-throughput, phenotyping, field, RGB-D, morphological

1. INTRODUCTION

Cotton (*Gossypium*) is one of the most important textile fibers in the world, accounting for about 25% of total world textile fiber use (USDA-ERS, 2017). Thus, improvement of cotton production is vital to fulfilling the fiber requirements of over nine billion people by 2050 (Reynolds and Langridge, 2016). Plant canopy structure is an important trait, affecting crop functions such as light-energy production and utilization (Norman and Campbell, 1989). Optimal canopy structure can improve plant photosynthesis and thus crop yield potential (Reta-Sánchez and Fowler, 2002; Stewart et al., 2003; Giunta et al., 2008). One key to increasing yield is to figure out the optimal canopy structure for maximizing plant photosynthesis (Murchie et al., 2009; Zhu et al., 2010). Canopy size is an important aspect of canopy structure and critical to plant photosynthesis, fruiting, and biomass accumulation. However, assessment of canopy size becomes a bottleneck, which limits breeding programs and genetics studies (White et al., 2012; Cobb et al., 2013; Araus and Cairns, 2014; Barabaschi et al., 2016), especially for large crop populations and high-dimensional traits (e.g., canopy volume). Accurate and high throughput techniques for quantifying canopy size would facilitate cotton (and other) breeding programs and genetics studies (Araus and Cairns, 2014; Pauli et al., 2016; Reynolds and Langridge, 2016).

Canopy size is spatially and temporally variable, and morphological traits describing canopy size can be grouped based on different criteria (Norman and Campbell, 1989). From the spatial perspective, component traits of canopy size can be categorized as one-dimensional and multi-dimensional traits. One-dimensional (1D) traits quantify canopy size in a single dimension (e.g., canopy height and width), whereas multi-dimensional traits quantify canopy size by considering multiple dimensions (e.g., canopy area in two-dimensional (2D) space and volume in three-dimensional (3D) space). From the temporal perspective, the traits can be separated into static and dynamic categories. Static traits are directly measured for canopies at a certain time, whereas dynamic traits (e.g., growth rates) are the change of static traits over time. In addition to spatial and temporal criteria, data dimensionality is an important consideration of morphological traits for canopy size quantification. Based on the number of variables, morphological traits can be classified as univariate (e.g., maximum and average canopy height) or multivariate (e.g., cumulative height profile describing heights at different percentiles).

Univariate traits are widely used due to their simplicity (easy to define) and availability (most measurement instruments provide a single reading). One-dimensional traits can be measured manually (or automatically) using distance meters.

These 1D traits represent canopy size in only one dimension, creating a potential bias in structure assessment and comparison (Norman and Campbell, 1989). Leaf area as a 2D trait has been widely used for canopy size studies, and its derivative indicator (leaf area index, LAI) has demonstrated successes in estimating crop photosynthetic activities, biomass, yield, and biotic/abiotic stress tolerance (Sinclair and Horie, 1989; Kross et al., 2015; Feng et al., 2017). Leaf area can be directly measured using destructive methods, such as using a leaf area meter or image scanner to obtain the area of leaf samples that have been cut off the plant. Such destructive methods are usually arduous and hard to apply in large-scale experiments. To address those issues, LAI can be indirectly estimated using instruments such as the LI-COR LAI-2200 and Decagon LP-80 (Bréda, 2003; Weiss et al., 2004). Conventional 2D imaging methods can also be applied to calculate projected leaf area or canopy coverage ratio, estimating the true leaf area or LAI (Jonckheere et al., 2004; Zheng and Moskal, 2009). Nonetheless, all aforementioned methods require laborious data collection and do not provide 3D information, and therefore are inadequate to rapidly and comprehensively quantify canopy size and development.

Advanced 3D imaging approaches provide new opportunities to accurately quantify canopy size in multiple dimensions such as canopy height and width (1D), leaf area (2D), and volume (3D). The approaches can be categorized into passive and active 3D imaging (Li et al., 2014).

Passive 3D imaging methods reconstruct 3D structures of objects by expanding conventional 2D imaging methods. Stereo vision and the structure-from-motion (SfM) technique are two representative passive imaging methods. Small unmanned aerial systems (UASs) provide a means to quickly collect images for 3D reconstruction and trait extraction (e.g., crop height) using digital surface models and photogrammetry. However, images collected by UASs usually have lower quality (e.g., image resolution and sharpness) than images from ground systems, which could significantly affect the reconstruction accuracy (Shi et al., 2016) or even cause failures of 3D structure reconstruction (Jay et al., 2015). In addition, most of the passive techniques are computationally expensive (taking several hours for reconstruction of one plot) (Jay et al., 2015; Muller-Linow et al., 2015; Nguyen et al., 2016; Dong et al., 2017). Speeding up the process would require either using high-performance computing (HPC) resources, which would impose a considerable cost for large-scale breeding programs, or limiting the experimental scale and therefore breeding efficiency.

Active 3D imaging methods directly acquire 3D information by operating external light sources. Commonly used sensors include light detection and ranging (LiDAR), triangulation laser scanners, and time-of-flight (TOF) and structured light

cameras. With suitable camera setup and illumination, most active methods can accurately obtain 3D point clouds of plants in field conditions for canopy size analysis, but the instruments are usually expensive (\$2,000–100k or more) compared with cameras used in passive methods (Li et al., 2014). Recent advances in consumer-grade RGB-D cameras (e.g., Microsoft Kinect and ASUS Xtion) provide an inexpensive solution for 3D scanning (Nock et al., 2013; Paulus et al., 2014; Andujar et al., 2016). In particular, the Microsoft Kinect v2 camera uses the TOF principle with upgraded color and depth resolution, creating the possibility for inexpensive and high-resolution 3D sensing in field conditions. Several previous studies explored the use of Kinect v2 camera in measuring canopy size (height and volume) in field conditions, finding that the Kinect v2 was a promising tool for field-based phenotyping (Andujar et al., 2016; Jiang et al., 2016; Andujar et al., 2017). However, two aspects need to be further improved: (1) data processing should be fully automated to improve throughput, and (2) more spatially (single and multi-dimensional traits) and temporally (static and dynamic traits) morphological traits need to be measured and studied.

The overall goal of the study was to develop a 3D imaging approach to automatically and quantitatively analyze canopy size of cotton plants in field conditions. Specific objectives were to (1) develop algorithms to reconstruct colored point clouds of individual plots using depth and color images collected by Kinect v2 camera, (2) develop algorithms to segment canopy point clouds from the ground and weeds, (3) extract canopy size traits (canopy height, width, projected area, and volume) and their dynamic changes over time, and (4) explore the potential of using the extracted traits for fiber yield prediction.

2. MATERIALS AND METHODS

2.1. Experimental Design and Field Data Collection

A field (33.727239 N, 83.299097 W) was planted with cotton seed at the Iron Horse Farm of the University of Georgia in Watkinsville, Georgia, USA on 13 June 2016. The field contained 128 plots (16 rows with 8 plots per row) of length 3.05 m, with 1.83 m alleys between consecutive plots in a row, and row-spacing of 1.52 m (see Figure S1 in Supplementary Material). Three experimental genotypes (GA2011158, GA2009037, and GA2010074) and one commercial variety (Americot conventional) were used, each having 32 plot replicates. A completely randomized design was used to assign genotypes to individual plots. In each plot, 15 cotton seeds were manually planted at a spacing of 0.15 m. The first seed was planted 0.15 m away from the plot starting point, and hence a total of 15 seeds occupied 2.4 m out of 3.05 m in each plot. Cotton fiber was handpicked and weighed for 96 randomly selected plots (24 per genotype) on 4 November 2016 (which was 144 days after planting, DAP 144).

A field-based phenotyping system, GPhenoVision, was developed (Jiang et al., 2017) and used to scan the experimental field during midday (around 1,200–1,330 h) on 8 days in 2016 including 28 July (DAP 45), 4 August (DAP 52), 19 August (DAP

67), 26 August (DAP 74), 9 September (DAP 88), 16 September (DAP 95), 23 September (DAP 102), and 30 September (DAP 109). The scanning period covered two growth stages: canopy development (DAP 45–74), and flowering and boll development (DAP 74 to DAP 109). During the scanning period, the Kinect v2 camera of the GPhenoVision system was consistently used at 2.4 m above ground level (Figure 1A). The system ran in a continuous scanning mode at a constant speed of 1 m/s, but the operator manually controlled the data acquisition software to start/stop saving images at the beginning/end of each row to save storage space (Figure 1B).

2.2. Image Processing Pipeline

2.2.1. Reconstruction of Colorized Point Cloud

The entire image processing pipeline included three sections: point cloud reconstruction, cotton canopy segmentation, and trait extraction. The aim of the point cloud reconstruction section was to reconstruct colorized point clouds for individual plots using collected depth and color images with their corresponding GPS information (Figure 2). This section contained four steps:

2.2.1.1. Step 1.1: data grouping

All acquired GPS records were firstly converted from the geographic coordinate system to the universal transverse mercator (UTM) coordinate system. Altitude measures were used as the height of image acquisition. Collected depth and color images were segregated into individual plots based on their GPS coordinates, and the following processes were executed within each plot.

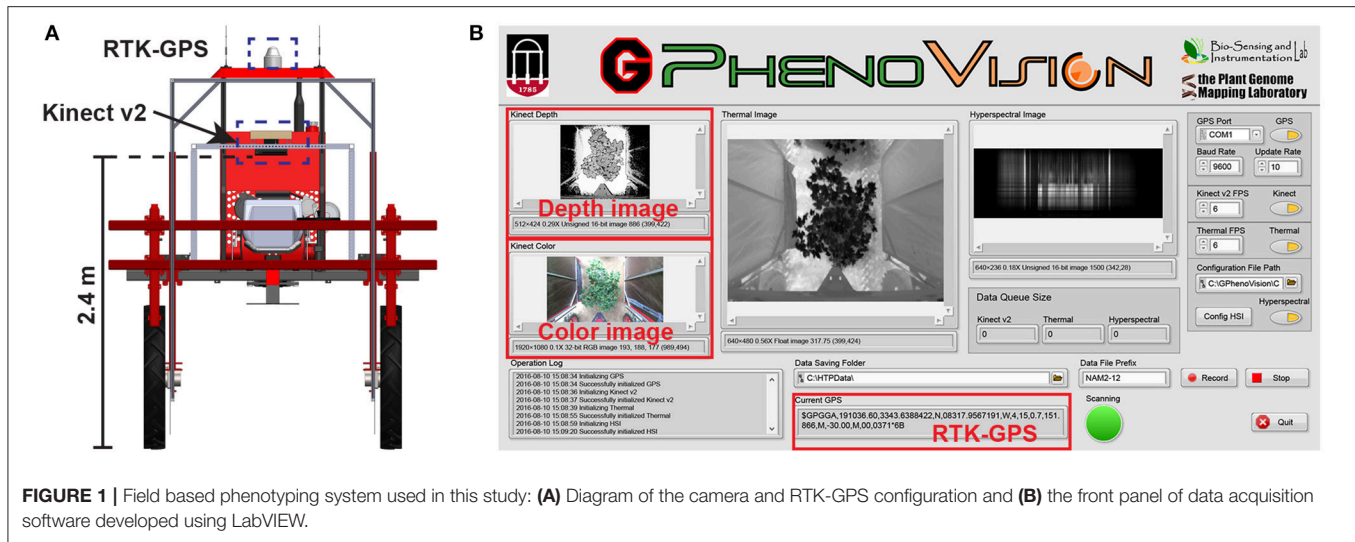
2.2.1.2. Step 1.2: camera position adjustment

Each plot used both global and local coordinate systems. The global (UTM and altitude) system indicated exact positions of image acquisition (also image center), whereas the local system represented image pixel coordinates in which the y-axis was aligned with the vehicle moving (row) direction and the x-axis was aligned with the direction that is perpendicular to the vehicle moving (across-row) direction. Global (UTM and altitude) coordinates were converted to local coordinates, so image acquisition positions were aligned with image pixel coordinates. First, global coordinates were rotated to be aligned with y-axis in the local system using Equation (1).

$$\begin{bmatrix} p_x^i \\ p_y^i \\ p_z^i \end{bmatrix} = \begin{bmatrix} \cos \theta_{G-L} & \sin \theta_{G-L} & 0 \\ -\sin \theta_{G-L} & \cos \theta_{G-L} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{easting}^i - p_{easting}^1 \\ p_{northing}^i - p_{northing}^1 \\ p_{altitude}^i - p_{altitude}^1 \end{bmatrix} \quad (1)$$

where θ_{G-L} was the rotation angle (the angle between the fitting curve and north axis) in radians, p_x^i , p_y^i , p_z^i were the x, y, z coordinates of the acquisition position of the i th frame in the local system, and $p_{easting}^i$, $p_{northing}^i$, and $p_{altitude}^i$ were the UTM coordinates and altitude values, respectively, of the acquisition position of the i th frame in the global system. p^1 was the starting (first) frame acquired in each plot.

After the coordinate system conversion, image acquisition positions were aligned with image pixel coordinates and the



starting point of each plot became the origin point in the local system. Individual depth and color images were reconstructed to colorized point clouds using functions provided by the manufacturer's software development kit (SDK).

2.2.1.3. Step 1.3: camera orientation adjustment

The Kinect v2 camera might be slightly tilted during data collection due to uneven terrain, and thus it was necessary to estimate camera orientation in each frame and to correct point cloud offset due to camera orientation changes. As individual frames would be eventually stitched, only a part of the frame, the region of interest (ROI), needed to be processed. To save processing time, camera orientation estimation and adjustment were performed on the ROI of each point cloud. ROIs were selected based on image acquisition positions using Equation (2).

$$\begin{bmatrix} ROI_X^{Lower} & ROI_X^{Upper} \\ ROI_Y^{Lower} & ROI_Y^{Upper} \\ ROI_Z^{Lower} & ROI_Z^{Upper} \end{bmatrix} = \begin{bmatrix} -\frac{W_{enclosure}}{2} & \frac{W_{enclosure}}{2} \\ -\frac{Dist(i,j)-1}{2} & \frac{Dist(i,j)+1}{2} \\ 0 & H_{camera} \end{bmatrix} \quad (2)$$

where ROI_X^{Lower} and ROI_X^{Upper} were the lower and upper limits on a specific axis of an ROI in the point cloud of a single frame, $W_{enclosure}$ was the width of the enclosure (1.52 m in the present study), $Dist(i,j)$ was the absolute distance difference in acquisition position between the i th and j th frames, and H_{camera} was the height of the camera above the ground level (2.4 m in the present study).

In each selected ROI, ground surface was detected using maximum likelihood estimation sample consensus (MLESC) (Torr and Zisserman, 2000), and the normal of the detected ground surface was calculated. Angle differences in the normal between the detected and ideal (X-Y plane) ground surfaces were calculated, and points in the ROI were rotated based on the angle differences with respect to the X-Z and Y-Z planes using

Equation (3).

$$\begin{bmatrix} x_{adj} \\ y_{adj} \\ z_{adj} \end{bmatrix} = \begin{bmatrix} \cos(\theta_{xz}) & 0 & \sin(\theta_{xz}) \\ 0 & 1 & 0 \\ -\sin(\theta_{xz}) & 0 & \cos(\theta_{xz}) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_{yz}) & \sin(\theta_{yz}) \\ 0 & -\sin(\theta_{yz}) & \cos(\theta_{yz}) \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} \quad (3)$$

where θ_{xz} and θ_{yz} were the angles between the detected and ideal ground surfaces with respect to the X-Z and Y-Z planes, and x_{adj} , y_{adj} , and z_{adj} were the adjusted coordinates of the original points (x_o , y_o , and z_o) in the ROI.

If vegetation covered the entire ground surface and resulted in a failure of ground detection, the ground normal of the ROIs was assumed to be the same as the X-Y plane normal ($\vec{n}_{X-Y} = [0, 0, 1]$). As a consequence, there was no adjustment to account for effects caused by camera orientation changes.

2.2.1.4. Step 1.4: stitching of selected ROIs

The final step in the reconstruction section was to stitch all of the selected ROIs into an integrated colorized point cloud for individual plots. Based on camera positions, the transformation of points in ROIs to the local coordinate system would naturally result in stitching of ROIs and produce the colorized point cloud for a plot. The coordinate transformation was conducted using Equation (4).

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{adj} \\ y_{adj} \\ z_{adj} \end{bmatrix} + \begin{bmatrix} p_x^i - p_x^1 \\ p_y^i - p_y^1 \\ p_z^i - p_z^1 \end{bmatrix} \quad (4)$$

where x , y , and z were the coordinates of points in the final colorized point cloud of a plot, and p_x^i , p_y^i , and p_z^i were the coordinates of the acquisition position of the i th frame in the local coordinate system.

2.2.2. Segmentation of Cotton Plant Canopies

Point clouds of cotton plots could contain irrelevant objects such as system frames, ground surface, and weeds. Therefore,

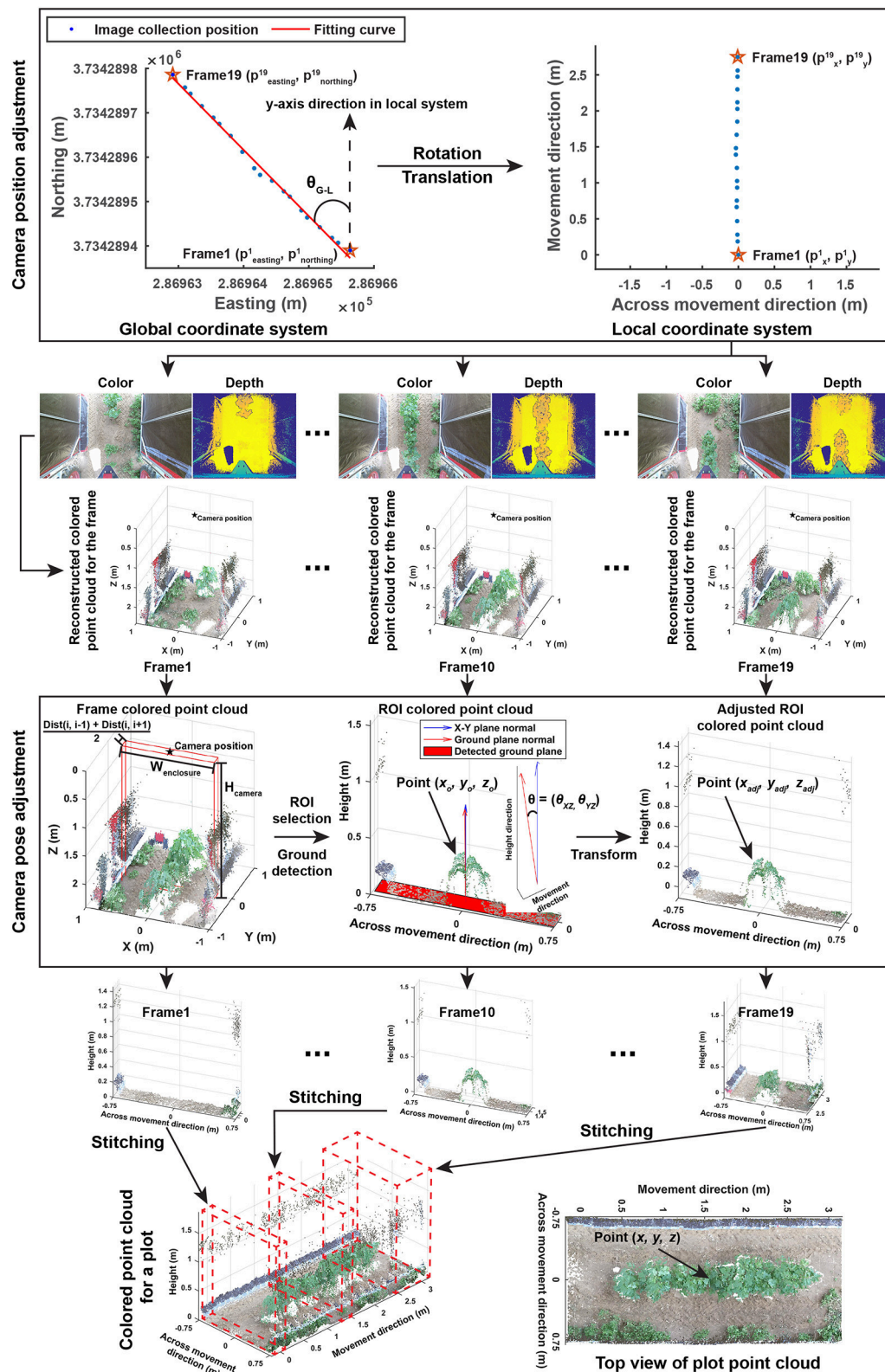


FIGURE 2 | Processing pipeline of reconstructing colorized point clouds for individual plots using color and depth images and GPS collected by the GPhenoVision system.

it was necessary to segment the cotton plant canopy from irrelevant objects (Cotton canopy segmentation in **Figure 3**). The segmentation of cotton canopies involved two steps:

2.2.2.1. Step 2.1: vegetation segmentation

Unlike point cloud data collected by conventional instruments such as LiDAR, point cloud data in the present study also contained color information (RGB values) for each point. A color filter was used to segment vegetation from background objects. According to previous studies (Hamuda et al., 2016), excess green (ExG) index was an effective indicator to obtain vegetation objects in an image, and thus ExG values were calculated for each point of a plot. A preliminary test was performed on a small subset of the collected images, showing that a threshold of 0.15 provided adequate separation between vegetation and background.

2.2.2.2. Step 2.2: cotton canopy segmentation

Color filtering was able to remove most irrelevant objects but not weeds, because plants and weeds are both vegetation and thus hard to differentiate based on color information. In this step, point clouds were rasterized to depth images, so pixels of cotton canopies were differentiated from those of weeds in the depth images by using 2D computer vision algorithms with spatial information including height (depth), area, and position. Identified canopy pixels were back-projected to 3D space to select points representing cotton canopies in point clouds. Point cloud data (x , y , and z coordinates) were rasterized to depth images using Equation (5).

$$\begin{aligned}
 I_D(i, j) &= \max(z_k), \\
 \forall k \in \{k | x_{Lower} + S_g \times (i - 1) &\leq x_k \leq x_{Lower} + S_g \times i\} \\
 &\cap \{k | y_{Lower} + S_g \times (j - 1) \leq y_k \leq y_{Lower} + S_g \times j\} \\
 i &= 1, 2, \dots, \lceil \frac{Range(x)}{S_g} \rceil \\
 j &= 1, 2, \dots, \lceil \frac{Range(y)}{S_g} \rceil
 \end{aligned} \quad (5)$$

where I_D represented a rasterized depth image, and i and j were pixel indices in I_D . x_{Lower} and y_{Lower} were the lower limits of vegetation point clouds, and S_g was the grid size (0.05 m in the present study). x , y , and z were coordinates of vegetation points, and k was the index of a point in vegetation point clouds. $Range(\cdot)$ was a function to calculate the maximum length along a particular axis in vegetation point clouds, and $\lceil \frac{Range(\cdot)}{S_g} \rceil$ was the total number of grid cells along a particular axis.

Most in-row weeds were short and small, and removed using height and area information. Depth images were binarized by the 30th percentile of depth value of all pixels in order to exclude those representing short weeds. Connected components (CCs) were labeled in binary depth images, and small CCs representing weeds were removed using Equation (6). Thresholds of depth and

area were empirical values based on observations of height and area of weeds in the present study.

$$CC_i^{included} = \begin{cases} 1, & Area(CC_i) > Th_{Area} \\ 0, & \text{otherwise} \end{cases}, i = 1, 2, \dots, n \quad (6)$$

where $CC_i^{included}$ is a flag for a connected component, and “1” or “0” indicated inclusion/exclusion of the connected component in the binary depth image. $Area(\cdot)$ was a function to count the number of pixels in a connected component, and Th_{Area} was the threshold (15 pixels in the present study) for including a CC.

Weeds between plots were tall and large, and removed using position information. The CC with the largest area (pixel counts) was selected as the main canopy in a plot, and its bounding box was calculated. Based on their positions relative to the main canopy, remaining CCs were classified as weed or cotton canopy using Equation (7).

$$CC_i^{class} = \begin{cases} 1, & Dist(CC_i, CC_m) \leq W_{CC_m^b}, \forall i \in \{i | CC_i^{included} = 1\} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where CC_i^{class} is the class marker for a CC with “1” for cotton canopy and “0” for weeds. $Dist$ was a function to calculate the distance between the center position of the i th CC and main canopy CC (CC_m), and $W_{CC_m^b}$ was the width of the bounding box for main canopy CC.

According to the rasterization process, a pixel in depth images represented a grid cell in vegetation point clouds. Based on Equation (5), each pixel in the identified canopy CCs was back-projected to a grid cell, and vegetation points in that cell needed to be included in canopy point clouds. Vegetation points were selected using Equation (8) to form canopy point clouds that were used for trait extraction in each plot.

$$PtCloud_{canopy} = \{\forall pt | pt \in CC_i\}, i = \{j | CC_j^{class} = 1\} \quad (8)$$

where $PtCloud_{canopy}$ represented point clouds of cotton canopy, pt was a point in vegetation point clouds, CC_i was the i th identified canopy CC, and CC_j^{class} was the class marker of the j th CC in a depth image.

2.2.3. Extraction of Morphological Traits

Morphological traits were extracted from point clouds of cotton plant canopies (Trait extraction in **Figure 3**). Trait extraction included two parts: extraction of static traits from multiple dimensions (one- and multi-dimensional traits) and calculation of dynamic traits (growth rates).

2.2.3.1. Step 3.1: extraction of one-dimensional traits

One-dimensional traits contained canopy height (maximum and mean), cumulative height profile, and width (maximum and mean) at the plot level. The maximum and mean canopy heights were defined as the tallest and average height values of all canopy

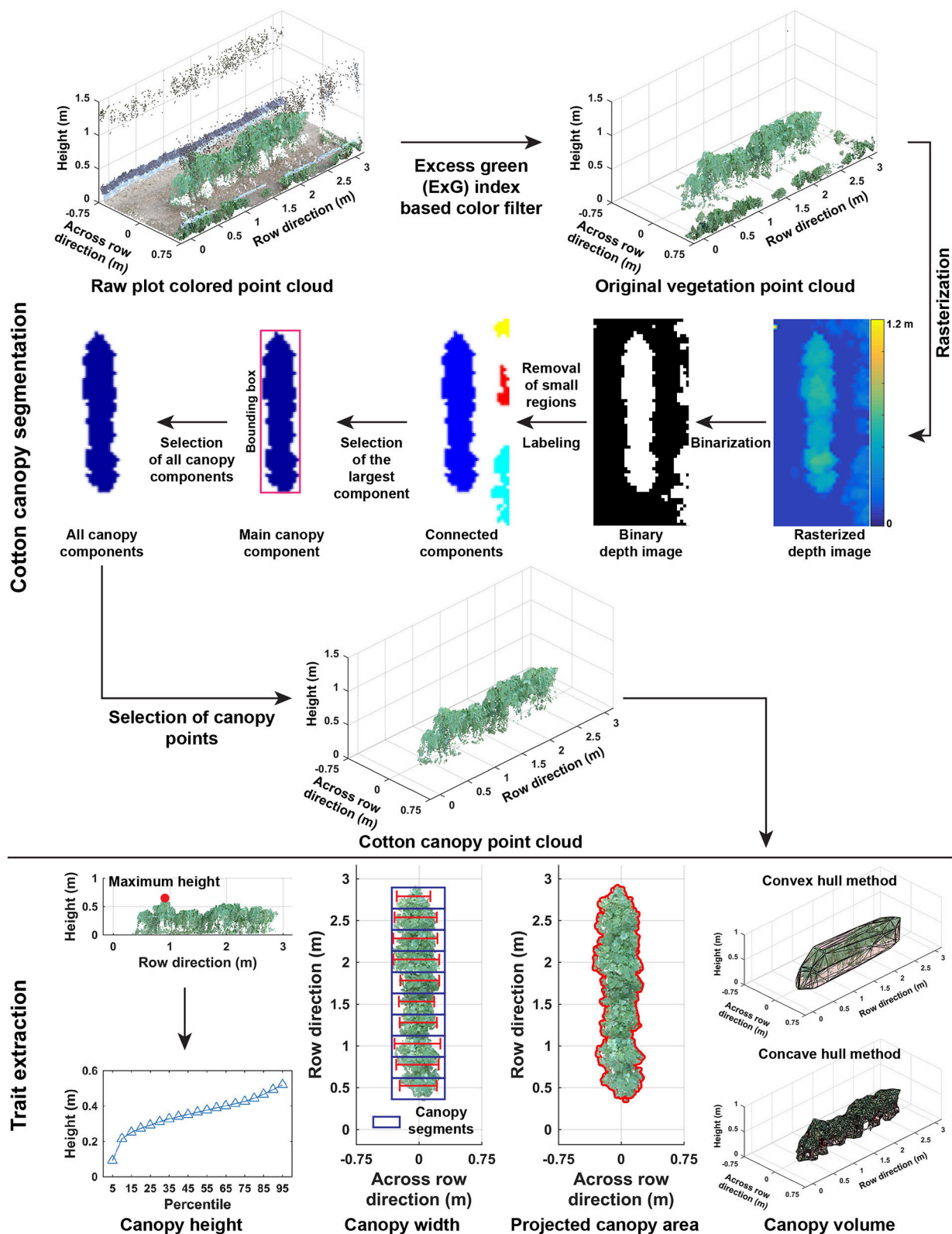


FIGURE 3 | Processing pipeline of segmenting cotton canopy from the reconstructed point clouds and of extracting morphological traits from the canopy point clouds.

points. Cumulative height profile was the combination of canopy height from the 5th percentile to 95th percentile with an interval of 5 %. For canopy width, the point cloud of a cotton canopy was segregated into ten segments along the row direction, and the maximum length across the row direction was calculated in each segment. Maximum and mean widths were the maximum and average values of widths in the ten segments.

2.2.3.2. Step 3.2: extraction of multi-dimensional traits

Multi-dimensional traits contained projected canopy area and canopy volume. All canopy points were projected onto the X-Y plane, and the boundary of the projected shape was identified to calculate the projected canopy area. Convex and concave hulls were detected on canopy point clouds, and canopy volume was estimated using the detected convex and concave hulls.

2.2.3.3. Step 3.3: calculation of growth rate

In addition to static traits on a specific date, growth rates (dynamic changes) could provide information about growth behavior for cotton plants. Growth rate was defined and calculated between every two consecutive data collection dates using Equation (9).

$$G_{T,P} = \frac{T_{d_{last}} - T_{d_{first}}}{d_{last} - d_{first}} \quad (9)$$

where G was the growth rate of a trait (T) during a period (P). d_{first} and d_{last} were the first and last days after planting in the period (P).

2.3. Performance Evaluation

2.3.1. Accuracy of Canopy Segmentation

Cotton canopy segmentation strongly affected the accuracy of trait extraction. In particular, weed removal would significantly influence values of extracted morphological traits. Ground truth segmentation results of depth images were manually generated, and accuracy, false positive rate (FPR; weed pixels identified as canopy), and false negative rate (FNR; canopy pixels identified as weeds) were calculated for each plot. Based on these three values, segmentation results were evaluated and classified into three categories: clean canopy (accuracy $\geq 95\%$), under-removal of weed (FPR $\geq 10\%$), and over-removal of canopy (FNR $\geq 10\%$). The percentages of the three categories were used as indicators to evaluate the performance of canopy segmentation.

2.3.2. Accuracy of Sensor Measurements

It was important to evaluate the accuracy of depth measurement because depth information was not only used for height measurement but also calculation of x and y coordinates of individual points, thereby affecting the overall accuracy of obtained point clouds. As height was directly derived from depth measurement, maximum canopy height was measured for 32 field plots that were randomly selected on each of four scanning dates (DAP 45, 52, 74, and 88). Therefore, a total of 128 data points were obtained for evaluating the accuracy of height measurement.

In addition, due to difficulties in field measurement, potted and artificial plants were used to assess accuracies of measuring other traits. A total of 8 potted plants were used for validating measurements of width, length, and volume, and an artificial plant was used for projected leaf/convex area (see Figure S2). For the potted plants, canopy width (cross-movement direction) and length (movement direction) were measured using a ruler, whereas volume was measured using a protocol for tree volume estimation (Coder, 2000) (see Figure S3). Following the protocol, a plant was virtually segregated into layers with a height interval of 5 cm. Diameters were measured for the middle of individual layers, from which volume could be estimated using a cylinder model, and the plant volume was the summation of all layer volumes. For the artificial plant, a total of 8 layouts were configured to form different plant leaf/convex areas (see Figure S4). In each layout, all leaves were laid on the base surface and imaged with a size marker by a digital single-lens reflex (DSLR) camera (A10, Fujifilm Holdings Corporation, Tokyo, Japan) so that projected leaf/convex area could be accurately measured using image processing. After taking color images, leaves were vertically lifted at various heights to form a 3D layout simulating real plants. It was noteworthy that real plants had many more leaves than the artificial plant, resulting in a denser canopy, and thus projected convex area would be closer to projected canopy area in real situations. In total, 8 data points were used for validating sensor measurements of canopy width, length, projected area, and volume.

Simple linear regression analyses were performed between sensor and manual measurements for all traits. Adjusted coefficient of determination (Adjusted R^2) and root mean squared error (RMSE) were used as indicators for performance assessment.

2.3.3. Efficiency of Image Processing

In addition to accuracy performance, the efficiency of the proposed algorithm was tested on a workstation computer that used an Intel i7-4770 CPU with 16 GB of RAM on a Windows 10 operating system. Processing time of point cloud reconstruction, cotton canopy segmentation, and trait extraction was recorded during the processing of all plot data collected on all eight data collection dates. For point cloud reconstruction, simple linear regression analysis was conducted between the number of frames in each plot and the reconstruction time. For cotton canopy segmentation and trait extraction, the percentage of processing time for each key step was calculated.

2.4. Statistical Analyses between Fiber Yield and Extracted Traits

Extracted traits were grouped into two categories: univariate and multivariate traits. For univariate traits, simple linear regression analyses were conducted between the traits and fiber yield, whereas for multivariate traits, multiple linear regression analyses were conducted between the traits and fiber yield. The adjusted R^2 and RMSE were used to assess the potential of the usefulness of extracted traits for prediction of cotton fiber yield. As regression models established using the maximum and mean canopy height were nested to those using cumulative height profile, F -tests were

TABLE 1 | Performance of the segmentation of cotton canopy.

Date	Clean canopy (%)	Weed under-removal (%)	Weed over-removal (%)
28 July 2016 (DAP 45)	99.2	0.8	0
4 August 2016 (DAP 52)	99.2	0.8	0
19 August 2016 (DAP 67)	95.2	4.8	0
26 August 2016 (DAP 74)	97.6	1.6	0.8
9 September 2016 (DAP 88)	99.2	0.8	0
16 September 2016 (DAP 95)	100	0	0
23 September 2016 (DAP 102)	100	0	0
30 September 2016 (DAP 109)	98.4	0	1.6
All dates	98.6	1.1	0.3

conducted to rigorously test the statistical significance of model differences. All regression and *F*-test analyses were performed in R software (R Core Team, 2016).

3. RESULTS

3.1. Performance of Segmentation of Cotton Canopy

Overall, the cotton canopy segmentation achieved an accuracy of 98.6% on all collected data, which indicates that the scanning system provided accurate canopy point clouds for trait extraction throughout the growing season (Table 1). The proposed algorithms precisely segmented cotton canopy under various plot conditions (Figures 4A–D). In-row weeds were usually shorter and smaller than cotton plants, so they were mostly removed by height and area filters. In contrast, weeds between rows were generally large and tall, but they could be removed using position information because they were located between plots and away from the main canopy in each plot. Therefore, the method could provide accurate segmentation of cotton plant canopy. However, the weed under-removal rate increased noticeably on DAP 67 and DAP 74. This was because no weeding activity was arranged during that period, and many weeds grew substantially and became comparable with cotton plants in height and size; as a consequence, cotton plants became hard to differentiate from weeds (Figure 4E). In late growth stages (e.g., flowering and boll development stage), cotton plant leaves started to shrink and fall down, resulting in a reduction of canopy size and overlap. Large canopies were segregated into small leaf areas, especially along the outer part of canopies. These small leaf areas were largely filtered out by the area filter, leading to over-removal of canopy (canopy pixels identified as weeds) (Figure 4F).

3.2. Representative Colorized Point Clouds

Representative colorized point clouds of cotton canopies were generated and demonstrated at the plot level over the growing season (Figure 5). Poorly germinated plots contained some empty areas (Figure 5A), which tended to be at least partly filled by neighboring plants after the canopy development stage (DAP 74). In well germinated plots, seeds generally sprouted around the same time and seedlings grew at comparable speeds,

maintaining similar canopy height along a plot. Even in cases with some variation of germination date (thus development speed), canopy height tended to reach a similar level in a plot, avoiding the occurrence of extremely tall or short sections (Figure 5B). Well germinated plots showed canopy overlap at an earlier date (approximately DAP 52 or earlier) than poorly germinated ones and were relatively flat (or slightly arched) along the row direction, whereas in poorly germinated plots the canopy close to empty areas was short. This slightly arched shape of the canopy is due to “edge effect,” and plants can be either bigger or smaller at the plot edge than those in the center.

3.3. Accuracy of Sensor Measurement

Sensor measurements were strongly correlated (adjusted $R^2 > 0.87$) with manual measurements for all extracted traits (Figure 6). In particular, height measurements were obtained from multiple days and the RMSE was 0.04 m, suggesting a high accuracy and repeatability of depth measurement (and therefore point clouds) in field conditions. Consequently, point clouds acquired by the Kinect v2 sensor could be used to accurately measure traits such as width, length, projected leaf/convex area. It should be noted that although correlation was strong (adjusted $R^2 > 0.87$) between sensor and manual measurements for volume, the absolute values were significantly different (see the slope in regression equations for convex and concave hull volumes). Convex and concave hull volumes were smaller than reference measurements. In the present study, the Kinect v2 sensor acquired point clouds from the top view, and thus only canopy surface could be imaged, resulting in a lack of information from two sides (especially sections under the canopy). As a consequence, the volume of a plant (or a plot) estimated by convex and concave hulls was a portion of the ground truth value. Since the canopy of most cotton plants was roughly a cone, the sensed portion could generally represent the entire plant. This was also supported by the high correlation between sensor and reference measurements. In addition, convex hull volume showed better correlation with reference measurement than concave hull volume, because both convex hull and reference measurements included volume of empty space among branches of a plant. Nonetheless, as both convex and concave hull demonstrated strong correlation, they could be used as indicators for growth pattern analyses and/or yield prediction.

3.4. Efficiency of the Proposed Approach

The proposed approach used on average 215 s for processing a plot, including 184 s for point cloud reconstruction and 31 s for canopy segmentation and trait extraction (Figure 7). Variations in reconstruction time primarily came from the different number of frames acquired for a plot (Figure 7A). Although the GPhenoVision system was set at a constant speed, the actual system speed would vary due to different terrain conditions (dry vs. wet) and scanning status slower at the two ends of a row than in the middle). Consequently, the number of frames could be different in various plots on different dates. Generally, the reconstruction time increased linearly with the number of frames acquired in a plot (Figure 7B). However, for plots with equal number of frames,

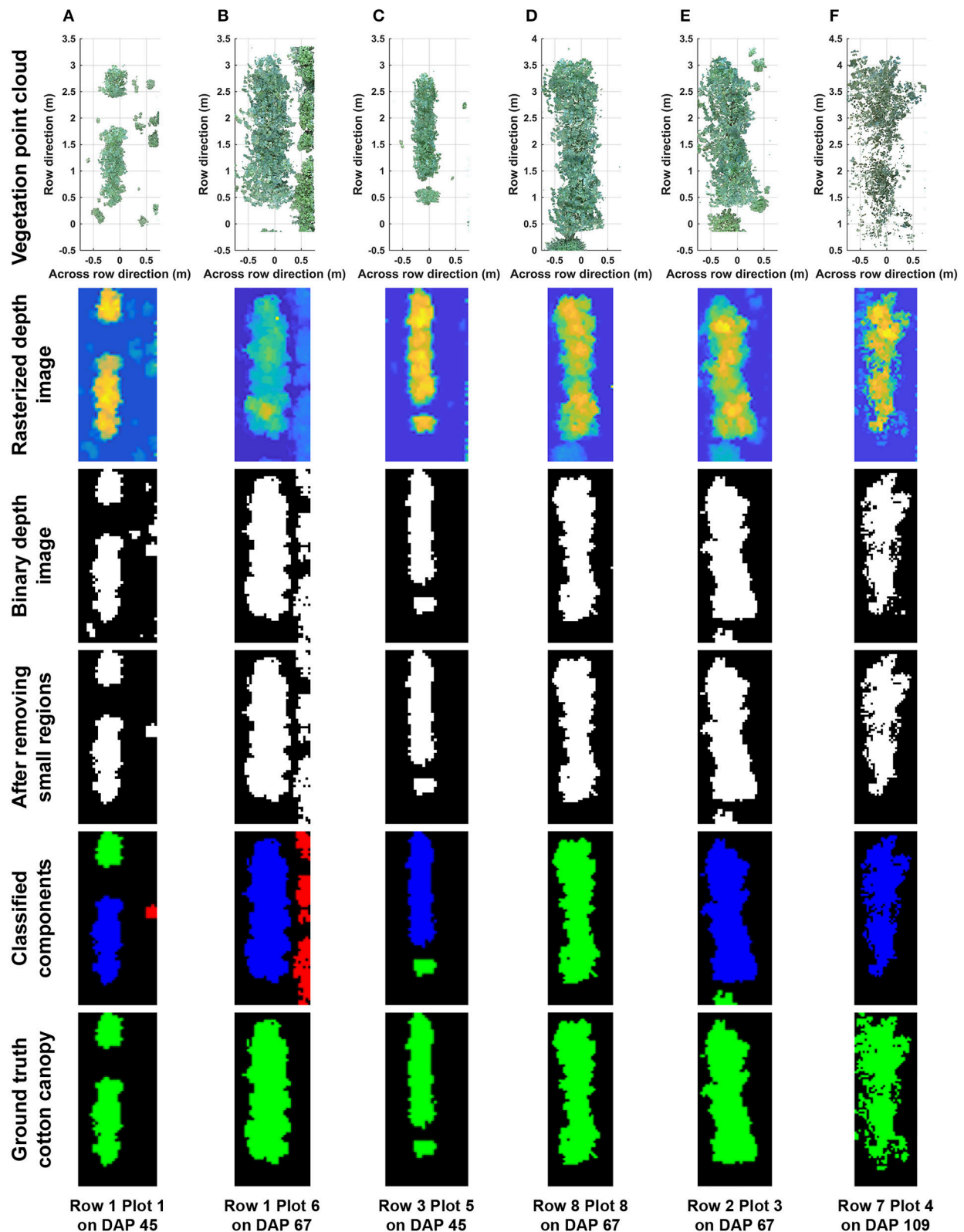
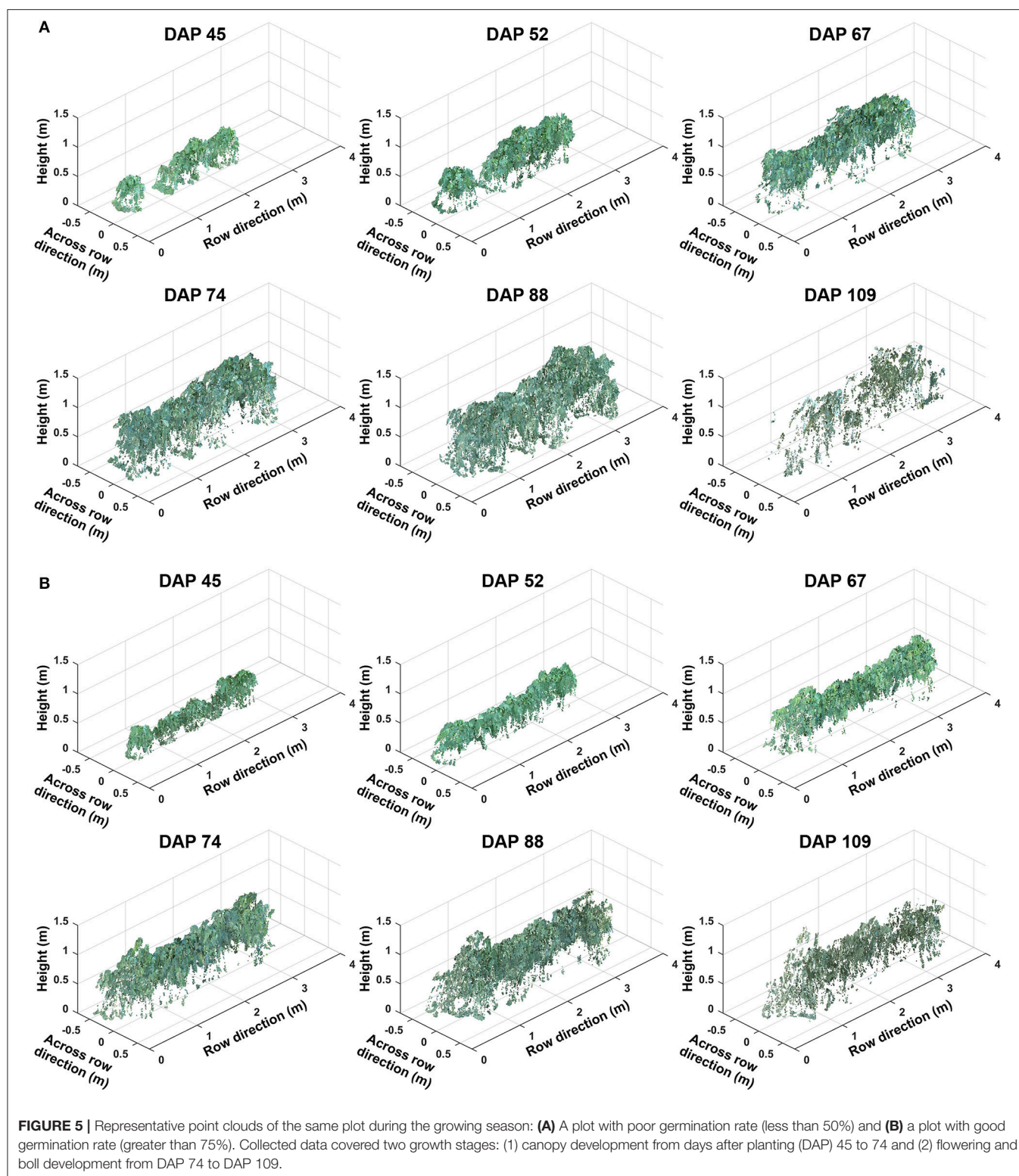


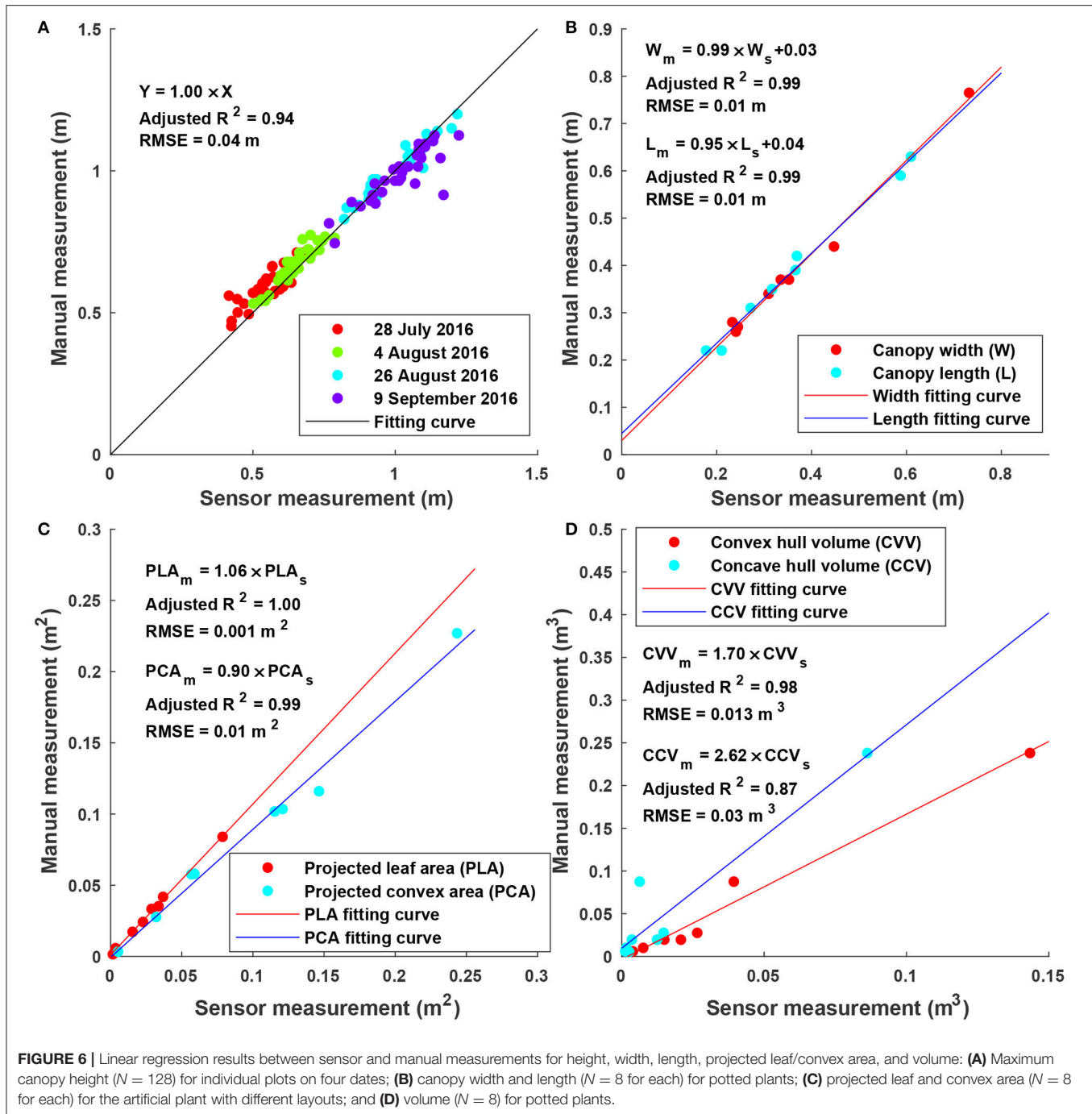
FIGURE 4 | Representative results of cotton canopy segmentation: successful segmentation under (A) poorly germinated plot with weeds, (B) well germinated plot with weeds, (C) plot with a segregated plant, and (D) well germinated plot with connected weeds; (E) failure of weed removal; and (F) over-removal of cotton plant canopy. In classified connected component images, blue, green, and red colors indicated main canopy, cotton canopy, and weeds, respectively. Ground truth images were manually generated for including cotton canopy.



variations in processing time were mainly due to different number of points in frames. For instance, point clouds of poorly germinated plots contained more points of ground than those of well germinated plots, resulting in variation

in ground surface detection and thus in reconstruction time.

For canopy segmentation and trait extraction, efficiency variations occurred due to different numbers of points in point



cloud data (Figure 7C). The number of points in a point cloud was determined by the canopy size, which was affected by both germination condition and growth stage. Plant canopies were larger in well germinated plots or plots in the canopy development stage than those in poorly germinated plots or plots in the flowering and boll development stage. Thus, the number of canopy points varied from plot to plot and stage to stage, resulting in processing time variations. Among the three operations, canopy segmentation and trait extraction were more time consuming than vegetation segmentation (Figure 7D).

Vegetation segmentation was based on color filtering, in which all points were processed simultaneously. In contrast, canopy segmentation and trait extraction processed information point by point, thus using a significantly longer time than vegetation segmentation.

3.5. Growth Trend of Cotton Plants

3.5.1. Static Traits

Overall, plants elongated substantially from DAP 45 to DAP 74 (canopy development stage) and expanded considerably from

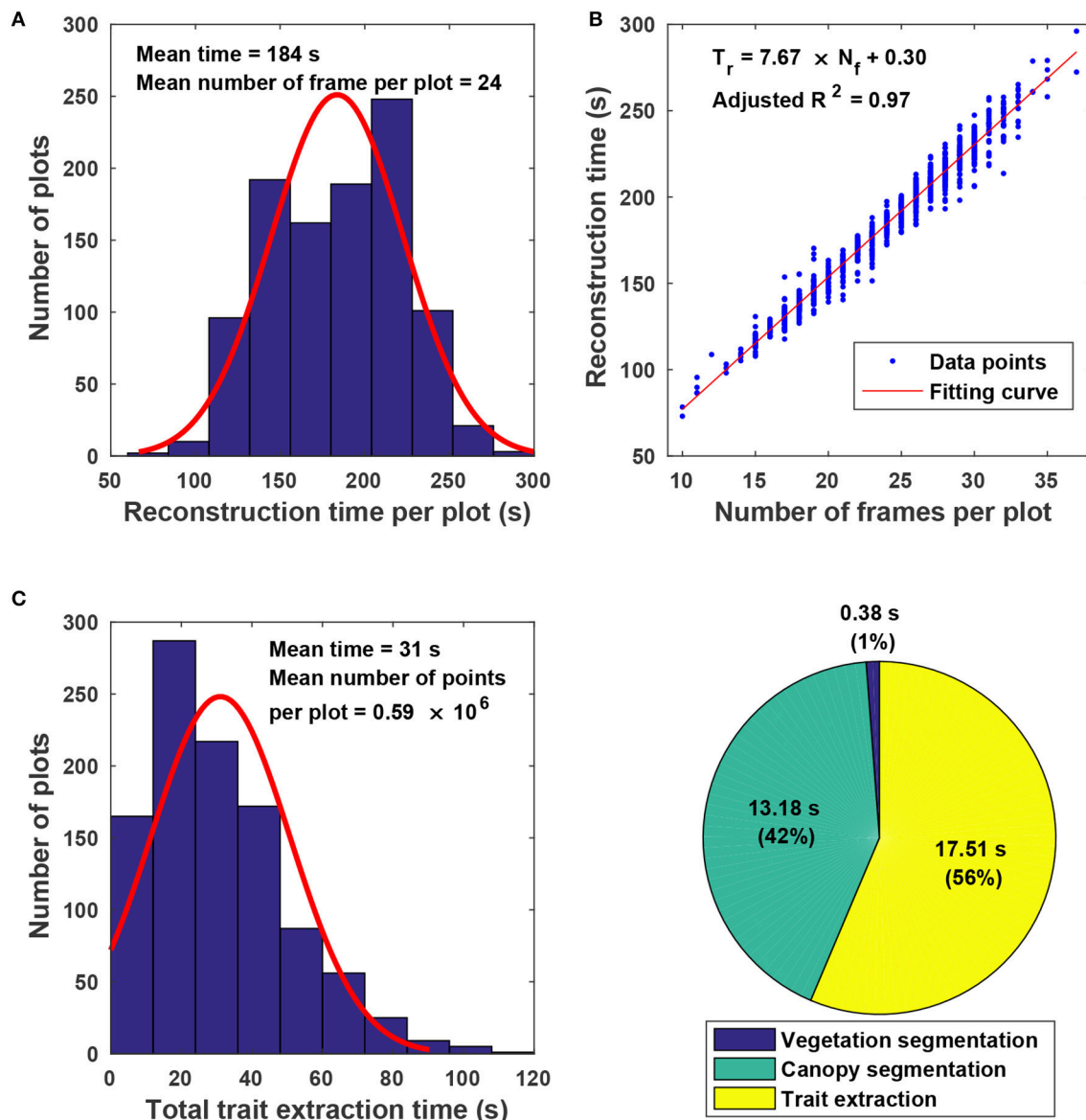


FIGURE 7 | Algorithm efficiency of point cloud stitching and trait extraction: **(A)** Histogram of reconstruction time for individual plots; **(B)** relationship between the number of images collected in each plot and the reconstruction time of stitching; **(C)** histogram of the total trait extraction time for individual plots; and **(D)** percentages of processing time for various steps in canopy segmentation and trait extraction.

DAP 45 to DAP 88 (canopy development stage and early flowering stage) (**Figure 8**). The values of canopy morphological traits gradually decreased during the rest of the growing season (flowering and boll development stages). Among the genotypes, the Americot conventional showed the shortest and smallest canopy, compared with the other three experimental genotypes. Genotype GA2009037 was the tallest genotype but with the least width, indicating a tall and narrow canopy (statistical comparison results are provided in Tables S1–S7).

Cumulative height profiles of all four genotypes showed logarithmic growth during the canopy development stage (DAP 45 to DAP 74) and linear growth during the flowering and boll

development stage (DAP 74 to DAP 109) (**Figure 9**). In the canopy development stage, plant leaves increase in transpiration capability and maintain a solid shape to receive more sunlight for photosynthesis. Therefore, at the canopy development stage, upper leaves occluded lower ones, resulting in less variation in canopy height after the 25th percentile. The large difference between the 5th and 25th percentiles was due to un-occluded leaves on expanded branches at lower positions along the canopy border. When transitioning into flowering and boll development stage, plant leaves receive fewer nutrients (due to nutrient demands of maturing fruit) and start to shrink. Previously occluded leaves could be captured, and thus canopy height was

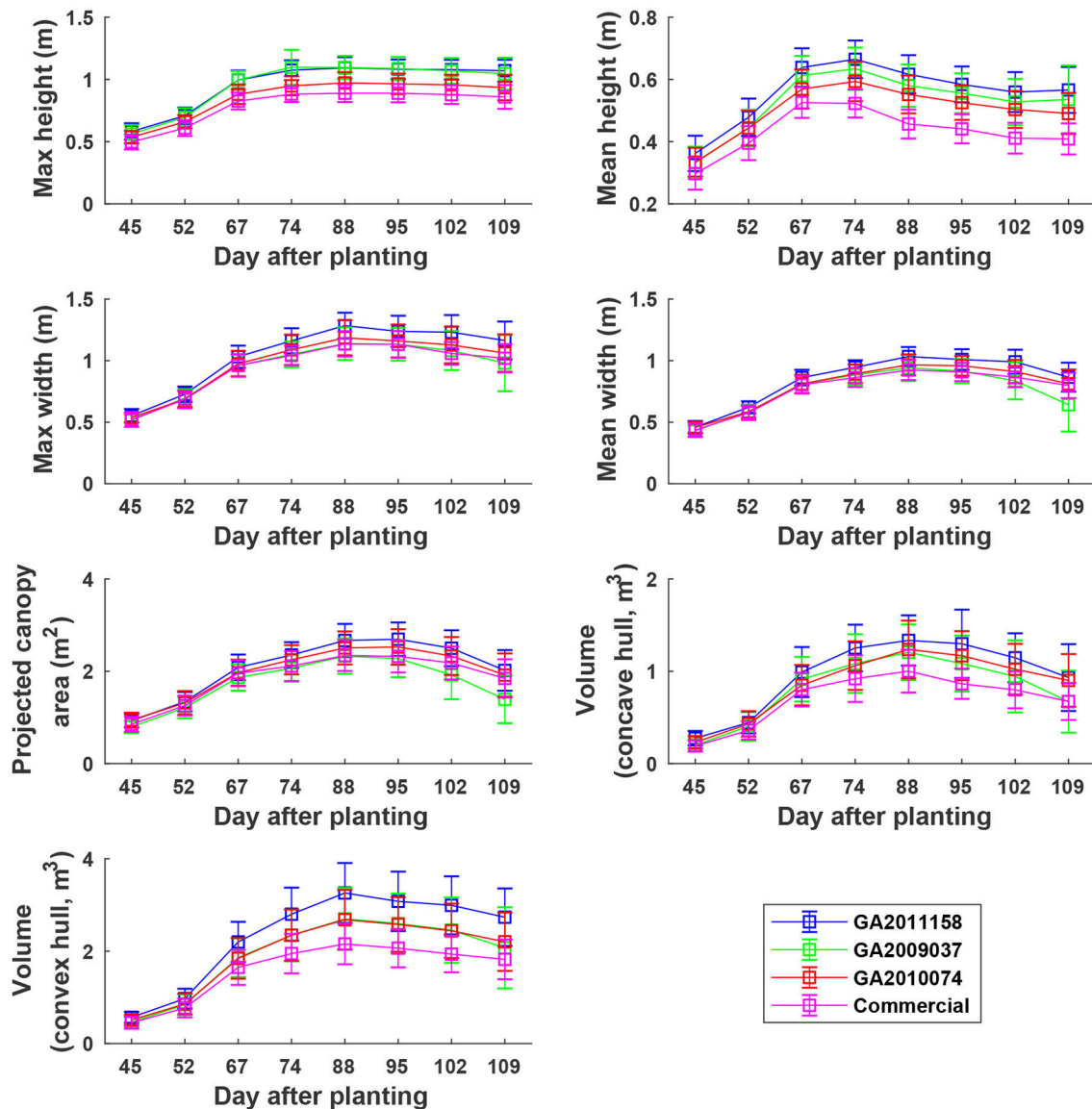


FIGURE 8 | Extracted static traits of individual plots in ENGR field during the growing season. Extracted traits covered two growth stages: (1) canopy development from DAP 45 to DAP 74 and (2) flowering and boll development from DAP 74 to DAP 109.

represented by leaves at a wider range of positions, resulting in an increase of height difference at various percentiles. This information might be useful for predicting the change of growth stages for cotton plants.

3.5.2. Dynamic Traits

The four genotypes showed different patterns of canopy development in different growth stages (Figure 10). Generally, growth rates of all four genotypes approached zero in P5 (DAP 74 to DAP 88, which was the first data collection period in the flower and boll development stage) for maximum and mean canopy height, and in P6 (DAP 88 to DAP 95, which was the second data collection period in the flower and boll development stage) for the other traits, because the four

genotypes were short season cottons and had similar growth stage transitions. In addition, all genotypes stopped canopy elongation approximately a week earlier than canopy expansion. Although the overall growth trend was similar, there were differences in growth behavior among the four genotypes (statistical comparison results are provided in Tables S8–S14). The Americot conventional was generally the slowest in developing its canopy, whereas the three experimental genotypes grew rapidly in different dimensions. Genotype GA2011158 showed the fastest growth in all dimensions, resulting in large canopy height, width, projected area, and volume. Genotype GA2009037 primarily elongated rather than expanded, resulting in tall canopies with the least projection coverage, whereas Genotype GA2010074 mainly expanded rather than elongated, resulting in short

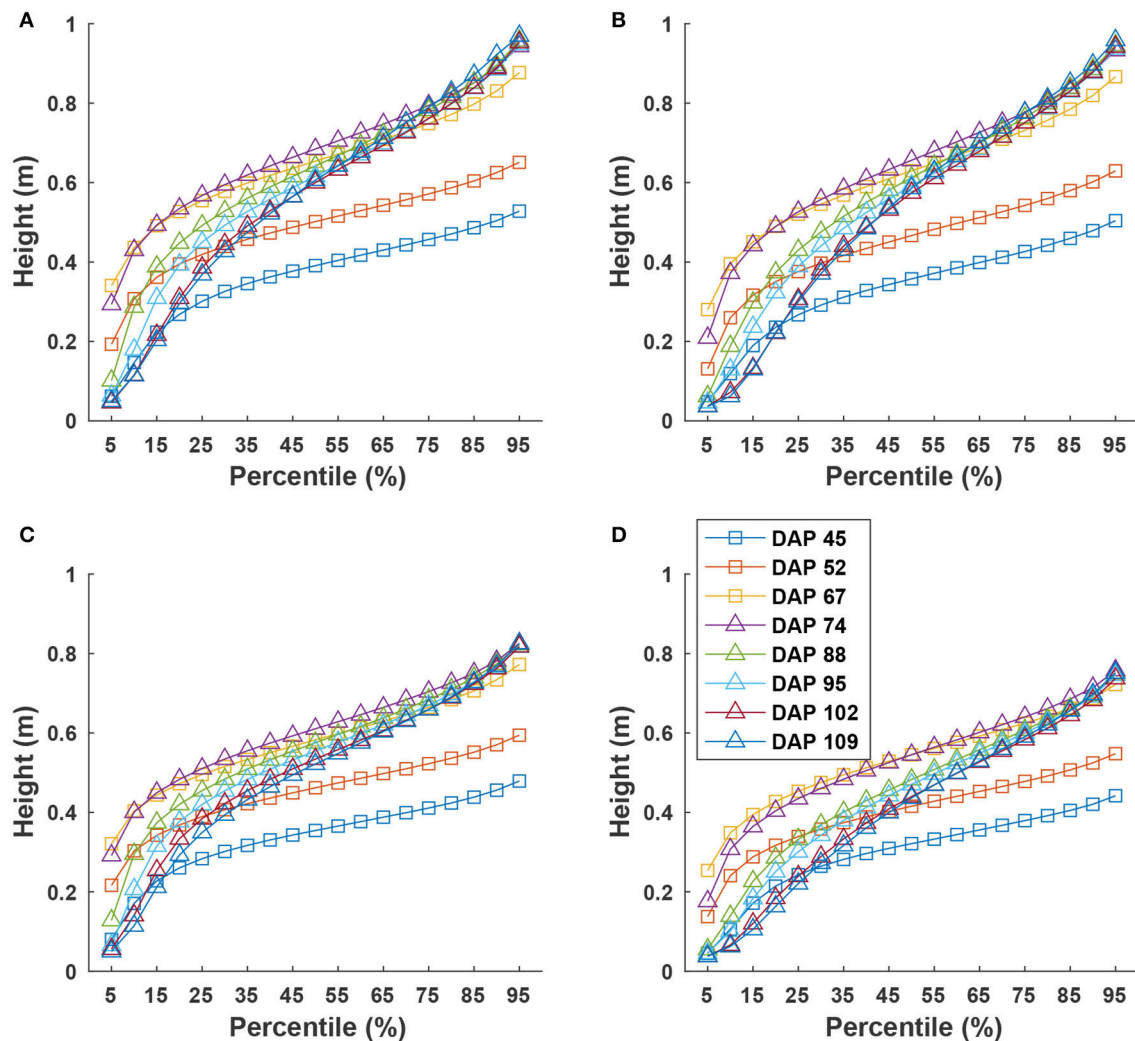


FIGURE 9 | Extracted canopy height at various percentiles during the growing season: (A–D) were for GA2011158, GA2009037, GA2010074, and commercial variety. Extracted traits covered two growth stages: (1) canopy development from DAP 45 to DAP 74 and (2) flowering and boll development from DAP 74 to DAP 109.

canopies with large coverage and volume. Detection of such variations might permit identifying genes controlling cotton plant growth patterns and/or selecting genotypes suitable for different production or harvesting strategies.

3.6. Performance of Yield Prediction

In general, static traits had some value ($R^2 = 0.12$ – 0.71) for predicting cotton fiber yield (Figure 11A). Among univariate traits, multi-dimensional traits (e.g., projected canopy area and volume) considerably outperformed one-dimensional traits such as canopy height and width, presumably because multi-dimensional traits could depict canopy size more completely. For instance, genotypes GA2011158 and GA2009037 had similar canopy height but different fiber yield, resulting in a low correlation between canopy height and fiber yield. However, the two genotypes had obvious differences in projected canopy area and canopy volume, with improved correlation between

the morphological traits and fiber yield. This indicated the usefulness of extracting multi-dimensional traits using advanced 3D imaging techniques. Compared with univariate traits (the maximum and mean canopy height), multivariate traits (cumulative height profile) showed a significant improvement in yield prediction on most days (see detailed *F*-test results in Table S15). This was because cumulative height profiles intrinsically incorporated spatial distribution: high percentile ranks represented height information in the middle of the canopy, whereas low percentile ranks represented the borders. The capability of extracting multivariate traits was an advantage of using 3D imaging modalities. Most static traits achieved the highest correlation with fiber yield after the canopy development stage (DAP 67), and then the correlation decreased slightly during the rest of the growing season. In the growing season that was studied, the best period to use morphological traits for yield prediction was the transition period between canopy

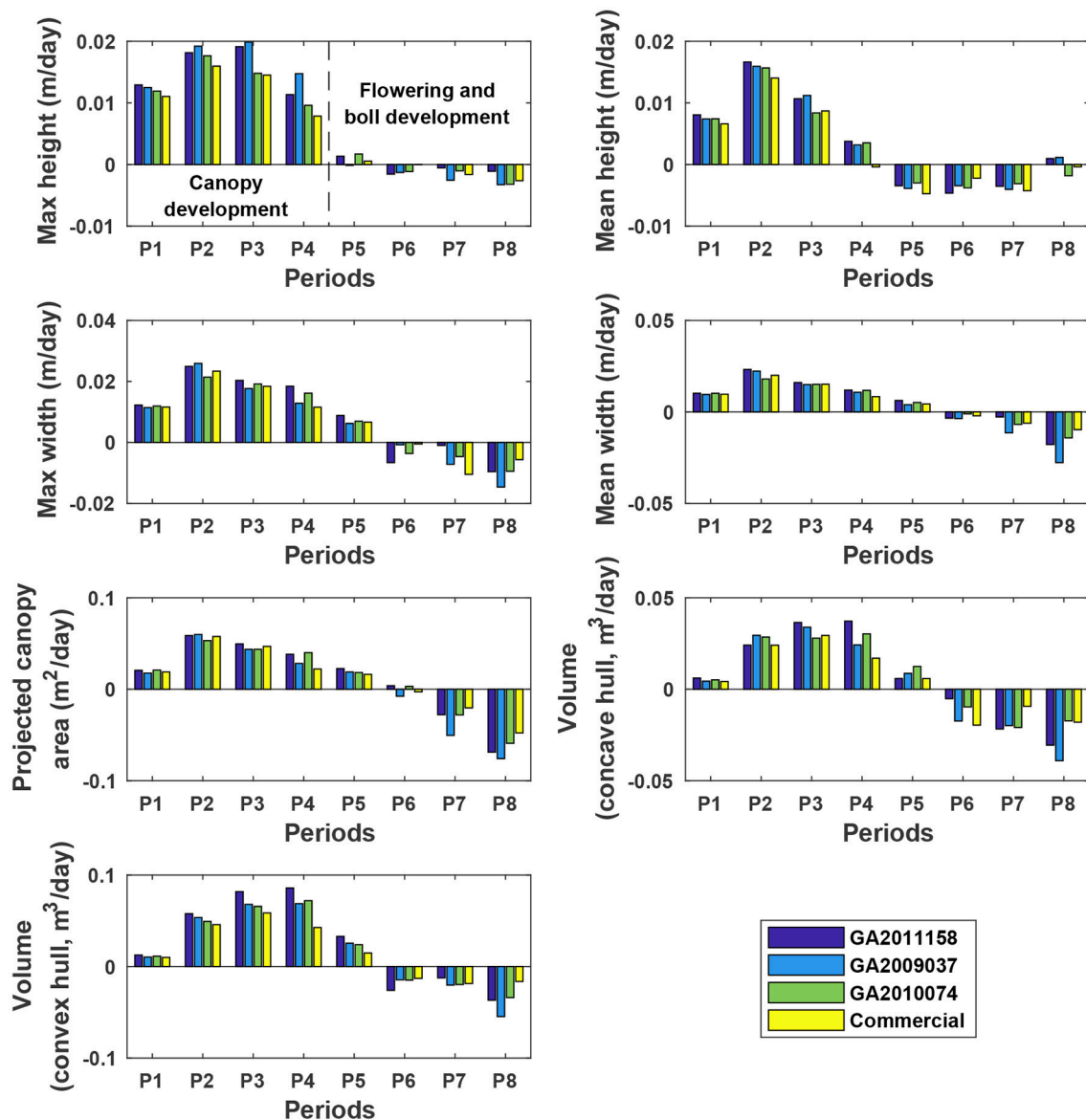


FIGURE 10 | Extracted growth rates (dynamic traits) of individual plots in ENGR field during the growing season. P1 was the period from the day of planting to DAP 45, and P2–P8 were the periods between two consecutive data collection dates. Extracted traits covered two growth stages: (1) canopy development from P1 to P4 (DAP 67 to DAP 74) and (2) flowering and boll development from P4 to P8 (DAP 102 to DAP 109).

development and flowering and boll development stages. Canopy size during the transition period reached the maximum size and reasonably represented the potential of flower and boll development. However, mean canopy width and projected canopy area achieved the highest correlation at later dates (DAP 95 and DAP 102) because all genotypes kept expanding until DAP 95 and most of the expansion was due to the development of sympodial (fruiting) branches.

Compared with static traits, dynamic traits (growth rates) demonstrated less capability for yield prediction, but multi-dimensional traits that could comprehensively quantify canopy

size still outperformed one-dimensional traits (**Figure 11B**). This further supported the advantage of using 3D imaging for measuring morphological traits. In contrast to static traits, growth rates were more informative for yield prediction in early stages (early periods in canopy development stage) rather than late stages. This may have been because all four genotypes were short season cotton and had similar overall growth trends. In early stages, growth rates were positive, reflecting canopy development, and were well correlated with the final vegetative structure and productive capacity. However, in late stages, growth rates were negative, reflecting leaf and

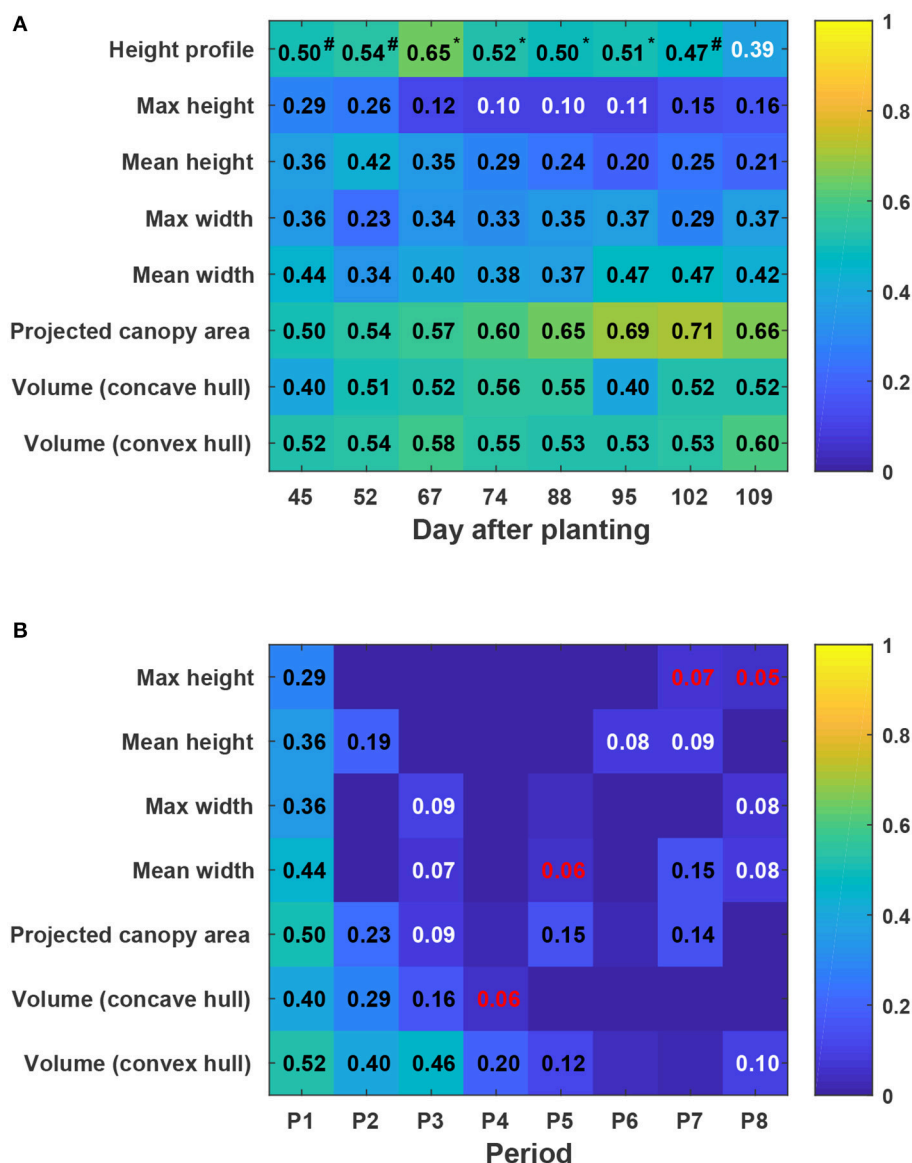


FIGURE 11 | Coefficients of determination (R^2) of linear regressions between extracted traits and cotton fiber yield: **(A)** Static traits and fiber yield and **(B)** dynamic traits and fiber yield. Black, white, and red colors indicated statistical significance of “<0.001,” “<0.01,” and “<0.05,” respectively. Insignificant R^2 values were not shown in blocks. P1 was the period from the day of planting to DAP 45, and P2–P8 were the periods between two consecutive data collection dates. Superscripts of R^2 values for height profile denoted the F -test results between regression models: a number sign (#) indicated a significant model improvement by using height profile rather than maximum height, and an asterisk (*) indicated a significant model improvement by using height profile rather than both maximum and mean heights (see Table S15 for detailed F -test results). In addition, Pearson’s correlation coefficients were calculated for all univariate traits (see Figure S5).

plant senescence, and were not well correlated with fiber yield.

4. DISCUSSION

Consumer-grade RGB-D image-based measurement of plant canopies in field conditions could lower the cost and increase the use of imaging techniques in phenotyping, benefitting the plant science community. The entire system used in the present study cost \$70,500, including an RGB-D

camera (\$200), high-clearance tractor platform (\$60,000), RTK-GPS (\$8,000), power unit (\$300), and rugged laptop (\$2,000). This is still a considerable investment, but can be reduced. The tractor platform can be replaced with a low-cost pushcart (usually less than \$2,000) if experiments are less than a hectare (Bai et al., 2016). In addition, a pushcart system allows for “stop-measure-go” mode, in which image acquisition locations can be predefined and manually controlled, and thus RTK-GPS is not required. However, these low-cost solutions decrease the efficiency and

throughput of data collection. The consumer-grade RGB-D camera is an inexpensive 3D sensing solution for plant phenotyping, but building a phenotyping system requires comprehensive consideration of balance between research budget, experimental scale, and data collection and processing throughput.

The processing algorithms reported here can accurately extract morphological traits of plant canopies at the plot level, providing useful information for genomic studies and/or breeding programs. However, we note two limitations of the proposed approach. First, the Kinect v2 camera cannot be directly used under strong illumination (e.g., during midday in the field), so a mechanical structure is required to provide shade. Second, the reconstruction step currently relies on functions provided by Kinect v2 SDK, and must be performed on a computer running on an operating system of Windows 8 (or later) with connection to the Kinect v2 camera. This may significantly decrease the post-processing throughput, because cloud computing services usually run on remote Unix/Linux systems and users cannot connect hardware components to them while processing. This limitation could be addressed with third party libraries or user-performed registration between depth and color images (Kim et al., 2015). We acknowledge that the developed method was tested in a 1-year experiment, and altered conditions may change results. For example, predictors of yield were most effective near the end of canopy development in this experiment—however, dramatic differences in growing conditions (such as drought or cold) could change that in other years. Additional experiments might also consider different degrees of replication—here, a total of 32 replicates per genotype were used, which is considerably more than the number that is generally used in genetics/genomics studies. Variability among these samples provides a basis for estimating the minimal number of samples that might be used to discern differences between treatments of pre-determined magnitude, which is important in experiments involving large numbers of genotypes such as breeding programs. Based on theoretical calculation (Cochran and Cox, 1992), most static traits required three replicates, whereas most dynamic traits needed more than ten replicates (see Figure S6). This is because the four genotypes used in this study are very similar in growth patterns, requiring a high number of replicates to increase the statistical power for genotype differentiation. However, the minimal number of replicates could be reduced for growth rates in experiments involving genotypes with distinctive growth patterns (e.g., wild and elite cotton germplasm lines).

Two important findings were observed for the extracted traits. Firstly, static and dynamic traits showed the highest correlation with fiber yield in different periods; dynamic traits were informative in early canopy development stages, whereas most static traits were useful in the transition period between canopy development and flower and boll development. Canopy size (canopy height, width, projected area, and volume) remained relatively constant near its peak for a period after canopy development, indicating to some extent the capability to develop flowers and bolls. However, growth rates were negative in late stages, reflecting plant senescence as resources were redirected to the maturing bolls. Secondly, multivariate

traits consistently showed better yield prediction than univariate traits, presumably because the multivariate traits intrinsically incorporated information such as the spatial variation of canopy height. Several previous studies reported methods to find the best percentile of canopy height to use to increase the accuracy of yield prediction (Friedli et al., 2016; Weiss and Baret, 2017). However, according to our results, it may be more efficient to explore the use of multivariate traits to predict yield. It is noteworthy that environment is an important factor, and analysis of these parameters in different growing seasons with various environments may yield different results. These differences may help to better explain interactions between genotype and environment ($G \times E$), and make breeding programs more effective.

5. CONCLUSIONS

The 3D imaging system and data processing algorithms described here provided an inexpensive solution to accurately quantify cotton canopy size in field conditions. The extracted morphological traits showed potential for yield prediction. Multidimensional traits (e.g., projected canopy area and volume) and multivariate traits (e.g., cumulative height profile) were better yield predictors than traditional univariate traits, confirming the advantage of using 3D imaging modalities. Early canopy development stages were the best period in which to use dynamic traits for yield prediction, whereas late canopy development stages were the best period in which to use static traits. Future studies will be focused on improving the data processing throughput (via methods such as parallel computing) and extracting new multivariate traits for canopy size quantification.

AUTHOR CONTRIBUTIONS

YJ, CL, and AP conceived the main concept of the study; YJ, JR, SS, RX, AP, and CL designed the experiment and performed field data collection; YJ developed the data processing pipeline; YJ, CL, AP, SS, and RX analyzed the data; AP and CL provided plant materials, instruments, and computing resources used in the study; YJ, CL, and AP contributed to writing the manuscript. All authors reviewed the manuscript.

FUNDING

This study was funded jointly by the Agricultural Sensing and Robotics Initiative of the College of Engineering, and the College of Agricultural and Environmental Sciences of the University of Georgia. The project was also partially supported by the National Robotics Initiative (NIFA grant No: 2017-67021-25928) and Georgia Cotton Commission.

ACKNOWLEDGMENTS

The authors gratefully thank Dr. Peng Chee (University of Georgia) for providing plant materials and insights to

the present study; and Mr. Zikai Wei, Mr. Jamal Hunter, Mr. Aaron Patrick, Mr. Jeevan Adhikari, Ms. Mengyun Zhang, Mr. Shuxiang Fan, and Dr. Tariq Shehzad for their assistance in planting, field management, data collection, and harvesting.

REFERENCES

- Andújar, D., Dorado, J., Fernández-Quintanilla, C., and Ribeiro, A. (2016). An approach to the use of depth cameras for weed volume estimation. *Sensors* 16:972. doi: 10.3390/s16070972.
- Andújar, D., Dorado, J., Bengochea-Guevara, J. M., Conesa-Muñoz, J., Fernández-Quintanilla, C., and Ribeiro, A. (2017). Influence of wind speed on RGB-D images in tree plantations. *Sensors* 17:E914. doi: 10.3390/s17040914
- Andujar, D., Ribeiro, A., Fernandez-Quintanilla, C., and Dorado, J. (2016). Using depth cameras to extract structural parameters to assess the growth state and yield of cauliflower crops. *Comput. Electr. Agricult.* 122, 67–73. doi: 10.1016/j.compag.2016.01.018
- Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008
- Bai, G., Ge, Y. F., Hussain, W., Baenziger, P. S., and Graef, G. (2016). A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. *Comput. Electr. Agricult.* 128, 181–192. doi: 10.1016/j.compag.2016.08.021
- Barabaschi, D., Tondelli, A., Desiderio, F., Volante, A., Vaccino, P., Vale, G., et al. (2016). Next generation breeding. *Plant Sci.* 242, 3–13. doi: 10.1016/j.plantsci.2015.07.010
- Bréda, N. J. J. (2003). Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. *J. Exp. Bot.* 54, 2403–2417. doi: 10.1093/jxb/erg263
- Cobb, J. N., DeClerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theor. Appl. Genet.* 126, 867–887. doi: 10.1007/s00122-013-2066-0
- Cochran, W. G., and Cox, G. M. (eds.). (1992). “Number of replications,” in *Experimental Designs* (New York, NY: Wiley), 17–22.
- Coder, K. D. (2000). *Tree Biomechanics Series: Crown Shape Factors and Volumes*. Available online at: <https://urbanforestrysouth.org/resources/library/citations/Citation.2005-12-07.0632>
- Dong, J., Burnham, J. G., Boots, B., Rains, G., and Dellaert, F. (2017). “4d crop monitoring: spatio-temporal reconstruction for agriculture,” in *Robotics and Automation (ICRA), 2017 IEEE International Conference on (IEEE)* (Singapore), 3878–3885.
- Feng, L., Dai, J., Tian, L., Zhang, H., Li, W., and Dong, H. (2017). Review of the technology for high-yielding and efficient cotton cultivation in the northwest inland cotton-growing region of china. *Field Crops Res.* 208, 18–26. doi: 10.1016/j.fcr.2017.03.008
- Friedli, M., Kirchgessner, N., Grieder, C., Liebisch, F., Mannale, M., and Walter, A. (2016). Terrestrial 3d laser scanning to track the increase in canopy height of both monocot and dicot crop species under field conditions. *Plant Methods* 12:9. doi: 10.1186/s13007-016-0109-7
- Giunta, F., Motzo, R., and Pruneddu, G. (2008). Has long-term selection for yield in durum wheat also induced changes in leaf and canopy traits? *Field Crops Res.* 106, 68–76. doi: 10.1016/j.fcr.2007.10.018
- Hamuda, E., Glavin, M., and Jones, E. (2016). A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electr. Agricult.* 125, 184–199. doi: 10.1016/j.compag.2016.04.024
- Jay, S., Rabatel, G., Hadoux, X., Moura, D., and Gorretta, N. (2015). In-field crop row phenotyping from 3d modeling performed using structure from motion. *Comput. Electr. Agricult.* 110, 70–77. doi: 10.1016/j.compag.2014.09.021
- Jiang, Y., Li, C., Sun, S., Paterson, A. H., and Xu, R. (2017). “Gphenovision: a ground mobile system with multi-modal imaging for field-based high throughput phenotyping of cotton,” in *2017 ASABE Annual International Meeting (American Society of Agricultural and Biological Engineers)*, 1 (Spokane, WA).
- Jiang, Y., Li, C. Y., and Paterson, A. H. (2016). High throughput phenotyping of cotton plant height using depth images under field conditions. *Comput. Electr. Agricult.* 130, 57–68. doi: 10.1016/j.compag.2016.09.017
- Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., et al. (2004). Review of methods for *in situ* leaf area index determination: part I. Theories, sensors and hemispherical photography. *Agricult. For. Meteorol.* 121, 19–35. doi: 10.1016/j.agrformet.2003.08.027
- Kim, C., Yun, S., Jung, S.-W., and Won, C. S. (2015). “Color and depth image correspondence for Kinect v2,” in *Advanced Multimedia and Ubiquitous Engineering*, eds J. Park, H. C. Chao, H. Arabnia, and N. Yen (Hanoi: Springer), 111–116.
- Kross, A., McNairn, H., Lapen, D., Sunohara, M., and Champagne, C. (2015). Assessment of rapideye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Observat. Geoinform.* 34, 235–248. doi: 10.1016/j.jag.2014.08.002
- Li, L., Zhang, Q., and Huang, D. F. (2014). A review of imaging techniques for plant phenotyping. *Sensors* 14, 20078–20111. doi: 10.3390/s141120078
- Müller-Linow, M., Pinto-Espinosa, F., Scharr, H., and Rascher, U. (2015). The leaf angle distribution of natural plant populations: assessing the canopy with a novel software tool. *Plant Methods* 11:11. doi: 10.1186/s13007-015-0052-z
- Murchie, E. H., Pinto, M., and Horton, P. (2009). Agriculture and the new challenges for photosynthesis research. *New Phytol.* 181, 532–552. doi: 10.1111/j.1469-8137.2008.02705.x
- Nguyen, T. T., Slaughter, D. C., Townsley, B. T., Carriedo, L., Maloof, J. N., and Sinha, N. (2016). “In-field plant phenotyping using multi-view reconstruction: an investigation in eggplant” in *International Conference on Precision Agriculture* (St. Louis, MO).
- Nock, C. A., Taugourdeau, O., Delagrangue, S., and Messier, C. (2013). Assessing the potential of low-cost 3d cameras for the rapid measurement of plant woody structure. *Sensors* 13, 16216–16233. doi: 10.3390/s131216216
- Norman, J. M., and Campbell, G. S. (1989). “Canopy structure,” in *Plant Physiological Ecology: Field Methods and Instrumentation*, eds R. Pearcy, H. A. Mooney, and P. Rundel (Dordrecht: Springer), 301–325.
- Pauli, D., Chapman, S., Bart, R., Topp, C. N., Lawrence-Dill, C., Poland, J., et al. (2016). The quest for understanding phenotypic variation via integrated approaches in the field environment. *Plant Physiol.* 172, 622–634. doi: 10.1104/pp.16.00592
- Paulus, S., Behmann, J., Mahlein, A. K., Plümer, L., and Kuhlmann, H. (2014). Low-cost 3d systems: suitable tools for plant phenotyping. *Sensors* 14, 3001–3018. doi: 10.3390/s140203001
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reta-Sánchez, D. G., and Fowler, J. L. (2002). Canopy light environment and yield of narrow-row cotton as affected by canopy architecture. *Agron. J.* 94, 1317–1323. doi: 10.2134/agronj2002.1317
- Reynolds, M., and Langridge, P. (2016). Physiological breeding. *Curr. Opin. Plant Biol.* 31, 162–171. doi: 10.1016/j.pbi.2016.04.005
- Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., et al. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE* 11:e0159781. doi: 10.1371/journal.pone.0159781
- Sinclair, T., and Horie, T. (1989). Leaf nitrogen, photosynthesis, and crop radiation use efficiency: a review. *Crop Sci.* 29, 90–98. doi: 10.2135/cropsci1989.0011183X002900010023x

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.02233/full#supplementary-material>

- Stewart, D., Costa, C., Dwyer, L., Smith, D., Hamilton, R., and Ma, B. (2003). Canopy structure, light interception, and photosynthesis in maize. *Agron. J.* 95, 1465–1474. doi: 10.2134/agronj2003.1465
- Torr, P. H., and Zisserman, A. (2000). Mlesac: a new robust estimator with application to estimating image geometry. *Comput. Vis. Image Understand.* 78, 138–156. doi: 10.1006/cviu.1999.0832
- USDA-ERS (2017). *Cotton and Wool: Overview*. Available online at: <https://www.ers.usda.gov/topics/crops/cotton-wool/> (Accessed November 11, 2017).
- Weiss, M., and Baret, F. (2017). Using 3D point clouds derived from UAV RGB imagery to describe vineyard 3D macro-structure. *Remote Sens.* 9:111. doi: 10.3390/rs9020111
- Weiss, M., Baret, F., Smith, G., Jonckheere, I., and Coppin, P. (2004). Review of methods for *in situ* leaf area index (LAI) determination: part II. Estimation of LAI, errors and sampling. *Agricult. For. Meteorol.* 121, 37–53. doi: 10.1016/j.agrformet.2003.08.001
- White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003
- Zheng, G., and Moskal, L. M. (2009). Retrieving leaf area index (LAI) using remote sensing: theories, methods and sensors. *Sensors* 9, 2719–2745. doi: 10.3390/s90402719
- Zhu, X. G., Long, S. P., and Ort, D. R. (2010). Improving photosynthetic efficiency for greater yield. *Annu. Rev. Plant Biol.* 61, 235–261. doi: 10.1146/annurev-arplant-042809-112206

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Jiang, Li, Paterson, Sun, Xu and Robertson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Leaf Movements of Indoor Plants Monitored by Terrestrial LiDAR

Mónica Herrero-Huerta^{1,2,3*}, Roderik Lindenbergh¹ and Wolfgang Gard⁴

¹ Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, Netherlands, ² TIDOP Research Group, Higher Polytechnic School of Avila, University of Salamanca, Avila, Spain, ³ Agronomy Department, Purdue University, West Lafayette, IN, United States, ⁴ Department of Structural and Building Engineering, Delft University of Technology, Delft, Netherlands

OPEN ACCESS

Edited by:

Eetu Puttonen,
National Land Survey of Finland,
Finland

Reviewed by:

Jan Behmann,
University of Bonn, Germany
Gerben van Ooijen,
University of Edinburgh,
United Kingdom

*Correspondence:

Mónica Herrero-Huerta
monicaherrero@usal.es

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 11 October 2017

Accepted: 31 January 2018

Published: 16 February 2018

Citation:

Herrero-Huerta M, Lindenbergh R and
Gard W (2018) Leaf Movements of
Indoor Plants Monitored by Terrestrial
LiDAR. *Front. Plant Sci.* 9:189.
doi: 10.3389/fpls.2018.00189

Plant leaf movement is induced by some combination of different external and internal stimuli. Detailed geometric characterization of such movement is expected to improve understanding of these mechanisms. A metric high-quality, non-invasive and innovative sensor system to analyze plant movement is Terrestrial LiDAR (TLiDAR). This technique has an active sensor and is, therefore, independent of light conditions, able to obtain accurate high spatial and temporal resolution point clouds. In this study, a movement parameterization approach of leaf plants based on TLiDAR is introduced. For this purpose, two *Calathea roseopicta* plants were scanned in an indoor environment during 2 full-days, 1 day in natural light conditions and the other in darkness. The methodology to estimate leaf movement is based on segmenting individual leaves using an octree-based 3D-grid and monitoring the changes in their orientation by Principal Component Analysis. Additionally, canopy variations of the plant as a whole were characterized by a convex-hull approach. As a result, 9 leaves in plant 1 and 11 leaves in plant 2 were automatically detected with a global accuracy of 93.57 and 87.34%, respectively, compared to a manual detection. Regarding plant 1, in natural light conditions, the displacement average of the leaves between 7.00 a.m. and 12.30 p.m. was 3.67 cm as estimated using so-called deviation maps. The maximum displacement was 7.92 cm. In addition, the orientation changes of each leaf within a day were analyzed. The maximum variation in the vertical angle was 69.6° from 12.30 to 6.00 p.m. In darkness, the displacements were smaller and showed a different orientation pattern. The canopy volume of plant 1 changed more in the morning (4.42 dm³) than in the afternoon (2.57 dm³). The results of plant 2 largely confirmed the results of the first plant and were added to check the robustness of the methodology. The results show how to quantify leaf orientation variation and leaf movements along a day at mm accuracy in different light conditions. This confirms the feasibility of the proposed methodology to robustly analyse leaf movements.

Keywords: leaf movements, plants, terrestrial LiDAR, indoor, temporal series

INTRODUCTION

Plant canopy structure properties and their spatial changes, are linked to different vegetation processes, such as radiation absorption, plant water balance, precipitation interception and photosynthetic activity (Harley and Baldocchi, 1995). Canopy structural and biochemical variables are frequently used as constraints to model interactions between the land surface and the

atmosphere (Sellers et al., 1997). Moreover, many studies that investigated the implications of interception and light transmission for species competition, biodiversity, ecosystem and agro-ecosystem dynamics, as well as wood production, depend on the spatial distribution of leaves and branches (Pretzsch, 2009).

In ecology and plant physiology, circadian rhythms are activities that occur on a near-24-h cycle due to ecologically useful adaptations, regarding plant's physiology and its environment (Sadava et al., 2009). In this context, the growth patterns of roots and leaves are determined by the circadian clock and leaf starch metabolism (Ruts et al., 2012). Therefore, diel leaf movements are a well-typified symbol of the circadian clock (Farré, 2012). These movements have been studied intensively for a long time (Barak et al., 2000). Two reasons are notably identified by scientific researches as operators of these fluctuations, plant water balance (Chapin et al., 2011), and photoperiodism (Sysoeva et al., 2010), related to the fluency rate of photosynthetically active radiation in the plant. To optimize the interception of incoming light and to avoid temperature related stress, the leaf angle is adjusting (Medina et al., 1978; Ehleringer, 1998). Plant imaging aims to analyse and quantify the development, growth, physiological and other phenotypic plant properties by different accurate and automated processes. A complete review focuses on the latest advances in high-throughput image-based plant phenotyping (Fahlgren et al., 2015). Although many successes are reported, image based techniques may require the use of a flash during low-light conditions, while some effort is needed in setup and/or processing to manually or automatically obtain 3D results.

TLiDAR has newly arisen as a promising tool to fast measure 3D vegetation structure at plot level with high spatial resolution and millimeter accuracy (Dassot et al., 2010). Light Detection And Ranging (LiDAR) is an active remote sensing technique that accurately measures distances by transmitting laser energy and detecting the time of arrival of the return energy. As an active technique, it is almost insensitive to varying external lighting conditions and is able to capture data even in absence of light. TLiDAR data are extensively used in engineering applications to monitor morphological terrain changes (Herrero-Huerta et al., 2016). The interest on TLiDAR for vegetal plot measurements started in the past decade in forestry (van Leeuwen and Disney, 2018) and (Thies and Spiecker, 2004). Recently, a fully-automatic approach for tree structural modeling at plot level has been proposed in Raunonen et al. (2013). Circadian rhythms in tree geometry can be accurately monitored at millimeter scale and sub-hour temporal resolution by TLiDAR in outdoor conditions (Puttonen et al., 2016). Dornbusch et al. analyse circadian movements and rhythmic leaf growth of *Arabidopsis* rosettes using a near-infrared laser scanner (Dornbusch et al., 2014). Low cost 3D imaging tools were used for this purpose in Paulus et al. (2014) and evaluated with respect to the metric quality of estimated geometric plant features. To summarize, TLiDAR has potential to monitor vegetation geometry at different scale levels. In addition, scanning can be done outdoors, is insensitive to vary lighting conditions and is able to operate at up to temporal resolutions of minutes. The

scanning mechanism is based upon a fast rotating multi-facet polygonal mirror, so little effect is transmitted to plants. As a disadvantage, point clouds cannot provide direct biochemical parameters from plants, as hyperspectral and thermal chlorophyll fluorescence imaging do (Fiorani and Schurr, 2013). On the other hand, image based-3D modeling needs a significant amount of post-processing, while TLiDAR requires a higher cost.

Goal of this research is to show how plant moves can efficiently be captured and parameterized using TLiDAR. In general, data capturing using TLiDAR is considered relatively easy. However, there is still a lack of methodology and corresponding software tool to automatically extract geometric parameters from point cloud data sampling moving plants. Thus, the goal of this study is to propose a phenotyping pipeline to automatically determine movement from leaves in indoor plants through TLiDAR.

The paper is organized as follows: after this brief introduction, the proposed methodology is described in section Methodology. Next, the results and the discussion are presented in section Experimental Results and Discussion, respectively. Finally, conclusions are drawn.

METHODOLOGY

Experimental Setup

An experiment was carried out with the understory species *Calathea roseopicta* (Linden) Regel, a native species from Tropical America, with basal leaves growing in a rosette shape. Because of the photo- and thermotropism of most tropical plants, *Calathea* unfurl its leaves and moves the whole system including the stem toward the light to optimize the surface for sunlight absorption. During the dark period (night) the leaves coil up and minimize the surface to prevent the plant from cooling-off. Therefore, this species is suitable to monitor the complex, spatial movement of the leaf-stem system. It needs a minimum temperature of 16°C so in temperate systems, it is commonly kept in indoor conditions. Its non-lignified tissues allow the presence of soft tissues susceptible to soil water and light conditions (Chao-ying, 2009). Two plants were tested, planted in an 850 cm³ PVC container filled with a substrate composed of sand and a high content of organic material. Studied plant #1 had a mean height of ~45 cm and a canopy diameter of ~31 cm during the measurements, while plant #2 has ~44 cm and ~29 cm, respectively, as **Figure 1** illustrates.

The experiment was performed in December 2016. Air temperature (°C) and relative humidity (%) were continuously monitored during the sampling period with a HOBO U30-NRC weather station, with sampling intervals of 5 min. Light conditions were measured during the experiment by the "LuxMeterPro Advanced App" from AM PowerSoftware, with 12 lux-accuracy for a distance <5 m. A Leica Scan-Station C10 scanner, a high speed TLiDAR implementing the so-called time-of-flight principle, is used to collect data (Vosselman and Maas, 2010)¹. Obstructions in the visibility w.r.t. the position of the

¹Available online at: <https://hds.leica-geosystems.com> (Accessed January 11, 2018).

laser scanner is hindering complete sampling of the plant surface. Indeed, leaves in the front, as seen from the scanner position, may occlude leaves in the back. Thereby, the relative horizontal position of the TLiDAR w.r.t. the scanned plant should be carefully chosen to have best coverage of leaves within a single scan. Using multiple scans could mitigate this occlusion effect but would (i) include larger time differences within one acquisition and (ii), would require an additional matching or registration step to align the scans as obtained from different positions into a common coordinate system.

Both plants were located in a room, 1.6–1.8 m from the single window oriented in SW direction. Scans were performed at 2 different days. The first sampling day (test A) was done under natural light conditions. Previously to these TLiDAR measurements, the plants were kept in full light conditions. After a 2 day-period in which the plants were kept in complete darkness, the second sampling (test B) was carried out in the dark. An overview of the acquired data is shown in **Table 1**.

Method of Leaf Movement Estimation

Outlier Removal

The point cloud sample of the plants may contain outliers and noisy points caused by various reasons such as backscatter from interference effects (Thies and Spiecker, 2004). Such points are not regarded as samples of the actual plant and the first step is to filter them from the point cloud. To remove separate isolated points or few-point clusters, a statistical analysis of the distances is performed as follow. First, all points are visited. For each query point, all neighboring points within a pre-set distance are determined and all distances from the query point to these points are determined. After visiting all points, the mean and standard deviation of these distances (\bar{d} and σ_d , respectively) is stored.

Next, a threshold (L_d) is defined as:

$$L_d = \bar{d} + p \cdot \sigma_d \quad (1)$$

where p is the confidence level, expressed as the critical value associated in the standard normal density curve (Palnick, 2014).

Again, all points are visited. For a given point, when its average distance to its k neighbors is above L_d , it is marked as an outlier and removed (Rusu et al., 2008). This method is reasonably fast once the points are organized in a data structure like an octree.

Point Cloud Registration and Deviation Map Calculation

Scans are taken without moving the TLiDAR during the sampling periods. Still, it is not guaranteed that the scans obtained at different moments are fully aligned. Therefore, the registration or alignment of the 3D point clouds is checked using 7 external spherical targets of 10 cm-diameter from Leica®. This registration step estimates a rigid body transformation (only translations and rotations can be applied to the point clouds). The Iterative Closest Point algorithm by Besl and McKay (1992) is used, reaching a negligible mean absolute error among targets from multi-temporal datasets.

After that, all the point clouds are cropped using a common bounding box. At this point, the recognition of correspondences among point clouds is essential to compute plant movements of the plant as displacements between multi-temporal datasets. Therefore, first cloud-to-cloud differences between sequential TLiDAR datasets are computed using a chamfer distance approach that exploits an octree organization (Akmal Butt and Maragos, 1998). Second, the distance statistics resulting from this first step are used to refine the displacements. A more precise cloud-to-cloud distance between two multi-temporal point clouds is extracted by applying a local approximation model to the reference cloud by a quadric surface. As a result, a deviation map of displacements between point clouds is obtained which precisely outlines the coordinates and enables the quantification of plant movements.

Segmentation in Individual Leaves

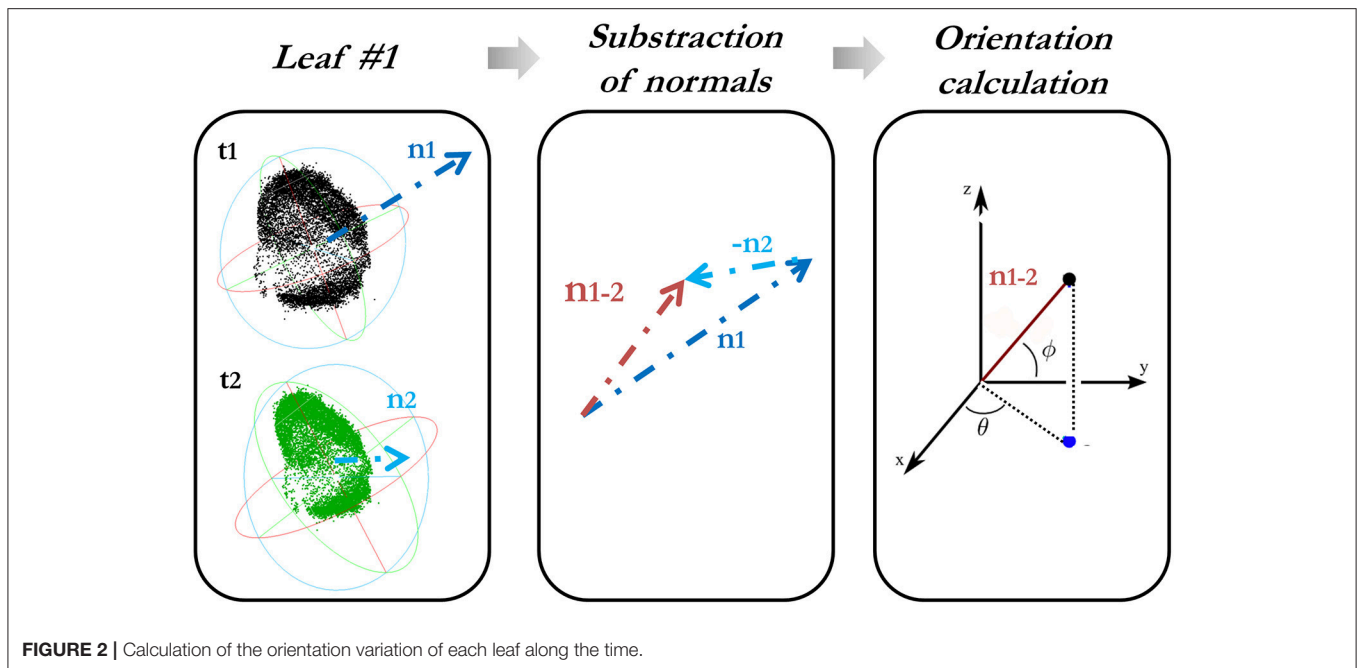
The changes in the orientation of the leaves along time are estimated from the point cloud data. For that purpose, individual plant leaves are extracted from the point clouds using a segmentation method based on an organization of the point cloud in an octree structure, as proposed in Woo et al. (2002).



FIGURE 1 | Picture of *Calathea roseopicta* plant 2.

TABLE 1 | Overview of the acquired data.

Test	Condition	Plant	#points in Point Cloud		
			7:00 a.m	12:30 a.m	6:00 p.m
A	Natural light	1	98,868	94,043	93,235
		2	84,966	82,727	88,597
B	Darkness	1	94,355	96,456	95,765
		2	81,916	82,254	77,008



Initially, the point normals are estimated by Delaunay triangulation (Golias and Dutton, 1997). Next, the 3D points are organized in an octree structure, where the local point normals are used as a criterion for splitting an octree cell: if points at a certain octree depth all have similar normals, subdivision terminates; otherwise, the octree cell is divided. This process is repeated until normals stabilize at a certain depth or the maximal octree depth is reached. After splitting is finished, adjacent cells having similar orientation are merged in a bottom-up procedure. First, cells with a large variation in normals are simply removed. Such behavior might, for example, occur at the edges of leaves. Next, an optimally flat cell is selected as a seed cell and adjacent cell are added as long as a homogeneity criterion is met. When such segment is finished, a new seed is selected and the procedure is repeated using the remaining cells. The standard deviation of the normal vectors in a cell is taken as homogeneity criterion. After termination of this procedure, the point cloud is automatically separated into certain regions by considering the obtained segments. Using a suitable criterion value, it is possible to tune the method such that the resulting segments correspond to individual leaves.

The last step is to match individual leaves, which is difficult as the individual leaves are moving and changing the shape between the different data acquisitions. Automatic co-registration of each leaf point cloud segment is done by least squares matching of overlapping surfaces (Least Squares 3D Surface Matching). The transformation parameters of the 3D compared surface (leaf segment after movement) to a 3D reference surface (leaf segment before movement) is approximated by a so-called Generalized Gauss–Markov model, based on minimizing the squared sum of Euclidean distances between surfaces (Gruen and Akca, 2008).

Leaf Normal Calculation

Once the leaves are segmented, leaf-wise orientation is approximated by the leaf normal. For this purpose, Principal Component Analysis (PCA) is used (Weinmann et al., 2014). This statistical analysis uses the first and second moments of the points and results in three orthogonal vectors centered on the center of gravity of the point cloud. The PCA synthesizes the distribution of points along the three dimensions and therefore models the principal directions and magnitudes of variation of the point distribution around the center of gravity.

The coordinates x_i , y_i , and z_i for each point $i=1, \dots, k$ from the point cloud of each leaf, is considered. The covariance matrix (Σ) (2) of each leaf point cloud (X) is defined by:

$$\Sigma = \frac{(X - \bar{X})^T (X - \bar{X})}{k} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix} \quad (2)$$

where $(\sigma_x^2, \sigma_y^2, \sigma_z^2)$ are the variances of the coordinate directions and the elements outside the main diagonal of Σ are the covariances. \bar{X} contains k copies of the mean of the three coordinates of the leaf point cloud (X). The principal leaf normal is the third eigenvector of Σ (Jolliffe, 2002).

Computation of Orientation Variation

The orientation variation of each leaf is characterized in a spherical coordinate system by determining the change in both spherical azimuth (θ) (aspect) and elevation angle (ϕ) (hyponasty). For this reason, firstly, a subtraction of the normals from each leaf at different times is done, always in temporal order with respect to the first epoch. These angles obtained from the

difference-in-normal vector (x, y, z) are calculated as follows:

$$\theta = \arctan \frac{y}{x} \quad (3)$$

$$\varnothing = \arctan \frac{z}{\sqrt{x^2 + y^2}} \quad (4)$$

Figure 2 explains how the computation of the orientation variation is carried out.

Volume Variations by Canopy Mesh

The foliage shape of the full plant may help to understand the radiation regime and canopy structure. However, the single scan acquisitions only give a partial point cloud due to the one-sided field of view of the TLiDAR. This paragraph explains how to quantify the canopy volume variation from the point cloud of the plant.

The approach is supported by a 3D modeling of the plant using a 3D Delaunay triangulation (Golias and Dutton, 1997), as used before for the leaf segmentation (section Segmentation in Individual Leaves). The result is a 3D mesh. An additional filter phase is required to obtain a plant canopy model close to reality. The approximation shown in Attene (2008) is used in this step which incorporates several smoothing operators to repair issues in the mesh like holes and different types of noise (Fan et al., 2008).

Once the canopy mesh is fixed, its shape is approximated by a 3D convex hull (Barber et al., 1996). The convex-hull is chosen due to its capacity of enveloping the effective area which potentially can be related to the radiance absorption. As an alternative in future, an alpha-shape approximation of the hull could be used, which is able to more flexible estimate the shape of the plant canopy, by varying the alpha coefficient.

As single scans only provide partial point clouds, the recovery of the less well-visible backside is achieved through symmetry. Therefore, the symmetry axis is estimated by a 180° field of view from the dataset. Finally, the canopy shape is obtained which allows the direct volume estimation and changes along time.

EXPERIMENTAL RESULTS AND DISCUSSION

Sunrise and sunset times were similar during the experiment. Therefore, the sampling periods are comparable. The plant substrate was maintained well-watered to reduce the effect of root water shortage. The room temperature remained constant during the sampling period; however the relative humidity changed between test A and B (**Table 2**). This variation in humidity could have led to stress or more water movement within the plants in test B.

The scan pre-processing settings to remove noise (section Outlier Removal) from the data were fixed to a neighborhood size of a 1.6 cm sphere diameter. The leaf movements of *C. roseopicta* plant 1 are illustrated in **Figure 3**, by superimposing the scan obtained at 7.00 a.m. to the one obtained at 12.30 p.m. during test A (**Figure 3A**) and by calculating the deviation map from the compared point cloud (at 12.30 p.m.) to the reference point cloud

TABLE 2 | Room conditions during test A and B [sunrise and sunset times provided in GMT(+1)].

Test	Date	Timing		Variables	
		Sunrise	Sunset	Temp.	Rel. humidity
A	02/12/2016	8:30 a.m	16:34 p.m	21.2°	47.4 %
B	05/12/2016	8:34 a.m	16:32 p.m	21.2°	32.2 %

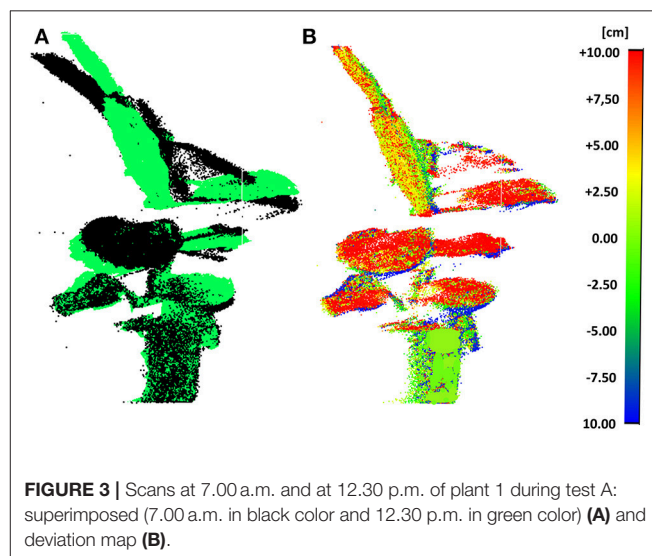


FIGURE 3 | Scans at 7.00 a.m. and at 12.30 p.m. of plant 1 during test A: superimposed (7.00 a.m. in black color and 12.30 p.m. in green color) (**A**) and deviation map (**B**).

(at 7.00 a.m.) (**Figure 3B**), as section Point Cloud Registration and Deviation Map Calculation explains. This time interval was chosen as during this time the biggest movement in terms of leaf displacement occurred. Further estimations can be obtained from this product. For instance, the maximum displacement of 7.92 cm between the two scans occurs at leaf #7. The average distance movement of this leaf is 3.67 cm.

In order to apply the segmentation method to extract individual leaves, some parameters have to be fixed; notably, the minimum voxel size and the number of iterations. These parameters were set as 1.8 cm and 5 iterations, respectively. Furthermore, the standard deviation of point normals within a cell as a tolerance for subsequent subdivisions, was set at 6 cm. To perform the Least Square Matching for individually matching the leaves among multi-temporal datasets, the iteration criteria threshold was set as 3.8 cm. As a result, 9 leaves of plant 1 and 11 leaves of plant 2 were recognized from all the scans. **Figure 4** shows the individual leaves extracted from both plants at 7.00 a.m. (test A) in different colors, where normals in red are obtained by PCA. Outliers are colored in dark blue.

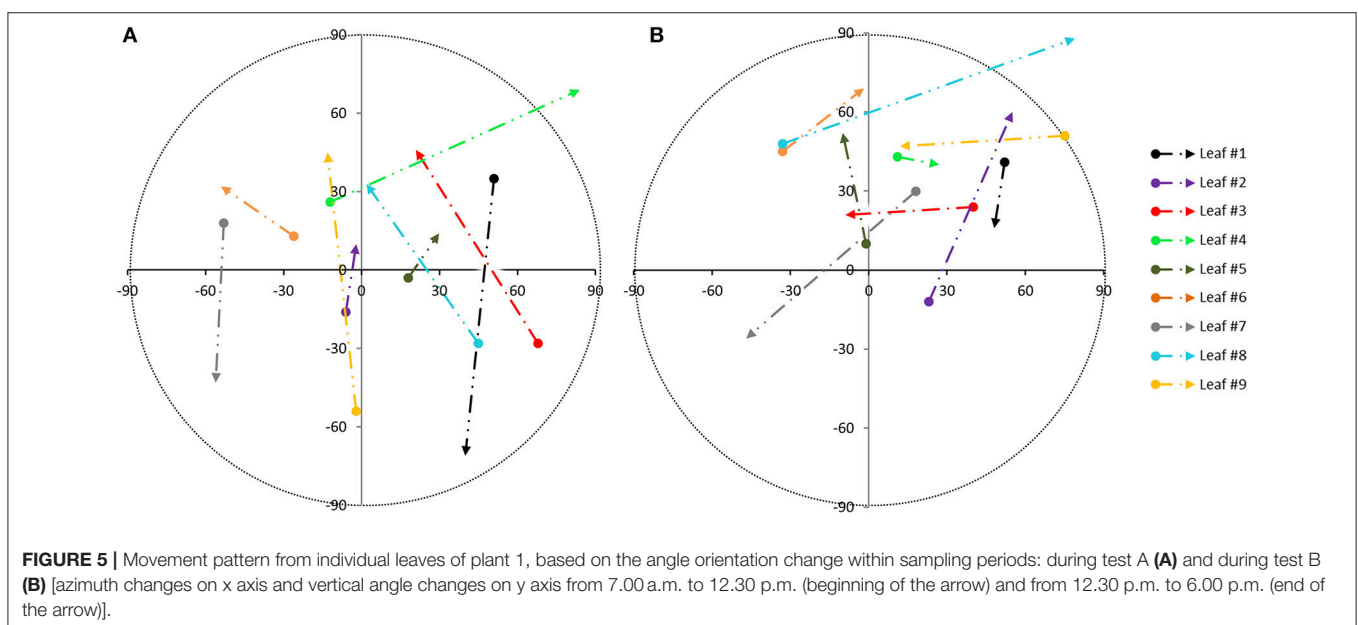
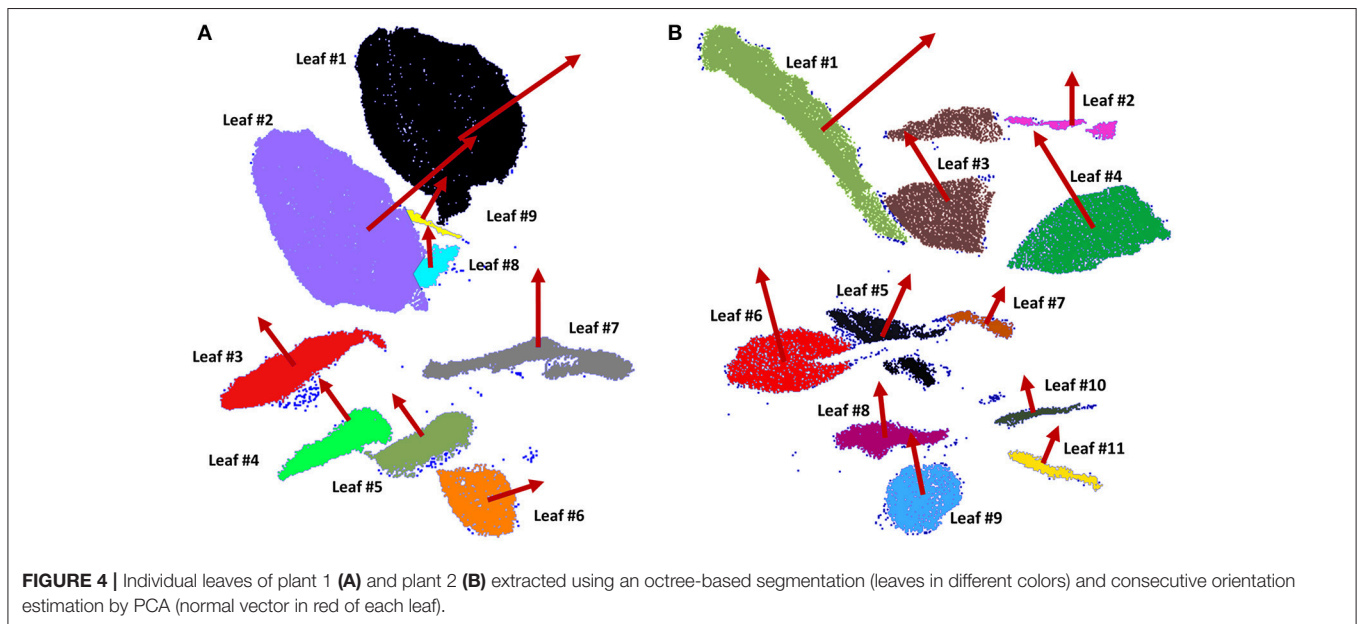
The global accuracy reached was analyzed by comparison to a manual segmentation of the leaves. To do so, each leaf is considered a target class and for each point. It can be determined if it is classified in the correct class. Results are collected in a confusion matrix (not shown), (Sasidharan et al., 2010). The resulting global accuracy was found to be 93.57% for plant 1

and 87.34% for plant 2. The plant species morphology gives the opportunity to use the proposed automatic segmentation technique due to the fact that the leaves and stems have dissimilar orientations. Once the individual leaves are obtained, the variability of each leaf orientation along the sampling tests is estimated. **Figure 5** shows the variation in orientation from the leaves of plant 1 during test A (**Figure 5A**) and during test B (**Figure 5B**). These variations were first derived between 7.00 a.m. and 12.30 p.m. (beginning of the arrow) and next between 12.30 and 6.00 p.m. (end of the arrow). These intervals were chosen because during these intervals, movement was bigger. The azimuth changes are represented on the x -axis while the vertical angle change is on the y -axis (both in sexagesimal

degrees), taking into account the normal direction of each leaf as obtained by PCA. The color of each trajectory corresponds to the leaf color in **Figure 1**. Different light conditions clearly correspond to different movement directions. The orientation change in darkness (**Figure 5B**) is smaller than during natural light conditions (**Figure 5A**).

The results of the orientation variations from plant 2 are displayed in **Table 3**. A similar magnitude order in the orientation variation is reached regarding plant 1 during both tests.

From the scan analysis results, we conclude that *C. roseopicta* displayed significant movement in the leaves during light and dark periods, with different patterns for both tests and plants.



This species shows conspicuous movement between the tested light conditions, however this leaf movement can be influenced by growing patterns of young leaves (Puttonen et al., 2016).

Leaf #1 from plant 1, for example, moves in the same direction during both light conditions, but its movement is larger in natural light conditions. In both situations, the starting point is similar: it means that the angle variation between 7.00 a.m. and 12.30 p.m. is comparable. Conversely, the end point is really different (from 12.30 to 6.00 p.m.). A possible explanation is that the movement of this leaf during the afternoon is related to the photosynthesis process, so the leaf movement has less influence for the sun-light in the morning. It can probably be related

to the different transpiration rates produced by leaf #1 within a day.

Canopy variations along time can actually be quantified. As section Volume Variations by Canopy Mesh describes, the convex-hull of the entire plant is generated from the point cloud. This process is illustrated in **Figure 6**, where the shape variation of the plant 2 during test A at 12.30 p.m. (**Figure 6A**) and 6.00 p.m. (**Figure 6B**) can be clearly defined as 'closing the leaves' (**Figure 6C**). The convex-hull volume was analyzed to extract the variations at different times. **Table 4** characterizes these values during the morning and afternoon in test A and B from the two studied plants, proving that in natural light conditions the variation is always bigger.

Thus, our data shows that light conditions influence diel leaf movement, characterized by movement patterns and leaf movement rate. As a general idea, we saw that leaves were opening during the day and closing at night (changing the elevation angle along sampling periods), generating an up and down movement as previous studies already indicated (Manel et al., 2001).

Notice that leaves can generate shadows within plants, occluding certain parts, therefore the obtained 3D point cloud of the plant surface, usually is incomplete. To overcome this challenge, the individual leaf architecture could be reconstructed via a model, already proposed for cereal leaves (Dornbusch et al., 2012).

The biological significance of diel leaf movements is related to favorable conditions of energetic requirements, water availability and auxin responsiveness (Stitt and Zeeman, 2012). This proposition agrees with the maximal peak of movement and

TABLE 3 | Results of orientation variation from plant 2 during test A and B.

Test	#Leaf	Azimuth changes		Vertical angle changes	
		7.00–12.30	12.30–18.00	7.00–12.30	12.30–18.00
A	#1	–58°	–54°	–56°	–64°
	#2	47°	–22°	51°	27°
	#3	20°	69°	–3°	–53°
	#4	–7°	15°	–20°	–24°
	#5	–42°	18°	45°	–14°
	#6	–83°	24°	–88°	–21°
	#7	–11°	–72°	–32°	–10°
	#8	–55°	81°	–70°	–64°
	#9	31°	31°	34°	–55°
	#10	13°	–13°	–45°	–39°
	#11	–7°	–2°	60°	–47°
B	#1	–63°	75°	75°	–71°
	#2	19°	–18°	35°	49°
	#3	–70°	7°	85°	–45°
	#4	52°	37°	34°	–52°
	#5	–5°	–18°	5°	–28°
	#6	9°	47°	9°	58°
	#7	–70°	–79°	–50°	83°
	#8	6°	81°	46°	61°
	#9	18°	–23°	–39°	12°
	#10	79°	28°	–88°	1°
	#11	0°	18°	63°	36°

TABLE 4 | Canopy volume analysis by convex-hull.

Test	#Plant	Canopy volume (dm ³)			Volume variation (dm ³)	
		7:00	12:30	18:00	7:00–12:30	12:30–18:00
A	1	33.12	28.70	31.27	–4.42	2.57
B	1	30.6	30.49	28.96	–0.16	–1.53
A	2	37.24	41.98	25.00	4.74	–16.98
B	2	45.52	43.36	44.80	–2.16	1.44

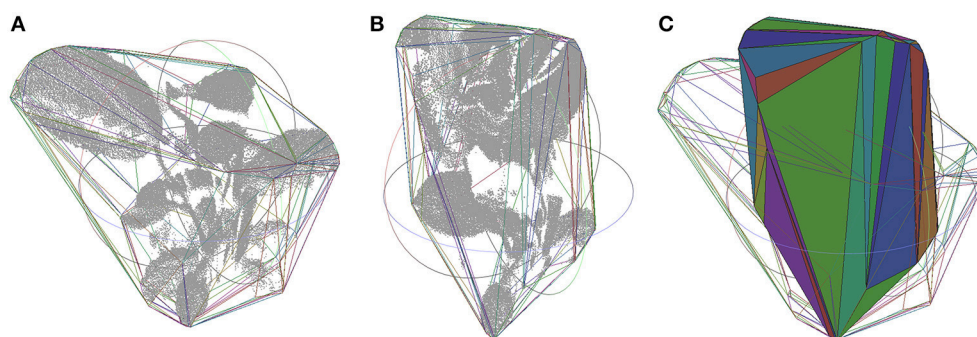


FIGURE 6 | Canopy point cloud at 12.30 p.m. (A) and at 6.00 p.m. (B) during test A from plant 1, together with its convex hull plotted by lines in random colors depending on the face they belong to. Both convex hulls superimposed, one in lines and the other in solid (C).

orientation variation during the early morning. Moreover, the leaf vertical angle tracks the usual daily temperature oscillations with a top in the late-afternoon. Bridge et al. (Bridge et al., 2013) determined that moving up the leaves is favorable to cool them during the warm daily hours and reduces the radiation contents when it exceeds the photosynthetic capacity.

CONCLUSIONS

This paper shows how TLiDAR can be used to efficiently and non-invasively parameterize plant movement. It presents a phenotyping method to monitor leaf movements from plants by TLiDAR, even in absence of light. We show that this approach is robust and computationally easy to determine, providing a tool to accurately measure complex leaf-stem system movements from plants with high temporal and spatial resolution.

Calathea roseopicta species was studied, showing different leaf movements in presence of natural daylight conditions and in darkness. Moreover, canopy variations in volume were analyzed by a convex hull approach. In darkness, the displacements are smaller and with a dissimilar orientation pattern than in natural light conditions. The natural movements are expressed by opening of the leaves during the day and closing them at night, pointing to the source of natural light (the window, in our study case). In absence of light, the plant still showed displacements, but with different pattern and magnitude.

The low relative humidity conditions during both sampling periods may have resulted in an increment in transpiration rates, leading to possible leaf movement during the photosynthesis. As

further consideration, water availability should be controlled and sunlight received by the plant should be equal in all directions to rigorously compare multi-temporal datasets.

Using TLiDAR technology to derive vegetation movements open up fresh challenges which merit more methodological advance. The proposed methodology could help to support biological investigations of exchange of water, sunlight absorption and other bio-geochemical between plants and the atmosphere. Furthermore, these achievements would break through the barrier of knowledge in objectively determining the behavior of plants and even its structure and function in ecosystem, nowadays remaining poorly understood (Whippo and Hangarter, 2009). Furthermore, the behavior of interior plants is linked to indoor air quality and this connection will be explored as further studies.

AUTHOR CONTRIBUTIONS

MH-H designed the experiment, created the software, ran the data analysis and analyzed the data; MH-H and RL wrote the paper; WG edited the paper. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

Authors would like to thank Cesar Jimenez-Rodriguez from the Department of Water Management and Jacintha Tjia, bachelor' student, both from Delft University of Technology, for their collaboration during the experimental phase of this research.

REFERENCES

- Akmal Butt, M. and Maragos, P. (1998). Optimum design of chamfer distance transforms. *Trans. Image Process.* 7, 1477–1484. doi: 10.1109/83.718487
- Attene, M. (2008). A lightweight approach to repairing digitized polygon meshes. *Vis. Comput.* 26, 1393–1406. doi: 10.1007/s00371-010-0416-3
- Barak, S., Tobin, E. M., Andronis, C., Andronis, C., Sugano, S. and Green, R. M. (2000). All in good time: the Arabidopsis circadian clock. *Trends Plant Sci.* 5, 517–522. doi: 10.1016/S1360-1385(00)01785-4
- Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 22, 469–483. doi: 10.1145/235815.235821
- Besl, P. J., and McKay, N. D. (1992). A method for registration of 3-D shapes. *Trans. Pattern Anal. Mach. Intell.* 14, 239–256. doi: 10.1109/34.121791
- Bridge, L. J., Franklin, K. A., and Homer, M. E. (2013). Impact of plant shoot architecture on leaf cooling: a coupled heat and mass transfer model. *J. R. Soc. Interface* 10:0326. doi: 10.1098/rsif.2013.0326
- Chapin, S., Matson, P. A., and Vitousek, P. (2011). *Principles of Terrestrial Ecosystem Ecology*. Springer Science & Business Media.
- Chao-ying, M. A. (2009). Dormancy movements of *Calathea mauiqueen*. *J. Anhui Agric. Sci.* 5:59. Available online at: http://en.cnki.com.cn/Article_en/CJFDTOTAL-AHNY200905059.htm
- Dassot, M., Constant, T., and Fournier, M. (2010). The use of terrestrial LiDAR technology in forest science application fields benefits and challenges. *Ann. For. Sci.* 68, 959–974. doi: 10.1007/s13595-011-0102-2
- Dornbusch, T., Lorrain, S., Kuznetsov, D., Fortier, A., Liechti, R., Xenarios, I., et al. (2012). Measuring the diurnal pattern of leaf hyponasty and growth in Arabidopsis—a novel phenotyping approach using laser scanning. *Funct. Plant Biol.* 39, 860–869. doi: 10.1071/FP12018
- Dornbusch, T., Michaud, O., Xenarios, I., and Fankhauser, C. (2014). Differentially phased leaf growth and movements in arabidopsis depend on coordinated circadian and light regulation. *Plant Cell* 26, 3911–3921. doi: 10.1105/tpc.114.129031
- Ehleringer, J. R. (1998). Changes in leaf characteristics of species along elevational gradients in the Wasatch Front, Utah. *Am. J. Bot.* 75, 680–689. doi: 10.1002/j.1537-2197.1988.tb13490.x
- Fahlgren, N., Gehan, M. A., and Baxter, I. (2015). Lights camera action high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* 24, 93–99. doi: 10.1016/j.pbi.2015.02.006
- Fan, Y. Z., Tam, B. S., and Zhou, J. (2008). Maximizing spectral radius of unoriented laplacian matrix over bicyclic graphs of a given order. *Lin. Multilinear Algebra* 56, 381–397. doi: 10.1080/03081080701306589
- Farré, E. M. (2012). The regulation of plant growth by the circadian clock. *Plant Biol.* 14, 401–410. doi: 10.1111/j.1438-8677.2011.00548.x
- Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 267–291. doi: 10.1146/annurev-arplant-050312-120137
- Golias, N. A., and Dutton, R. W. (1997). Adaptive meshing, part 2 delaunay triangulation and 3D adaptive mesh generation. *Finite Elem. Anal. Des.* 25, 331–41. doi: 10.1016/S0168-874X(96)00054-6
- Gruen, A., and Akca, D. (2008). Least squares 3D surface and curve matching. *J. Photogramm. Remote Sens.* 59, 151–174. doi: 10.1016/j.isprsjprs.2005.02.006
- Harley, P. C., and Baldocchi, D. D. (1995). Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest. I. Leaf model parametrization. *Plant Cell Environ.* 18, 1146–1156. doi: 10.1111/j.1365-3040.1995.tb00625.x
- Herrero-Huerta, M., Lindenbergh, R., Ponsioen, L., and van Damme, M. (2016). Morphological changes along a dike landside slope sampled by 4d high resolution terrestrial laser scanning. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 41, 227–232. doi: 10.5194/isprsarchives-XLI-B3-227-2016

- Jolliffe, I. T. (ed.). (2002). "Graphical representation of data using principal components," in *Principal Component Analysis* (New York, NY: Springer), 78–110.
- Manel, S., Williams, H. C., and Ormerod, S. J. (2001). Evaluating presence-absence models in ecology: the need to account for prevalence. *J. Appl. Ecol.* 38, 921–931. doi: 10.1046/j.1365-2664.2001.00647.x
- Medina, E., Sobrado, M., and Herrera, R. (1978). Significance of leaf orientation for leaf temperature in an amazonian sclerophyll vegetation. *Radiat. Environ. Biophys.* 15, 131–140. doi: 10.1007/BF01323262
- Palnick, J. (2014). *Plane Detection and Segmentation for DARPA Robotics Challenge*. Doctoral dissertation, Worcester polytechnic Institute.
- Paulus, S., Behmann, J., Mahlein, A. K., Plümer, L., and Kuhlmann, H. (2014). Low-cost 3D systems: suitable tools for plant phenotyping. *Sensors* 14, 3001–3018. doi: 10.3390/s140203001
- Pretzsch, H. (ed.). (2009). "Forest dynamics, growth, and yield," in *Forest Dynamics, Growth and Yield, From Measurement to Model* (Berlin; Heidelberg: Springer), 1–39.
- Puttonen, E., Briese, C., Mandlbürger, G., Wieser, M., Pfennigbauer, M., Zlinszky, A., et al. (2016). Quantification of overnight movement of birch (*Betula pendula*) branches and foliage with short interval terrestrial laser scanning. *Front. Plant Sci.* 7:222. doi: 10.3389/fpls.2016.00222
- Raumonen, P., Kaasalainen, M., Akerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., et al. (2013). Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sens.* 5, 491–520. doi: 10.3390/rs5020491
- Rusu, R. B., Marton, Z. C., Blodow, N., Dolha, M., and Beetz, M. (2008). Towards 3D point cloud based object maps for household environments. *Robot. Auton. Syst. J.* 56, 927–941. doi: 10.1016/j.robot.2008.08.005
- Ruts, T., Matsubara, S., Wiese-Klinkenberg, A., and Walter, A. (2012). Aberrant temporal growth pattern and morphology of root and shoot caused by a defective circadian clock in *Arabidopsis thaliana*. *Plant J.* 72, 154–161. doi: 10.1111/j.1365-3113X.2012.05073.x
- Sadava, D. E., Hillis, D. M., Heller, H. C., and Berenbaum, M. (2009). *Life: The Science of Biology*. New York, NY: W. H. Freeman & Co.
- Sasidharan, R., Chinnappa, C. C., Staal, M., Elzenga, J. T. M., Yokoyama, R., Nishitani, K., et al. (2010). Light quality-mediated petiole elongation in *Arabidopsis* during shade avoidance involves cell wall modification by xyloglucan endotransglucosylase/hydrolases. *Plant Physiol.* 154, 978–990. doi: 10.1104/pp.110.162057
- Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., et al. (1997). Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* 275, 502–509. doi: 10.1126/science.275.5299.502
- Stitt, M., and Zeeman, S. C. (2012). Starch turnover: pathways, regulation and role in growth. *Curr. Opin. Plant Biol.* 15, 282–292. doi: 10.1016/j.pbi.2012.03.016
- Sysoeva, M. I., Markovskaya, E. F., and Shibaeva, T. G. (2010). Plants under continuous light a review. *Plant Stress* 4, 5–17. doi: 10.1016/j.tplants.2011.02.003
- Thies, M., and Spiecker, H. (2004). "Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories," in *Proceeding of International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 36, 192–197. Available online at: <http://www.isprs.org/proceedings/XXXVI/8-W2/THIES.pdf>
- van Leeuwen, M., and Disney, M. (2018). *3.08 Vegetation Structure (LiDAR)*. London, UK: Elsevier.
- Vosselman, G., and Maas, H. G. (2010). *Airborne and Terrestrial Laser Scanning*. Boca Raton, FL: CRC Press.
- Weinmann, M., Jutzi, B., and Mallet, C. (2014). Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features. *Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 2, 181–188. doi: 10.5194/isprsannals-II-3-181-2014
- Whippo, C. W., and Hangarter, R. P. (2009). The 'sensational' power of movement in plants a darwinian system for studying the evolution of behaviour. *Am. J. Bot.* 96, 2115–2127. doi: 10.3732/ajb.0900220
- Woo, H., Kang, E., Wang, S., and Lee, K. H. (2002). A new segmentation method for point cloud data. *Int. J. Mach. Tools Manufacture* 42, 167–178. doi: 10.1016/S0890-6955(01)00120-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Herrero-Huerta, Lindenberg and Gard. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Low-Cost Imaging Method for the Temporal and Spatial Colorimetric Detection of Free Amines on Maize Root Surfaces

Truc H. Doan¹, Tu A. Doan¹, Michael J. Kangas², AdreAnna E. Ernest¹, Danny Tran³, Christina L. Wilson², Andrea E. Holmes², Erin L. Doyle¹ and Tessa L. Durham Brooks^{1*}

¹ Department of Biology, Doane University, Crete, NE, United States, ² Department of Chemistry, Doane University, Crete, NE, United States, ³ Crete High School, Crete, NE, United States

OPEN ACCESS

Edited by:

Alexander Bucksch,
University of Georgia, United States

Reviewed by:

Hannes Schmidt,
University of Vienna, Austria
Diana Heuermann,
Leibniz-Institut für Pflanzengenetik
und Kulturpflanzenforschung (IPK),
Germany

*Correspondence:

Tessa L. Durham Brooks
tessa.durhambrooks@doane.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 30 May 2017

Accepted: 16 August 2017

Published: 30 August 2017

Citation:

Doan TH, Doan TA, Kangas MJ,
Ernest AE, Tran D, Wilson CL,
Holmes AE, Doyle EL and Durham
Brooks TL (2017) A Low-Cost
Imaging Method for the Temporal
and Spatial Colorimetric Detection
of Free Amines on Maize Root
Surfaces. *Front. Plant Sci.* 8:1513.
doi: 10.3389/fpls.2017.01513

Plant root exudates are important mediators in the interactions that occur between plants and microorganisms in the soil, yet much remains to be learned about spatial and temporal variation in their production. This work outlines a method utilizing a novel colorimetric paper to detect spatial and temporal changes in the production of nitrogen-containing compounds on the root surface. While existing methods have made it possible to conduct detailed analysis of root exudate composition, relatively less is known about where in the root system exudates are produced and how this localization changes as the root grows. Furthermore, there is much to learn about how exudate localization and composition varies in response to stress. Root exudates are chemically diverse secretions composed of organic acids, amino acids, proteins, sugars, and other metabolites. The sensor utilized for the method, ninhydrin, is a colorless substance in solution that reacts with free amino groups to form a purple dye. A detection paper was developed by formulating ninhydrin into a print solution that was uniformly deposited onto paper with a commercial ink jet printer. This “ninhydrin paper” was used to analyze the chemical makeup of root surfaces from maize seedlings grown vertically on germination paper. Through contact between the ninhydrin paper and seedling root surfaces, combined with images of both the seedlings and dried ninhydrin papers captured using a standard flatbed scanner, nitrogen-containing substances on the root surface can be localized and concentration of signal estimated for over 2 weeks of development. The method was found to be non-inhibiting to plant growth over the analysis period although damage to root hairs was observed. The method is sensitive in the detection of free amines at concentrations as little as 140 μ M. Furthermore, ninhydrin paper is stable, showing consistent color changes up to 2 weeks after printing. This relatively simple, low-cost method could contribute to a better understanding of root exudates and mechanisms used by plants to interact with the complex soil environment during growth and development.

Keywords: rhizosphere, root exudates, maize seedling, colorimetric detection, ninhydrin, printing sensors, root imaging

INTRODUCTION

The rhizosphere is a complex and dynamic microecosystem formed between plant roots and their associated microbes, single-celled eukaryotes, mycorrhizal fungi, and small soil-dwelling multicellular organisms (Badri and Vivanco, 2009). At its core, it is a partnership between roots, which exude a variety of metabolites, ions, and cofactors, and microbes that respond to, utilize, and process them (Badri and Vivanco, 2009; Dennis et al., 2010; Kumar et al., 2012). Interactions of exudates with microbes in the rhizosphere can have significant benefits to the plant including mobilization of soil nutrients, prevention of and enhanced response to infection, and promotion of plant growth and development (Badri and Vivanco, 2009; Kumar et al., 2012). Advances in DNA sequencing technology have shown that microbial composition and abundance of the rhizosphere differs from that of the bulk soil, as evidenced by a decline in diversity and an increase in specialization near the root surface (Micallef et al., 2009; Bulgarelli et al., 2013; Peiffer et al., 2013). Root exudates are thought to be a primary driver of microbial community makeup (Micallef et al., 2009; Dennis et al., 2010; Bulgarelli et al., 2013). Additionally, root exudates play a critical role in plant–plant communication and mediate a variety of positive and negative interactions between nearby plants such as the induction of defense responses (Walker et al., 2003). Therefore, understanding exudate composition and determining genetic and environmental factors that influence its production is a primary focus of crop improvement efforts (Bulgarelli et al., 2013; Chaparro et al., 2013).

Root exudates represent a significant portion (20–50%) of the total metabolism of the plant (Aguilar et al., 2014). They are produced within the root and transported into the nearby soil environment by root epidermal cells and root hairs, and by the production of mucilage and the sloughing off of cells at the root cap (Walker et al., 2003; Dennis et al., 2010). Root exudates are composed of a mix of primary metabolites including sugars, amino acids and organic acids. In smaller abundance are fatty acids, sterols, cofactors, and regulatory molecules (Aguilar et al., 2014). Their composition changes with the plant's genetic makeup, development of the plant and in response to stress. Both biotic and abiotic stressors are able to mediate both positive and negative interactions (via exudate components) between rhizosphere microbes and surrounding plants (Czarnota et al., 2001; Walker et al., 2003; Warembourg et al., 2003; Micallef et al., 2009; Dennis et al., 2010). Amino acids represent the second-most abundant component of root exudates (Chaparro et al., 2013; Moe, 2013). They contribute to nitrogen cycling in the soil and are important in attracting beneficial root microbes (Jones et al., 2002; Oku et al., 2012; Moe, 2013). The amino acid exudate gamma-aminobutyric acid can act as a signal to prime root defenses (Jakab et al., 2001; Moe, 2013). Finally, amino acid exudates, particularly tryptophan, can be important signals and metabolic intermediates affecting root growth and development through the production of phytoactive hormones by plant growth promoting bacteria (PGPB) (Sukumar et al., 2013; Koul et al., 2015).

Maize exudates contain a variety of free-amine containing compounds, primarily in the form of amino acids (Krafczyk et al., 1984; Baudoin et al., 2003; Fischer et al., 2010). Glutamate is thought to be a primary component of this portion of the exudate, though serine, alanine and glycine have also been observed in relatively high proportions (Baudoin et al., 2003; Fischer et al., 2010) and its concentration increases with the presence of growth-promoting rhizobacteria and certain nutrient deficiencies (Krafczyk et al., 1984; Carvalhais et al., 2011). Functionally, maize root exudates play roles similar to what they do in other species including nutrient mobilization, enhanced defense response, and attraction of growth-promoting bacteria (Mench and Martin, 1991; Baetz and Martinoia, 2014; Zhang et al., 2015).

Only recently have metabolomic studies been undertaken to more systematically describe the chemical makeup of exudates (Chaparro et al., 2013). While it is known that apical structures of the root are generally more active in producing exudates, the spatial distribution of exudates and how they change with time remain largely unexplored (Walker et al., 2003; Bulgarelli et al., 2013). Understanding the complex relationships between plant roots, root exudates, and rhizosphere microbes requires detailed measurements of exudate composition and their spatial/temporal distribution. A number of approaches have been published, all of which have advantages and costs. Root exudates may be collected from roots grown in soil, hydroponically, or on other solid support systems such as glass beads, nutrient agar or a capillary mat (Czarnota et al., 2001; Ernst et al., 2010; Dundek et al., 2011; Chaparro et al., 2013; Aguiar et al., 2014; Slota et al., 2016). For some exudate collection methods, roots are washed to remove residual soil, nutrients, and dead cells and are then incubated in a bath solution from which the exudates are recovered (Dundek et al., 2011). This approach, while providing a more natural growing environment, has the disadvantage of being destructive, thereby only allowing a single time point to be collected for each plant. Some experimental setups allow for exudates to be non-destructively sampled as hydroponic media is continuously perfused through growth system and collected, but localization of exudate production in these systems is difficult (Oburger et al., 2013; Slota et al., 2016). Another way to track exudates is to introduce radioactive isotopes in combination with imaging techniques (Oburger and Schmidt, 2016), but introduction of radioisotopes into the lab is not feasible in some research settings.

Water-soluble exudates are typically filtered and concentrated prior to analysis (Dundek et al., 2011). The recovered exudates may be fractionated using chromatography techniques and sample fractions are characterized using UV-vis, mass spectrometry, or NMR (Micallef et al., 2009; Kim et al., 2010; Dundek et al., 2011; Aguiar et al., 2014). In a targeted analysis approach known exudate compounds are detected in the sample. However, methods for non-targeted profiling of exudate composition are rapidly growing and can lead to discovery of new compounds in the exudate and determinations of how exudate composition changes with a given condition (Marti et al., 2013; Oburger et al., 2013). Any metabolomics approach necessarily requires use of highly specialized equipment. A much more accessible approach has been to use colorimetric indicators and assays to more generally assess total carbohydrate, sugar, or

protein content (Van Egeraat, 1975; Dundek et al., 2011). Intact roots can also be exposed to a variety of indicators to detect the location of production and release of specific molecules. For example, blotting exposed plant roots with agar sheets containing the indicators bromocresol purple or BPDS were used to localize production of protons or iron-reducing compounds, respectively (Marschner et al., 1982; Römheld et al., 1984; Neumann et al., 2009). However, the agar sheets must be carefully prepared and stored, which adds difficulty and tedium to the procedure. In summary, there have been a variety of methods developed for collecting, detecting and analyze root exudate, but a need remains for development of simple (and higher throughput) approaches to enable spatial and temporal descriptions of exudate production.

We have developed a simple, minimally disruptive method for detecting and measuring root exudate production on the surface of a growing root (**Figure 1**). The colorimetric indicator ninhydrin has been used to detect free amine groups in exudates on the root surface. Ninhydrin is a useful compound for this purpose due to the easily visualized purple-colored product that is formed in the presence of analyte. Ninhydrin is also sensitive to free amines at physiological concentrations (Van Egeraat, 1975; Dundek et al., 2011; Chutipongtanate et al., 2012). A novel process for printing the indicator on commercially available tissue paper was developed and roots grown vertically on germination paper were blotted with this paper to capture a snapshot of free-amines on the root surface. Images of blotted ninhydrin papers and seedlings were captured and image processing tools developed to visualize free amine localization and relative concentration. Because our method does not disturb the plant as it grows and the indicator formulation is non-toxic to the seedlings, it can be used to collect time course data for individual plants. Additionally, we have determined that the ninhydrin-printed paper is shelf-stable and reliably detects root exudates at biologically relevant concentrations. Here, we demonstrate the application of our method to maize roots, but it could readily be expanded to other plant systems and indicators.

MATERIALS AND METHODS

Preparation of Ninhydrin Paper

All reagents and chemical products were purchased from Sigma-Aldrich (St. Louis, MO, United States) at technical grade or better and were used without further purification. The print solution was composed of 71.2% (v/v) sodium acetate buffer (0.1 mM, pH 5), 14.8% glycerol, 8.2% diethylene glycol, and 5.8 % of triethylene glycol monobutyl ether.

Ninhydrin ink (3% w/v) was prepared using the print solution as the solvent and sonicated (Bransonic ultrasonic bath, CPX2800H) for 3 h at room temperature. The solution was stored in an amber bottle in a dark cabinet at room temperature.

The 3% ninhydrin ink was added to an empty cartridge (InkOwl, Champlain, NY, United States) using a syringe (with 0.8 μ m filter) and placed into the printer (Canon MG5520). The high quality setting on the printer was used. Commercially available white tissue paper (Hallmark, Kansas City, MO, United States) cut to size (8.5 in \times 12 in) was taped onto regular copy paper using all-purpose labels (Avery, 6737) to facilitate feeding of the tissue paper through the printer. The template provided in **Supplementary Presentation S1** was used to uniformly print ninhydrin on the paper, except for two 3 \times 3 grids used for paper calibration and quantification of exudates. After printing, the ninhydrin paper was placed in an opaque envelope and stored at ambient temperature and humidity until used.

Preparation of Mock Exudates

A mock exudate solution was prepared according to Baudoin et al. (2003) for use in determining ninhydrin paper sensitivity and longevity. The 70X mock exudate consisted of 18.4 mM glucose, 18.4 mM fructose, 9.2 mM sucrose, 9.2 mM citric acid, 18.4 mM lactic acid, 13.8 mM succinic acid, 9.2 mM alanine, 9.2 mM serine, and 5.5 mM glutamic acid. The total free amine concentration of 70X mock exudate was 23.9 mM. When diluted to 1X with distilled water, carbon content of mock exudate was at a concentration similar to

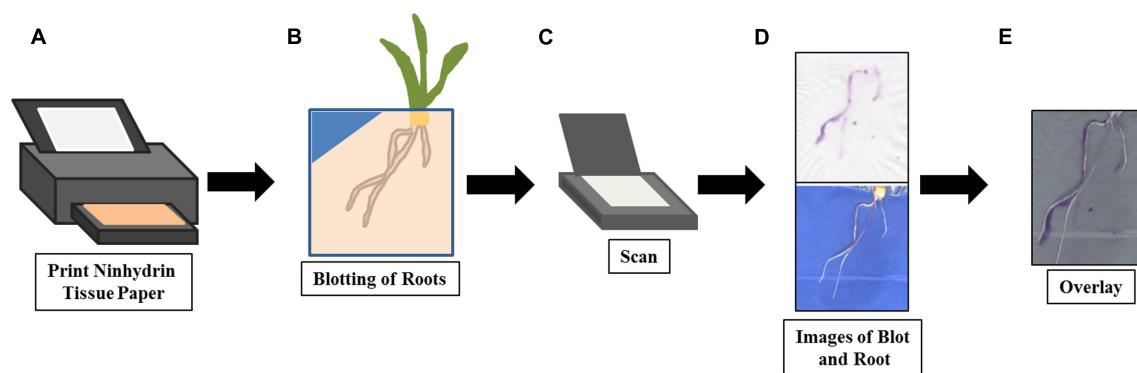


FIGURE 1 | Method overview for root exudate imaging with ninhydrin. The ninhydrin paper was printed (**A**) with 3% ninhydrin ink solution and allowed to dry. On data acquisition days, the maize root was blotted gently (**B**) with ninhydrin paper to absorb the exudates released by the root. After the blotted ninhydrin paper was allowed to develop, it and the root were scanned (**C**) using a flatbed scanner and images collected (**D**). Finally, the ninhydrin paper and seedling images were overlaid (**E**) to visualize localization of exudates.

previously published values (Iijima et al., 2000) and the C/N ratio was similar to previously reported values (Krafczyk et al., 1984).

Germination and Planting

Seeds of maize inbred genotype B73 were sterilized by washing with 70% ethanol for 3 min and 2% (w/v) bleach for 5 min (Naveed et al., 2014). The seeds were then washed three times with autoclaved distilled water, stirring for 1 min for each wash. The seeds were rolled in paper towels wetted with sterile distilled water. The paper rolls were then covered by aluminum foil and imbibed at 4°C for 2 days (due to improved germination of B73 during cold imbibition) then moved to a growth chamber (Percival, Geneva Scientific) set at 24°C with continuous light (to avoid circadian effects) and with a 75% relative humidity (van de Venter and Grobbelaar, 1985; Naveed et al., 2014).

Three-day old seedlings were planted using a pouch method (Hund et al., 2009). Blue germination paper (Anchor Paper Co.) was trimmed to 8.5 in × 12 in, and a small triangle was removed from the center of the top edge. Seedlings were positioned at the apex of the triangle. Polyethylene sheets were cut approximately to the size and shape of the germination paper and positioned over it. Seedlings were held to the germination paper by placing paper clips on each side of the triangle apex. Plastic sheets and blotter germination paper were clipped to a wooden dowel using two binder clips and the dowel was used to hang the entire seedling apparatus using either end of the rim of 8-gallon opaque plastic containers (**Supplementary Figure S1**). Each 8-gallon containers held up to ten seedlings and were filled with enough sterilized distilled water to reach the bottom of the germination paper. Lightweight black cloth was draped over the either side of the seedling to keep the roots relatively dark. Seedlings were grown at 24°C with continuous light and 75% relative humidity.

Blotting and Scanning

Pouches were removed from the growth chamber every 2 days for scanning and blotting. The polyethylene sheet covering the surface of the root was pulled back to expose the root surface for blotting. To minimize water contacting and wrinkling the ninhydrin paper during blotting, laminated paper was used to cover portions of the germination paper where the root had not grown. The ninhydrin paper was gently pressed to the root for approximately 10 s (**Supplementary Figure S2**). After blotting, the ninhydrin paper was dried for approximately 24 h at room temperature with compression in a press (Carolina, Cat #: 663050) to eliminate wrinkles in the paper during scanning. The blotted ninhydrin papers were placed on the flatbed surface and color images were acquired using an EPSON V700 scanner and VueScan x64 software with settings shown in Supplementary Table S1.

An EPSON V700 scanner using VueScan x64 software with settings shown in Supplementary Table S2 was used for collecting RGB images of the roots. To minimize bubbles on the blue blotter germination paper and to ensure a consistent background, a clear plastic sheet was placed on the surface of the scanner, and the surface of the scanner was flooded with a small volume

of water that was contained within the beveled edge of the scanner housing. The plastic sheet covering the root was lifted off the root surface and the seedling pouch was carefully placed onto the flooded scanner surface with the seedling facing down. Pouches were oriented in the same manner as the ninhydrin papers.

Image Analysis

ImageJ software, which is freely available, open-source, and able to run on nearly any operating system, was used for analysis of spot intensities and root lengths (Schneider et al., 2008). Primary root length was measured using the segmented line tool. For determining paper sensitivity and stability, average pixel intensity of a spot created by depositing 1 µL of mock exudate on ninhydrin papers was determined by visually selecting areas containing signal with the magic wand tool. The average pixel intensity was quantified using the RGB measure function. Average pixel intensity was calculated as the weighted sum of the red, green and blue channels (0.30 R + 0.59 G + 0.11 B). This method can be set as an option in ImageJ (Edit/Options/Conversions menu).

The program GIMP2 (GNU Image Manipulation Program)¹ was used to produce overlay images between signals in a ninhydrin paper and seedling roots growing on germination paper. After opening the seedling root image it was converted to grayscale using the Colors/Desaturate command. The ninhydrin paper image was imported by dragging it onto the seedling image or by using File/'Open as Layers' from the seedling image. The ninhydrin paper layer was overlaid using a multiply operation. The 'Multiply' option was selected from the 'Mode' pull-down menu in the 'Layers' toolbox to produce the overlay. The orientation of the ninhydrin paper was adjusted using the Move Tool as needed. The image overlay was exported in the desired format using the File/Export command.

Determination of Seedling Growth Rate

Seedling growth rate (R) for a given day (D) was determined using the following formula where (L) is length in cm:

$$R_D = \frac{L_{D+2} - L_D}{2} \quad (1)$$

Paper Calibration

To calibrate ninhydrin papers for root exudates quantification, three replicate 0.1 µL aliquots of mock exudate containing free amine concentrations of 1.71, 3.41, or 4.71 µM were deposited on one of the two printed grids on the paper immediately prior to time of blotting. After a 24 h incubation period in a press and imaging the paper as described above, the magic wand tool (ImageJ) was used to manually capture the developed areas for each aliquot. The average pixel intensity of each calibration spot was calculated in ImageJ as described above. For each of the three mock exudate concentrations, the lower and upper quartiles (Q3 and Q1) and the interquartile range (IQR, calculated as Q3-Q1) for average pixel intensity were

¹<http://www.gimp.org>

calculated using the calibration spots of that concentration from the blotted ninhydrin papers for 3-, 5-, and 7-day old seedlings. An individual calibration spot was considered to be an outlier if its average pixel intensity was less than $Q1 - 0.5 \cdot IQR$ or greater than $Q3 + 0.5 \cdot IQR$ for the concentration. Papers with two or more outlier calibration spots for a single mock exudate concentration were determined to be outlier papers and were removed from the dataset.

For the retained papers, a linear regression was used to model the relationship between the known concentration of mock exudate applied and the measured pixel intensity of each spot. The regression equation is shown below, where y is the average pixel intensity and x is the free amine concentration in the mock exudates ($R^2 = 0.72$, $p = 0.0$):

$$y = -20.98x + 241.71 \quad (2)$$

A 95% confidence interval (CI) was calculated for the regression line to determine the range of concentrations expected based on given average pixel intensity values.

In order to visualize exudate concentrations on the root surface, the blotted and scanned ninhydrin paper images were read using the Python OpenCV library and converted to gray-scale. Using the regression equation and computed CI, pixel intensities across the entire image were converted to estimated free amine concentrations. A heat map function was applied to re-draw the image with approximate exudate concentrations indicated by a color gradient. A custom Python script for executing this visualization is provided as **Supplementary Data Sheet S1**.

RESULTS

Method Overview

The method workflow is illustrated in **Figure 1**. Seedlings are grown using a modified pouch method, which allows easy access and visualization of the root surface (Hund et al., 2009). The roots are gently pressed against paper printed with a ninhydrin solution (**Figures 1A,B**). The blotted ninhydrin paper and the seedling roots are scanned and the resulting digital images overlaid to identify regions of the root that are active in nitrogenous exudate production (**Figures 1D,E**). After calculating the average pixel intensities of spots containing known free amine concentrations, outlier papers can be identified and free-amine containing exudate concentration in the remaining papers can be estimated (not shown).

Ninhydrin Paper Sensitivity to Free Amines

The sensitivity of ninhydrin paper to the detection of free amines was determined by measuring the average pixel intensity of selected regions of 3% ninhydrin paper spotted with 1 μ L of mock exudate solutions containing known free amine concentrations (**Figure 2**). Free amine concentrations between 0 and 3.41 mM were tested. A grid of nine 1 μ L volumes for each mock exudate concentration was deposited onto on four separate ninhydrin papers. The aliquoted volumes were allowed to spread, creating uniform spots on the paper. Papers were stored in an opaque envelope in the dark overnight in order to develop the indicator. The colorimetric transition from white/pale yellow

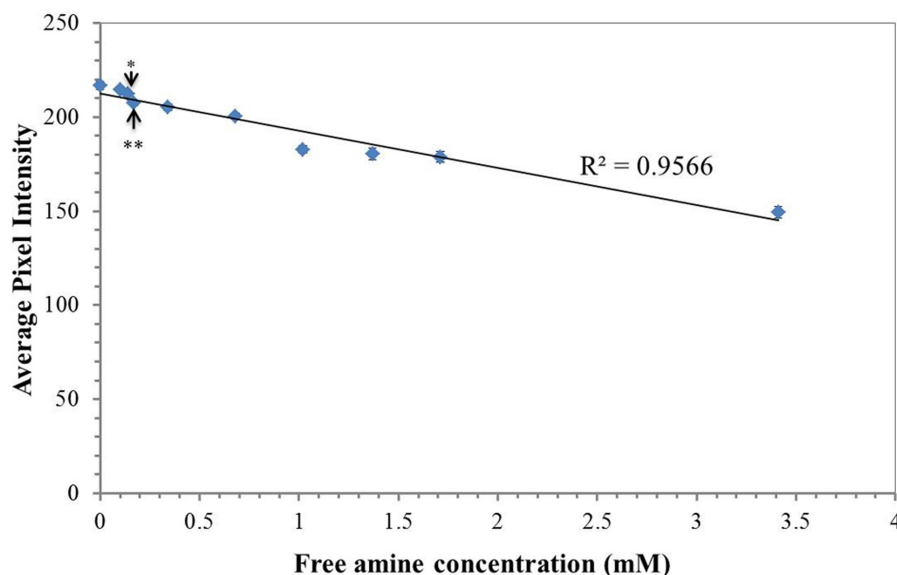


FIGURE 2 | Paper sensitivity for free amines. Average pixel intensity ($0.30 R + 0.59 G + 0.11 B$) of selected regions of 3% ninhydrin paper spotted with 1 μ L of a solution containing various concentrations of free amines in mock exudates. Points are the mean ($n = 36$, 4 sheets, 9 spots measured per sheet) \pm standard error. ** indicates a statistically significant difference ($p < 0.01$) between each sampling concentration to the control (water, 0 μ M free amine). Statistical significance was determined using a two-way ANOVA with exudate free-amine concentration and paper ($n = 4$) as the factors and average intensity as the dependent variable. Pairwise comparisons were made from a *post hoc* Tukey's HSD. All points more concentrated than the free amine concentration indicated by the double asterisks were significantly different from the control ($p < 0.01$).

to purple was apparent within this time frame. Papers were scanned and average pixel intensity determined using ImageJ. Free amine concentrations as little as $140\ \mu\text{M}$ were detected based on statistically significant differences compared to the control ($p < 0.01$, two-way ANOVA, Tukey *post hoc*). All concentrations tested were within the linear range of the indicator based on the linear trend of indicator response with concentration (Figure 2) and within the physiological range of the amine content expected in the exudate (Baudoin et al., 2003). Therefore, ninhydrin paper is sensitive to exudates within the physiological range.

Ninhydrin Paper Stability in Atmospheric Conditions

The optimal use period, or time from printing until which a stable signal could be achieved, was determined for conditions typical to the laboratory (room temperature and ambient humidity). Ninhydrin paper was treated with mock exudate containing $3.41\ \text{mM}$ free amines every day for 21 days after printing with 3% ninhydrin print solution (Figure 3). The ninhydrin paper was stored in an opaque envelope to protect it from light. The average pixel intensity was found to change most dramatically the first day after printing as observed by the increase from 134.77 ± 2.63 to 147.29 ± 3.09 . This is hypothesized to be the result of the ninhydrin ink drying which at atmospheric humidity and temperature takes at least 2 days to complete. On day three to fourteen the RGB intensity stabilized to an average value of 150 with insignificant fluctuation in signal during this period ($p > 0.05$, two-way ANOVA, Tukey *post hoc*). After 21 days in storage a significant drop in ninhydrin paper sensitivity was observed which is hypothesized to stem from interaction of the paper with humid air or light, which could likely be mitigated by storage of the paper under more controlled conditions. Based on

these data, ninhydrin paper was conservatively deemed optimal for blotting 4–13 days after printing.

Seedling Sensitivity to Ninhydrin Paper

As root growth is sensitive to environmental factors, such as chemical and mechanical stress, the impact of blotting with ninhydrin paper on root growth was determined. Insignificant differences were observed in growth length (Figure 4) and growth rate (Figure 5) between roots that were not blotted (Control), roots blotted with ninhydrin-free paper, and roots blotted with 3% ninhydrin paper. Seedlings that were blotted with ninhydrin paper were slightly, but significantly longer than roots used for the other two groups at the start of the experiment (Figure 4, $p < 0.01$, Student's *T*-test). Ninhydrin-free paper was used to determine the impact of mechanical stress of blotting while use of 3% ninhydrin paper indicates the potential impact of mechanical stress as well as chemical stress should ninhydrin be transferred from the paper to the root during contact. Roots achieved linear growth rates from day three through nine and plateaued in growth at around 15–18 cm after experimental day thirteen (Figure 4). The trend in growth rate was similar in all treatment groups (Figure 5) with roots growing at a steady rate of $4.5\ \text{cm/day}$ from days three through eight of blotting followed by incremental decreases to 3.3, 1.2, 0.5, and $0.25\ \text{cm}$ per day days 9, 11, 13, and 15 respectively. Overall, growth length and rate remained unaffected by mechanical stress and roots did not show any indication of toxicity to ninhydrin exposure.

Root Hair Growth and Exudate Localization

In order to assess the impact of the method on root hair development, digital images taken before and after scanning were

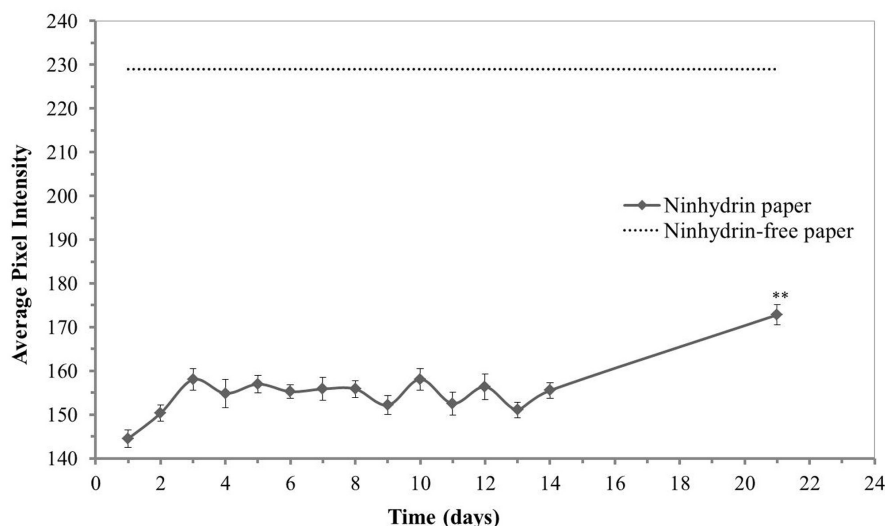


FIGURE 3 | Stability of ninhydrin paper. Average pixel intensity ($0.30\text{ R} + 0.59\text{ G} + 0.11\text{ B}$) of selected regions of 3% ninhydrin paper spotted with $1\ \mu\text{L}$ of a mock exudate solution containing $3.41\ \text{mM}$ free amines then developed overnight. Values presented are means \pm standard error ($n = 27$, 3 sheets, 9 spots measured per sheet). *Post hoc* analysis (Tukey's HSD) of a two-way ANOVA with paper ($n = 3$) and mock exudate free-amine concentration as factors and pixel intensity as the dependent variable determined that day 21 intensities were significantly different from every other day ($p < 0.01$).

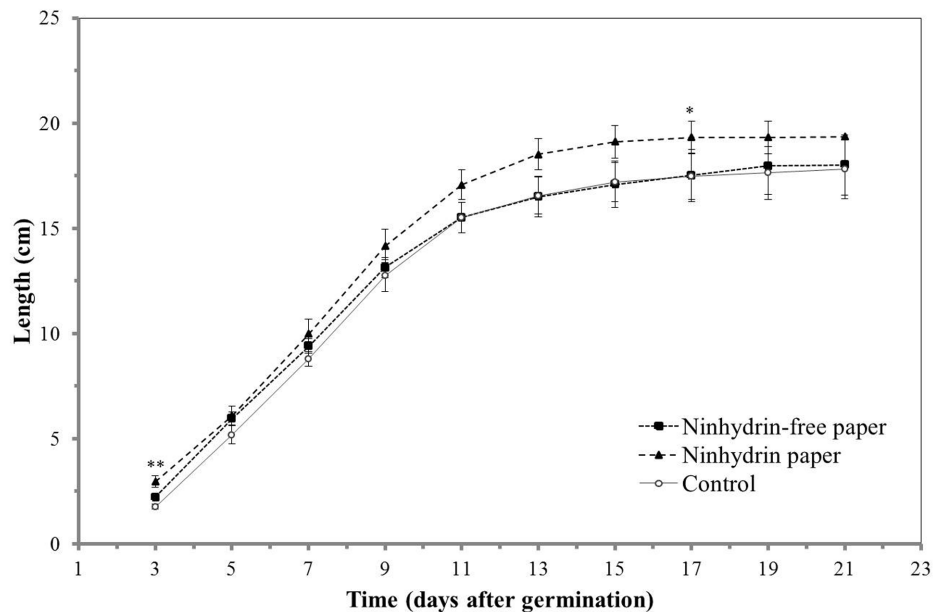


FIGURE 4 | Impact of ninhydrin paper blotting on root length. Growth of B73 seedling roots along the germination paper surface without blotting (Control), while blotting seedlings with plain paper every 2 days (Ninhydrin-free paper), or while blotting with 3% ninhydrin paper every 2 days (Ninhydrin paper). Germination occurred at day 0. The data are the mean of 10 seeds for control, 9 seeds for the ninhydrin-free paper control, and 10 seeds for ninhydrin paper. Error bars are the standard error of the mean. * indicates a statistically significant difference ($p < 0.05$) and ** indicates a statistically significant different ($p < 0.01$) between the ninhydrin paper condition and the control as determined by a Student's *t*-test. Note that seedlings blotted with ninhydrin paper were slightly, but significantly longer than roots used for the other two groups at the start of the experiment.

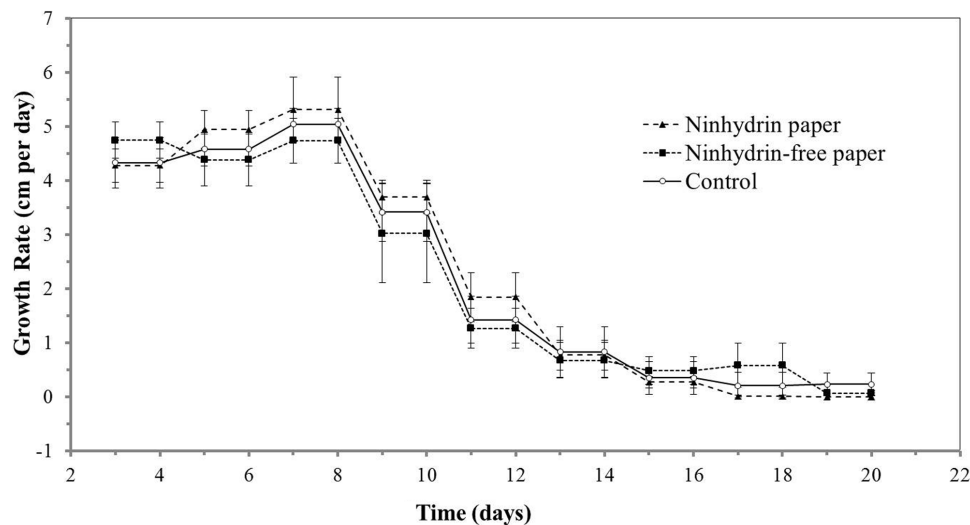


FIGURE 5 | Growth rates of roots exposed to ninhydrin paper. Growth rate (cm/day) of B73 seedling roots along the germination paper surface unperturbed without ninhydrin (Control), while blotting with plain paper every 2 days (Ninhydrin-free paper), or while blotting with paper containing 3% ninhydrin every 2 days (Ninhydrin paper). Germination occurred at day 0. The data are the mean of 10 seeds for control, 9 seeds for the blotting control, and 10 seeds for 3% ninhydrin. Error bars are the standard error of the mean.

inspected. The blotting method damaged root hairs. Regions of the root containing root hairs before blotting (**Figure 6A**) no longer contain them after blotting (**Figure 6C**). However, root hairs were again found on regions of new growth within 2 days after blotting (**Figure 6D** and **Supplementary Figure S3**).

While damage to root hairs contributes to some of the signal observed on the ninhydrin paper, signal was also observed in regions of the root that do not contain root hairs. For example, exudate production in the 3-day-old seedling in **Figure 6** is largely localized to the root tip and elongation

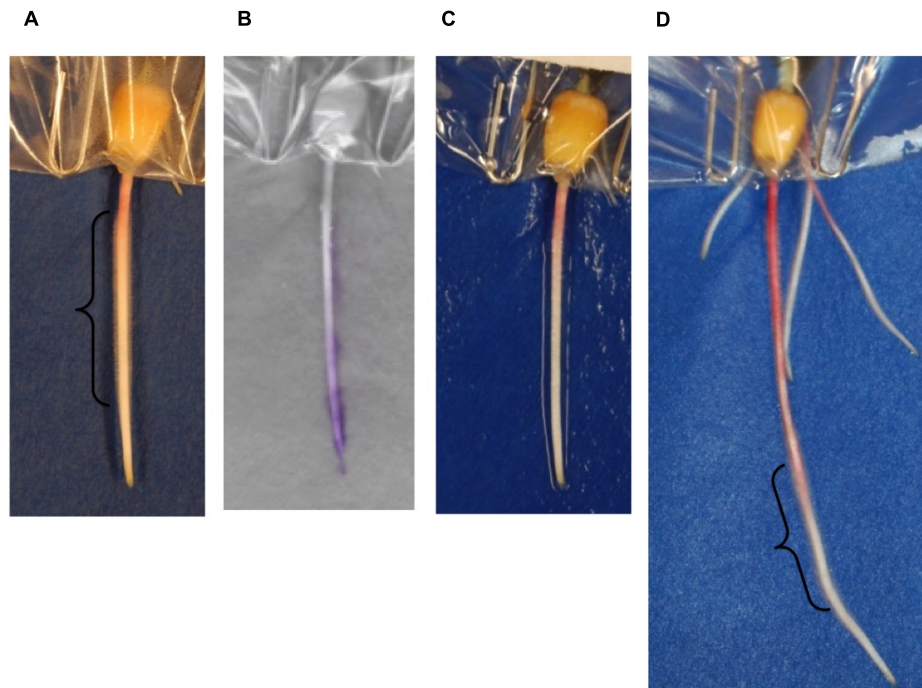


FIGURE 6 | Effects of blotting ninhydrin paper on root hairs. A 3-day-old seedling root **(A)** before blotting, **(C)** after blotting, and **(B)** the overlay of the same seedling with developed ninhydrin paper. **(D)** The same root 48 h later. Regions of the root in which root hairs were observed are shown with a bracket.

zone, rather than the region in which root hairs were present (**Figure 6B**).

In order to gain an initial characterization of root exudate production in young seedlings, roots growing in distilled water were blotted every other day over the course of several weeks with 3% ninhydrin paper. Grayscale images of scanned seedlings were overlaid with the corresponding developed ninhydrin paper. The first week of growth is shown for a representative B73 seedling in **Figure 7**. As expected, exudate production declined with age and was primarily localized to the apical regions of the root. By days five and seven, signal could also be observed at the base of the root and at the tip of seminal roots. Even though seedlings were grown in distilled water, signal could still be detected in some 17- and 19-day-old plants (**Figure 8**).

Quantification of Detected Free Amines

In order to facilitate relative quantification of signal within and across papers, two 3×3 grids of ninhydrin were printed on the bottom corner of each ninhydrin paper (**Supplementary Presentation S1**). Directly before blotting, mock exudates of known concentrations (three replicates of three standard concentrations) were applied to one of the grids (see Materials and Methods). After scanning, the developed spots on each grid were analyzed as described in the Materials and Methods to calculate mean pixel intensity of each spot. Outlier papers were detected and removed from the dataset using a Tukey fence as described in the Section “Materials and Methods.” Overall, 20 of 30 papers collected from seedlings 3–7 days old were retained for analysis.

After outlier papers were removed, a linear regression of all spots for the remaining 20 papers was calculated to predict free amine concentration based on average RGB intensity. The regression was significant ($p < 0$) with an R^2 of 0.72. The upper and lower 95% CIs were used to determine the range in signal intensities that could be expected at each concentration of standard. These values ranged from 19 intensity units at the lowest standard (1.71 mM) to 30 intensity units at the highest standard (4.71 mM). The slope of the regression was used to convert these ranges into concentrations, which ranged from 1.0 mM at the lowest standard to 1.6 mM at the highest standard. The average range in concentration across the three standards was 1.3 mM. This value estimates the approximate resolution in quantifying the concentration of signal on the papers. In order to visualize and compare free amine signal intensities across papers, a heatmap was applied to the grayscale images of developed ninhydrin paper using a custom Python script (**Supplementary Data Sheet S1**). The resulting visualization enabled estimation of exudate production levels within and between ninhydrin papers. The standardized color scale that was determined from the linear fit of the retained data is shown in **Figure 9A**.

Closer inspection of grids from papers retained after removal of outliers and application of the heatmap revealed that the standards for approximately 65% of papers were colored to indicate the expected concentrations on the color scale. An example of two grids displaying this pattern are shown in **Figure 9B**. The grids of standards and the corresponding root signals shown in **Figure 9B** come from blots of the same seedling at 3 and 5 days after germination. For papers in this category,

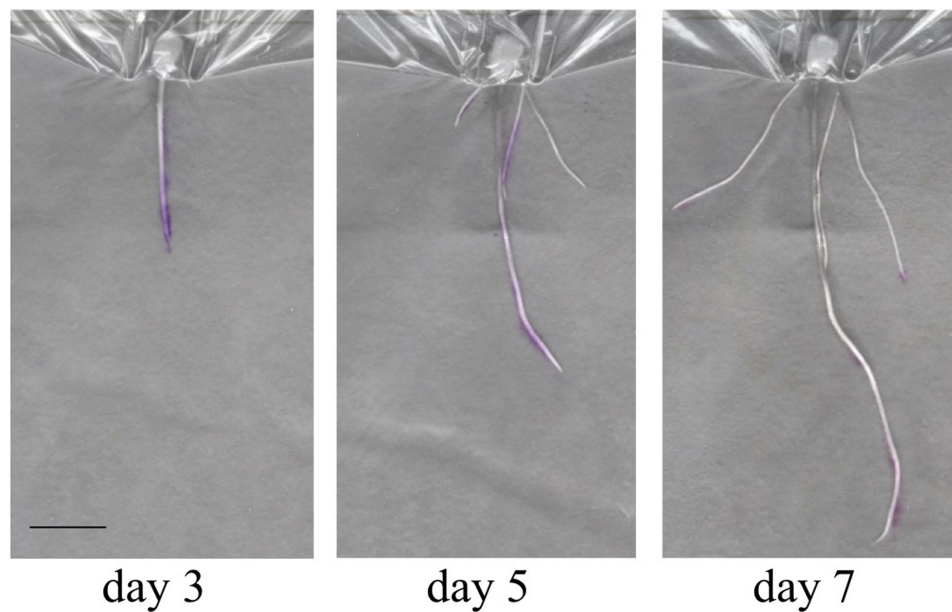


FIGURE 7 | Overlay images of a seedling during the first week of growth. A single seedling during the first week of development shows production of free amines in the root tip and elongation zone (Day 3), the root base (Day 5) and at the tips of seminal roots (Day 7). Scale bar is 2 cm.

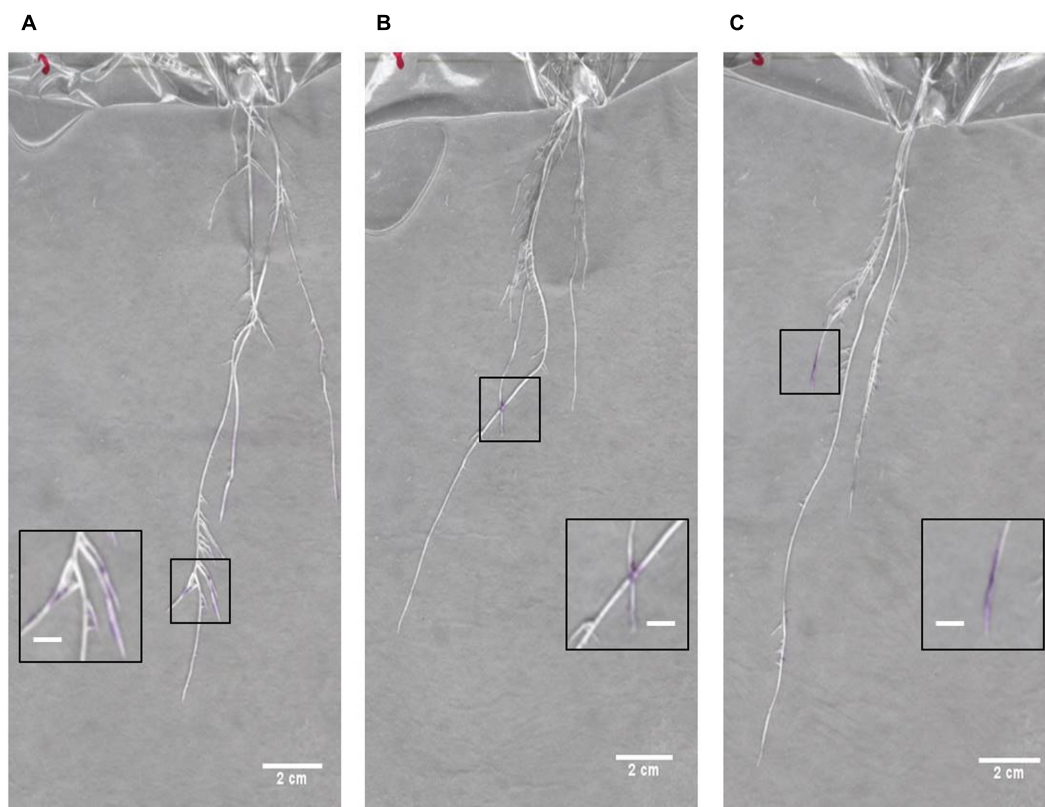
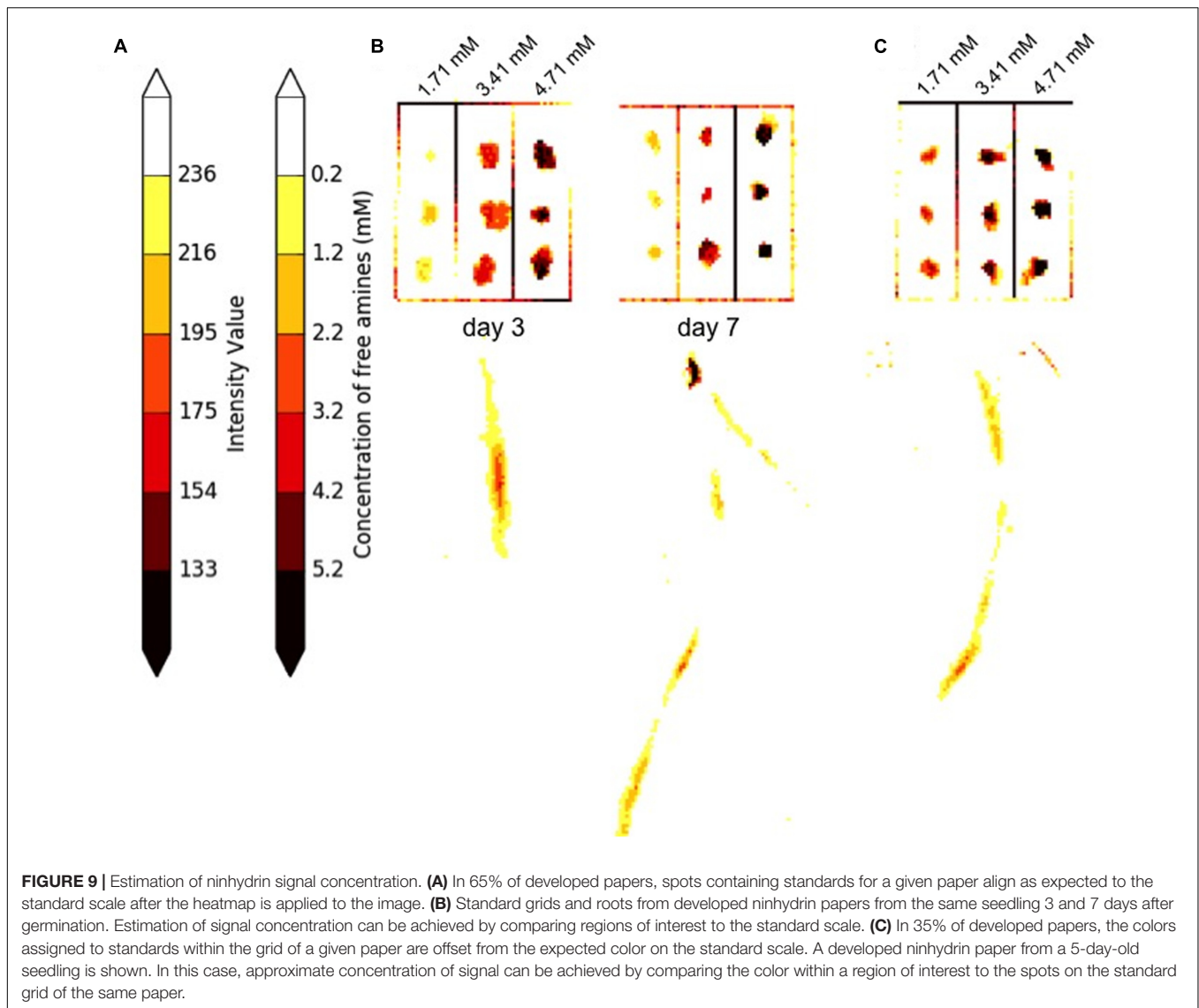


FIGURE 8 | Overlay images of seedlings at 17–19 days after germination. Overlays of seedlings and developed ninhydrin paper at **(A,B)** 17 days after germination and **(C)** 19 days after germination. Insets show zoomed-in regions containing ninhydrin signal on each root system. Scale bars of main panels are 2 cm. Scale bars of inset panels are 0.5 cm.



regions of interest can be referenced to the standard color scale shown in **Figure 9A** for approximation of signal concentration of signal within and between papers. The data shown for the seedling in **Figure 9B** indicates that free amine production did not exceed 4.2 mM at day 3 or 7. Free amines were concentrated at a relatively apical region of the root on day 3. Signal was still observed in apical regions of the root at day 7, but the most concentrated region of signal was found more basally than was observed at day 3.

A smaller proportion of retained papers contained grids with standards that were overexposed or underexposed compared to the expected values on the color scale. An example grid and corresponding root signal from a 5-day-old seedling are shown in **Figure 9C**. The data indicates that free amine production anywhere in this root was less than 4.71 mM and that much of the production was less than 1.71 mM. The signal concentration for papers in this category can be estimated by comparing regions of interest to the grid standards found on the same paper, but

these estimations should not be made by comparing signals to the standardized color scale.

DISCUSSION

This paper describes a cost-effective method for localization and low-resolution quantification of root exudate production during development. Roots are sensitive to chemical and physical stress, and exudate production or growth may be altered as a result of this stress, so any method developed for the purpose described here would aim to minimize such disruption (Badri and Vivanco, 2009; Baetz and Martinoia, 2014). Root hairs are recognized as important cells in production of root exudate and are estimated to make up as much as 77% of the total root surface area (Badri and Vivanco, 2009). While damage to root hairs was observed during our blotting procedure, regrowth of these cells was observed on new root tissue within 2 days (**Figure 6**). This observation and the

fact that the overall root growth rate was unaffected by blotting contribute to its utility in quickly gathering information about exudate production in a living plant. The indicator ninhydrin, which reacts with free amine groups, was chosen in part because amino acids represent a major component of root exudates and also because the indicator creates a color change that is easily detected using standard light imaging (Chutipongtanate et al., 2012; Chaparro et al., 2013; Moe, 2013). These features make it possible to explore adaptation of the method to other plant species and experimental conditions. Similarly, there are many more indicators that could be suitable for use in this method, such as pH indicators (Neumann et al., 2009). A future direction will be to identify additional chemical sensors for detection of other exudate components, including sugars.

Methods have been previously developed for visualization of root exudates within the soil environment and include use of rhizotrons or rhizobox systems for in situ sampling or visualization (Haase et al., 2007; Shi et al., 2011; Oburger et al., 2013; Oburger and Schmidt, 2016). Exudates can be sampled by exposing the root/soil interface with agar containing pH indicators or complexation and chelation agents (Neumann et al., 2009). Previous studies have also described use of colorimetric indicators for visualization of exudate production (Van Egeraat, 1975; Römheld et al., 1984; Dundek et al., 2011). However, while these methods have the advantage of allowing for more precise quantification of exudates within a more natural growth environment, the method described here is generally cheaper and simpler, while minimizing exposure of the root to the visualization chemicals making it possible to collect multiple samples from one plant and to potentially collect data at higher throughput.

Total free amine concentrations have been estimated at approximately 100–300 μM in soil extracts (Jones et al., 2002; Baudoin et al., 2003). However, growing conditions, particularly hydroponic growth systems, have previously been reported to limit free amine content of maize exudates (Oburger et al., 2013). Nitrogen limitation is an additional factor limiting free amine abundance in exudate (Carvalhais et al., 2011). The described method introduced both of these limiting factors. However, our ninhydrin paper, which was capable of detecting free amines in solution with a lower concentration limit of 140 μM (Figure 2), enabled visualization of free-amine containing compounds on the root surface despite these potential inhibiting effects. Therefore, we expect that our method could be useful for exudates detection across a range of plant species and experimental conditions. The root tip and apical meristem of the growing root, as well as root hairs, are known to be active regions of exudate production (Jones et al., 2002; Dennis et al., 2010). This agrees with the signal patterns observed in our data both for early seedling and later seedling growth (Figures 7–9). However, the method developed could be used for studies to detect exudate production patterns that may differ from this established scheme.

The method we have developed describes a semi-quantitative approach for visualization of root exudate production on the root surface. Next steps will include improving quantitative measurements of exudate production and automation of the

image processing tasks. Future work will also include exploring nitrogen sources that can be added to the growth media to encourage normal growth and enhance exudate production and root growth while minimizing the amount of background signal. Other colorimetric sensors will be explored to improve sensitivity and to allow for detection of other types of exudates.

CONCLUSION

The described method is simple and easy to use, consumables are inexpensive, and materials are commercially available. A flatbed scanner and open source software (Image J, GIMP2, Python/OpenCV) was used for overlaying and analyzing root images. The intermittent application of the ninhydrin paper did damage root hairs, but did not prevent normal root growth.

The rapid and low-cost detection of root exudates containing free amines was achieved by utilizing a novel colorimetric paper to detect spatial and temporal changes in exudate production on the root surface through the first 2 weeks of development.

AUTHOR CONTRIBUTIONS

THD and TAD made substantial contributions to the study concept and design, ninhydrin paper preparation, seedling growth, image data collection and analysis, prepared figures and references, and provided relative interpretation. MK made substantial contributions to the study concept and design, ninhydrin paper preparation, image data collection process, and handled manuscript preparation including preparation of figures and references, gave intellectual contributions, improvements and revisions. AE and DT wrote Python scripts for visualizing ninhydrin signal concentrations and helped prepare the related figure. CW, AH, and ED drafted and revised the article including preparation of figures, gave intellectual contributions, improvements, and revisions. TDB conceived and supervised the research in all aspects and made significant intellectual contributions in outlining the article, article content, and critically revising the article for final submission.

FUNDING

This work was funded by the National Science Foundation (NSF) EPSCoR Research Infrastructure Improvement (RII) Track-1 (1557417), National Institute for General Medical Science (NIGMS) (5P20GM103427), a component of the National Institutes of Health (NIH), SBIR Phase II - Chemical Biological Radiological Nuclear and Explosives (CBRNE) Reconnaissance Sampling Kit (W911SR-16-C-0051), and the Camille and Henry Dreyfus Foundation.

ACKNOWLEDGMENTS

We would like to acknowledge Rachel Lukowicz, and Armando Pliego for their assistance in training and help with formulation

of indicators. We would like to acknowledge Keeliann Mark and Michael Tross for assistance with printing and image analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01513/full#supplementary-material>

FIGURE S1 | Seedling growth pouch setup. A seedling growth apparatus was constructed according to Hund et al. (2009). Seedlings were grown vertically on germination paper and a polyethylene sheet was used to maintain moisture on the sheet and to help affix the seedling to the paper surface. Seedlings were grown in distilled water. Up to 10 plants were grown together in one container. Seedling

roots were shaded from light in the growth chamber by draping the containers with sheets of lightweight black fabric.

FIGURE S2 | Illustration of the blotting method. Shown is a sheet of ninhydrin paper that has been laid on top of a piece of a seedling grown as described in Section “Materials and Methods.” Regions of the germination paper that do not contain seedling have been shielded using laminated paper. The tissue paper is lightly blotted to ensure contact of all root surfaces with the ninhydrin paper.

FIGURE S3 | Detailed view of root hair regrowth after blotting. A detailed view of regions of the roots from **Figures 6A,D** are shown in **(A,B)**, respectively. **(A)** Well-developed root hairs of a 3-day-old seedling root that has never been blotted. **(B)** A section of the same root is shown 48 h later. The initiation of root hair development within this timeframe can be observed.

PRESENTATION S1 | Printing template for ninhydrin papers.

DATA SHEET S1 | Custom script for applying a heatmap to calibrated ninhydrin papers.

REFERENCES

- Aguiar, N. O., Vega, M. R. G., de Oliveira, R. R., Olivares, F. L., da Silva Lima, L., and Canellas, L. P. (2014). Root exudate profiling of maize seedlings inoculated with *Herbaspirillum seropedicae* and humic acids. *Chem. Biol. Technol. Agric.* 1, 23. doi: 10.1186/s40538-014-0023-z
- Badri, D. V., and Vivanco, J. M. (2009). Regulation and function of root exudates. *Plant Cell Environ.* 32, 666–681. doi: 10.1111/j.1365-3040.2008.01926.x
- Baetz, U., and Martinoia, E. (2014). Root exudates: the hidden part of plant defense. *Trends Plant Sci.* 19, 90–98. doi: 10.1016/j.tplants.2013.11.006
- Baudoin, E., Benizri, E., and Guckert, A. (2003). Impact of artificial root exudates on the bacterial community structure in bulk soil and maize rhizosphere. *Soil Biol. Biochem.* 35, 1183–1192. doi: 10.1016/S0038-0717(03)00179-2
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu. Rev. Plant Biol.* 64, 807–838. doi: 10.1146/annurev-arplant-050312-120106
- Carvalhais, L. C., Dennis, P. G., Fedoseyenko, D., Hajirezaei, M.-R., Borriss, R., and von Wirén, N. (2011). Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. *J. Plant Nutr. Soil Sci.* 174, 3–11. doi: 10.1002/jpln.201000085
- Chaparro, J. M., Badri, D. V., Bakker, M. G., Sugiyama, A., Manter, D. K., and Vivanco, J. M. (2013). Root exudation of phytochemicals in Arabidopsis follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLoS ONE* 8:e55731. doi: 10.1371/journal.pone.0055731
- Chutipongtanate, S., Watcharatanyatip, K., Homvises, T., Jaturongkakul, K., and Thongboonkerd, V. (2012). Systematic comparisons of various spectrophotometric and colorimetric methods to measure concentrations of protein, peptide and amino acid: detectable limits, linear dynamic ranges, interferences, practicality and unit costs. *Talanta* 98, 123–129. doi: 10.1016/j.talanta.2012.06.058
- Czarnota, M. A., Paul, R. N., Dayan, F. E., Nimbal, C. I., and Weston, L. A. (2001). Mode of action, localization of production, chemical nature, and activity of sorgoleone: a potent PSII inhibitor in *Sorghum* spp. Root exudates1. *Weed Technol.* 15, 813–825.
- Dennis, P. G., Miller, A. J., and Hirsch, P. R. (2010). Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiol. Ecol.* 72, 313–327. doi: 10.1111/j.1574-6941.2010.00860.x
- Dundek, P., Holík, L., Rohlík, T., Hromádka, L., Vranová, V., Rejšek, K., et al. (2011). Methods of plant root exudates analysis: a review. *Acta Univ. Agric. Silv. Mendel. Brun.* 59, 241–246. doi: 10.11118/actaun201159030241
- Ernst, L., Goodger, J. Q. D., Alvarez, S., Marsh, E. L., Berla, B., Lockhart, E., et al. (2010). Sulphate as a xylem-borne chemical signal precedes the expression of ABA biosynthetic genes in maize roots. *J. Exp. Bot.* 61, 3395–3405. doi: 10.1093/jxb/erq160
- Fischer, H., Eckhardt, K.-U., Meyer, A., Neumann, G., Leinweber, P., Fischer, K., et al. (2010). Rhizodeposition of maize: short-term carbon budget and composition. *J. Plant Nutr. Soil Sci.* 173, 67–79. doi: 10.1002/jpln.200800293
- Haase, S., Neumann, G., Kania, A., Kuzyakov, Y., Römhild, V., and Kandeler, E. (2007). Elevation of atmospheric CO₂ and N-nutritional status modify nodulation, nodule-carbon supply, and root exudation of *Phaseolus vulgaris* L. *Soil Biol. Biochem.* 39, 2208–2221. doi: 10.1016/j.soilbio.2007.03.014
- Hund, A., Trachsel, S., and Stamp, P. (2009). Growth of axile and lateral roots of maize: I development of a phenotyping platform. *Plant Soil* 325, 335–349. doi: 10.1007/s11104-009-9984-2
- Iijima, M., Griffiths, B., and Bengough, A. G. (2000). Sloughing of cap cells and carbon exudation from maize seedling roots in compacted sand: sloughing of root cap cells and carbon exudation. *New Phytol.* 145, 477–482. doi: 10.1046/j.1469-8137.2000.00595.x
- Jakab, G., Cottier, V., Toquin, V., Rigoli, G., Zimmerli, L., Métraux, J.-P., et al. (2001). β -aminobutyric acid-induced resistance in plants. *Eur. J. Plant Pathol.* 107, 29–37.
- Jones, D. L., Owen, A. G., and Farrar, J. F. (2002). Simple method to enable the high resolution determination of total free amino acids in soil solutions and soil extracts. *Soil Biol. Biochem.* 34, 1893–1902. doi: 10.1016/S0038-0717(02)00203-1
- Kim, H. K., Choi, Y. H., and Verpoorte, R. (2010). NMR-based metabolomic analysis of plants. *Nat. Protoc.* 5, 536–549. doi: 10.1038/nprot.2009.237
- Koul, V., Adholeya, A., and Kochar, M. (2015). Sphere of influence of indole acetic acid and nitric oxide in bacteria. *J. Basic Microbiol.* 55, 543–553. doi: 10.1002/jobm.201400224
- Krafczyk, I., Trollenier, G., and Beringer, H. (1984). Soluble root exudates of maize: influence of potassium supply and rhizosphere microorganisms. *Soil Biol. Biochem.* 16, 315–322. doi: 10.1016/0038-0717(84)90025-7
- Kumar, D., Shivay, Y. S., Dhar, S., Kumar, C., and Prasad, R. (2012). Rhizospheric flora and the influence of agronomic practices on them: a review. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* 83, 1–14. doi: 10.1007/s40011-012-0059-4
- Marschner, H., Römhild, V., and Ossenberg-Neuhaus, H. (1982). Rapid method for measuring changes in pH and reducing processes along roots of intact plants. *Z. Pflanzenphysiol.* 105, 407–416. doi: 10.1016/S0044-328X(82)80038-X
- Marti, G., Erb, M., Boccard, J., Glauser, G., Doyen, G. R., Villard, N., et al. (2013). Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. *Plant Cell Environ.* 36, 621–639. doi: 10.1111/pce.12002
- Mench, M., and Martin, E. (1991). Mobilization of cadmium and other metals from two soils by root exudates of *Zea mays* L., *Nicotiana tabacum* L. and *Nicotiana rustica* L. *Plant Soil* 132, 187–196. doi: 10.1007/BF00010399
- Micallef, S. A., Shiaris, M. P., and Colón-Carmona, A. (2009). Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *J. Exp. Bot.* 60, 1729–1742. doi: 10.1093/jxb/erp053
- Moe, L. A. (2013). Amino acids in the rhizosphere: from plants to microbes. *Am. J. Bot.* 100, 1692–1705. doi: 10.3732/ajb.1300033

- Naveed, M., Mitter, B., Reichenauer, T. G., Wieczorek, K., and Sessitsch, A. (2014). Increased drought stress resilience of maize through endophytic colonization by *Burkholderia phytofirmans* PsJN and *Enterobacter* sp. FD17. *Environ. Exp. Bot.* 97, 30–39. doi: 10.1016/j.envexpbot.2013.09.014
- Neumann, G., George, T. S., and Plassard, C. (2009). Strategies and methods for studying the rhizosphere—the plant science toolbox. *Plant Soil* 321, 431–456. doi: 10.1007/s11104-009-9953-9
- Oburger, E., Dell'mour, M., Hann, S., Wieshammer, G., Puschenreiter, M., and Wenzel, W. W. (2013). Evaluation of a novel tool for sampling root exudates from soil-grown plants compared to conventional techniques. *Environ. Exp. Bot.* 87, 235–247. doi: 10.1016/j.envexpbot.2012.11.007
- Oburger, E., and Schmidt, H. (2016). New methods to unravel rhizosphere processes. *Trends Plant Sci.* 21, 243–255. doi: 10.1016/j.tplants.2015.12.005
- Oku, S., Komatsu, A., Tajima, T., Nakashimada, Y., and Kato, J. (2012). Identification of chemotaxis sensory proteins for amino acids in *Pseudomonas fluorescens* Pf0-1 and their involvement in chemotaxis to tomato root exudate and root colonization. *Microbes Environ.* 27, 462–469. doi: 10.1264/jisme.2.ME12005
- Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6548–6553. doi: 10.1073/pnas.1302837110
- Römhelt, V., Müller, C., and Marschner, H. (1984). Localization and capacity of proton pumps in roots of intact sunflower plants. *Plant Physiol.* 76, 603–606. doi: 10.1104/pp.76.3.603
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2008). NIH Image to ImageJ: 25 years of image analysis: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Shi, S., Condrón, L., Larsen, S., Richardson, A. E., Jones, E., Jiao, J., et al. (2011). In situ sampling of low molecular weight organic anions from rhizosphere of radiata pine (*Pinus radiata*) grown in a rhizotron system. *Environ. Exp. Bot.* 70, 131–142. doi: 10.1016/j.envexpbot.2010.08.010
- Slota, M., Maluszynski, M., and Szarejko, I. (2016). An automated, cost-effective and scalable, flood-and-drain based root phenotyping system for cereals. *Plant Methods* 12, 34. doi: 10.1186/s13007-016-0135-5
- Sukumar, P., Legué, V., Vayssières, A., Martin, F., Tuskan, G. A., and Kalluri, U. C. (2013). Involvement of auxin pathways in modulating root architecture during beneficial plant-microorganism interactions. *Plant Cell Environ.* 36, 909–919. doi: 10.1111/pce.12036
- van de Venter, H. A., and Grobbelaar, N. (1985). Influence of sub-optimal imbibition temperatures on seed vigour and respiration in maize (*Zea mays* L.). *S. Afr. J. Plant Soil* 2, 203–206. doi: 10.1080/02571862.1985.10634170
- Van Egeraat, A. W. S. M. (1975). Exudation of ninhydrin-positive compounds by pea-seedling roots: a study of the sites of exudation and of the composition of the exudate. *Plant Soil* 42, 37–47. doi: 10.1007/BF02186972
- Walker, T. S., Bais, H. P., Grotewold, E., and Vivanco, J. M. (2003). Root exudation and rhizosphere biology. *Plant Physiol.* 132, 44–51. doi: 10.1104/pp.102.019661
- Warembourg, F. R., Roumet, C., and Lafont, F. (2003). Differences in rhizosphere carbon-partitioning among plant species of different families. *Plant Soil* 256, 347–357.
- Zhang, N., Yang, D., Wang, D., Miao, Y., Shao, J., Zhou, X., et al. (2015). Whole transcriptomic analysis of the plant-beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm formation regulated by maize root exudates. *BMC Genomics* 16:685. doi: 10.1186/s12864-015-1825-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Doan, Doan, Kangas, Ernest, Tran, Wilson, Holmes, Doyle and Durham Brooks. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



In-field High Throughput Phenotyping and Cotton Plant Growth Analysis Using LiDAR

Shangpeng Sun¹, Changying Li^{1*}, Andrew H. Paterson^{2,3}, Yu Jiang¹, Rui Xu¹, Jon S. Robertson², John L. Snider² and Peng W. Chee²

¹ School of Electrical and Computer Engineering, College of Engineering, University of Georgia, Athens, GA, United States,

² Department of Crop and Soil Sciences, College of Agricultural and Environmental Sciences, University of Georgia, Athens, GA, United States, ³ Department of Genetics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States

OPEN ACCESS

Edited by:

Eetu Puttonen,
National Land Survey of Finland,
Finland

Reviewed by:

Wei Guo,
The University of Tokyo, Japan
Marco Maccaferri,
University of Bologna, Italy

*Correspondence:

Changying Li
cyl@uga.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 30 September 2017

Accepted: 04 January 2018

Published: 22 January 2018

Citation:

Sun S, Li C, Paterson AH, Jiang Y, Xu R, Robertson JS, Snider JL and Chee PW (2018) In-field High Throughput Phenotyping and Cotton Plant Growth Analysis Using LiDAR. *Front. Plant Sci.* 9:16. doi: 10.3389/fpls.2018.00016

Plant breeding programs and a wide range of plant science applications would greatly benefit from the development of in-field high throughput phenotyping technologies. In this study, a terrestrial LiDAR-based high throughput phenotyping system was developed. A 2D LiDAR was applied to scan plants from overhead in the field, and an RTK-GPS was used to provide spatial coordinates. Precise 3D models of scanned plants were reconstructed based on the LiDAR and RTK-GPS data. The ground plane of the 3D model was separated by RANSAC algorithm and a Euclidean clustering algorithm was applied to remove noise generated by weeds. After that, clean 3D surface models of cotton plants were obtained, from which three plot-level morphologic traits including canopy height, projected canopy area, and plant volume were derived. Canopy height ranging from 85th percentile to the maximum height were computed based on the histogram of the z coordinate for all measured points; projected canopy area was derived by projecting all points on a ground plane; and a Trapezoidal rule based algorithm was proposed to estimate plant volume. Results of validation experiments showed good agreement between LiDAR measurements and manual measurements for maximum canopy height, projected canopy area, and plant volume, with R^2 -values of 0.97, 0.97, and 0.98, respectively. The developed system was used to scan the whole field repeatedly over the period from 43 to 109 days after planting. Growth trends and growth rate curves for all three derived morphologic traits were established over the monitoring period for each cultivar. Overall, four different cultivars showed similar growth trends and growth rate patterns. Each cultivar continued to grow until ~88 days after planting, and from then on varied little. However, the actual values were cultivar specific. Correlation analysis between morphologic traits and final yield was conducted over the monitoring period. When considering each cultivar individually, the three traits showed the best correlations with final yield during the period between around 67 and 109 days after planting, with maximum R^2 -values of up to 0.84, 0.88, and 0.85, respectively. The developed system demonstrated relatively high throughput data collection and analysis.

Keywords: field-based high throughput phenotyping, 3D point cloud, morphologic traits, plant growth analysis, LiDAR

INTRODUCTION

The global population is estimated to approach nine billion by 2050, and demand for food and fiber crops is expected to increase by 60% (Tilman et al., 2011; Gerland et al., 2014). New plant breeding approaches need to be developed to overcome these tremendous challenges. An important step in this direction is to gain a better understanding of the relationship between genotype and phenotype (Goggin et al., 2015; Großkinsky et al., 2015; Rahaman et al., 2015). However, in-field high throughput phenotyping technologies, which can facilitate automatic measurement of phenotypic traits over the entire growing season, are still considered to be a major bottleneck limiting crop improvement (Furbank and Tester, 2011; Cobb et al., 2013).

Plant morphologic traits can often be used for evaluating plant growth (Hosoi and Omasa, 2012; Taheriazad et al., 2016), which determines plant performance in terms of final crop biomass and yield (Dhondt et al., 2013). Several studies showed that morphologic traits such as canopy height and leaf area index (LAI) were strongly related to plant species, type of cultivation, plant growth rate, and final yield (Gebbers et al., 2011; Sharma and Ritchie, 2015; Friedli et al., 2016; Sun et al., 2017). Importantly, plant growth and yield is dependent upon leaf area development, the average photosynthetic efficiency of all leaves in the plant canopy (Gardner, 1985; Krieg and Sung, 1986), and partitioning of dry matter to the harvested portion of the crop (Earl and Davis, 2003). Thus, plant canopy development should provide some indication of the crop's capacity for growth and yield.

The traditional manual measurement of plant morphologic traits is time consuming, labor intensive, and sometimes destructive. Novel technologies for plant phenotyping in a non-invasive and high throughput manner with high spatial and temporal resolution offer improved efficiency (Furbank and Tester, 2011; Dhondt et al., 2013; Großkinsky et al., 2015). Over the past decade, several non-invasive approaches using sensing technologies were developed for in-field phenotyping (Lin, 2015; Simko et al., 2016). Computer vision was one commonly-used technology. Usually, plant traits were extracted from color (RGB) images. Li et al. (2016) introduced a method for in-field cotton boll detection based on color and texture features using 2D color images. Si et al. (2015) developed a machine vision system to automatically recognize and locate apples; over 89.5% accuracy was achieved. Such 2D image based methods provide potential to conduct phenotypic measurements with a high spatial resolution, but are limited by plant occlusion. In addition, one major challenge with 2D digital image methods is that image quality is significantly affected by highly variable illumination conditions in the field, which limits automatic data processing (Li et al., 2014).

The use of 3D model-based methods for plant phenotyping are receiving increasing attention, as they permit multiple morphologic traits such as canopy height, plant volume, and LAI

to be simultaneously extracted (Bietresato et al., 2016; Vazquez-Arellano et al., 2016; Gibbs et al., 2017) while mitigating plant occlusion. Moreover, 3D models have the potential to assist growers to continuously monitor and quantify plant growth and development, as well as plant responses to environmental stresses. A stereo-imaging based 3D reconstruction system was established to capture rape seedling leaf area and plant height (Xiong et al., 2017); two identical RGB cameras were utilized as an imaging unit. The mean error for leaf area and plant height measurements was 3.68 and 6.18%, respectively. The system was put in a well-designed box, in which homogenous illumination was provided. Thuy Tuong et al. (2015) developed a 3D reconstruction system based on 10 digital color cameras that were mounted on a custom structure, and an illumination system was used to enhance the visual texture of plants from all camera viewpoints. The system produced very high quality, dense, and complete point clouds. However, as both systems were designed for indoor use, they would need to be modified for field applications under natural illumination. 3D models can also be produced by time of flight (TOF) cameras; however, similarly to RGB image based methods, data quality would be significantly affected by sunlight under field conditions, which limits in-field applications. In Busemeyer et al. (2013) and Jiang et al. (2016) the TOF cameras were mounted inside an enclosure in order to mitigate the influence of sunlight.

Light detection and ranging (LiDAR) technology provides an alternative approach for 3D plant model reconstruction. LiDAR is a remote sensing technology to measure the distance between the sensor and an object of interest by illuminating the object with a laser and analyzing the TOF. LiDAR may be the best known and most widely used sensor for 3D canopy reconstruction (Deery et al., 2014; Gibbs et al., 2017). A 2D LiDAR collects two dimensional scans in a measured plane, and a 3D model can be obtained by moving the sensor along the perpendicular direction to the scanning plane. Although the spatial resolution of the 3D model produced by LiDAR is not as dense as those obtained by camera-based methods, it is sufficient for the extraction of most plant morphologic traits (Rosell-Polo et al., 2009; Bietresato et al., 2016; Sun et al., 2017). In Deery et al. (2014), a LiDAR (LMS400, SICK AG, Waldkirch, Germany) with a monochromatic red laser light source was used to generate intensity images of several crops including rice, wheat, and maize. It was concluded that LiDAR is a potential alternative to image-based methods for phenotyping morphologic traits at the plot or plant level under field conditions. Moreover, in contrast to image based methods, the LiDAR based method uses its own light source, mitigating problems with highly variable illumination conditions in the field. In addition, LiDAR can be used with a high scanning frequency and a large scanning range (Lin, 2015). Therefore, LiDAR has excellent potential for in-field plant phenotyping. 3D point clouds can also be generated by some other sensors such as triangulation line scanner (Paulus et al., 2014b) and ultrasonic sensing (Llorens et al., 2011). A very dense 3D model can be reconstructed using a triangulation line scanner and morphologic traits at the organ level could be extracted, however the relatively short working distance limits its application for large plants such as cotton.

Abbreviations: CH, canopy height; PCA, projected canopy area; PV, plant volume; 3PLM, three-parameter logistic model; CI, confidence interval.

Growth dynamics of plant morphologic traits provide important information toward determining plant productivity. Tessmer et al. (2013) described a high-throughput phenotyping platform for plant growth modeling and functional analysis (HPGA). Plant growth curves were generated by the platform, which were used to gain a deeper understanding of energy distribution. Friedli et al. (2016) introduced a terrestrial 3D laser scanner-based plant growth monitoring system which could be used for monitoring canopy height growth for different crops under field conditions. Paulus et al. (2014b) monitored the organ-specific growth dynamics of cereal plants with a high precision triangulation line laser scanner by scanning every 2–3 days. Three morphologic traits including leaf area, stem height, and plant volume were measured, allowing quantification of the growth dynamics of the barley plant.

In the present study, three morphologic traits—canopy height, projected canopy area, and plant volume—of cotton plants were derived based on data collected by a 2D LiDAR. Morphologic trait data was collected and growth analysis was conducted. Several improvements were made over previously mentioned studies. A 1-cm accuracy level RTK-GPS was used in this study in order to provide accurate spatial coordinates for LiDAR scans so that a precise 3D surface model could be reconstructed, which is important because the model is the basic dataset for further analysis. In addition, analyses of growth dynamics and correlation of morphologic traits with final yield were conducted over the growing season. The system repeatedly scanned plants over the growing season, permitting analysis of the effects of different cultivars on the growth and final yield of cotton plants.

The overall goal of this work was to develop a high throughput phenotyping system for morphologic traits of cotton plants using LiDAR under field conditions. This LiDAR-based system is one component of our broader effort to develop field-based high throughput phenotyping (HTP) systems, and it complements other image-based sensors such as Kinect V2, thermal camera, and hyperspectral camera. This LiDAR-based system provides accurate morphologic traits in a robust and fast way, and saves storage space and computing resources compared to image-based sensors. Specific objectives were to: (1) develop algorithms to extract multiple morphologic traits—canopy height, projected canopy area, and plant volume—from a 3D point cloud obtained with a 2D LiDAR; (2) conduct 4D monitoring (3D plant reconstructions over time) of the derived plant morphologic traits to detect growth patterns of plants from different cultivars; and (3) explore relationships between derived morphologic traits and final yield.

MATERIALS AND METHODS

Experimental Field

The study site was located at the Iron Horse Farm (IHF) in Greene County, GA, USA. The entire study included 128 plots arranged in 16 rows and 8 columns (**Figure 1A**), using a randomized complete block design with four cultivars of cotton and 32 replicate plots per cultivar. Four plots of each cultivar were planted in each column. The distribution of cultivars in each column was randomly assigned (**Figure 1B**). Plots were 3.05 m

wide. A total of 15 seeds were sowed in each plot at spacing of 0.15 m. Inter-row spacing was 1.52 m, and inter-column spacing was 1.83 m. Cotton seeds were sowed on June 13, 2016.

The four cotton cultivars were GA2011158 (cultivar 1), GA2009037 (cultivar 2), GA2010074 (cultivar 3), and UA48 (cultivar 4) which is commercialized by the private seed company Americot. All four cultivars were developed by conventional breeding possessing no transgenic insects or herbicides tolerant traits. However, they have different fiber quality, growth habits, and plant architecture due to adaptation to different production regions. Cultivars 1, 2, and 3 are elite breeding lines developed by the University of Georgia cotton breeding program for adaptation to the southeastern cotton production region, bred to have indeterminate growth habit to take advantage of the long growing season in the southern US cotton belt. Plants from these cultivars will continue adding vegetative growth at the same time as the reproductive development, therefore they can grow excessively tall and rank in high nitrogen environment or if there are severe insects damage causing excessive square loss. Cultivar 4, on the other hand, was released by the University of Arkansas cotton breeding program, bred to have an early maturity growth habit for adaptation to the northern region of the US cotton belt. It has a determinate growth habit which resulting in shorter statured, extended sympodial branches, and shorter flowering date.

Data Acquisition

The data collection system mainly consisted of a LiDAR (LMS 511 PRO SR, SICK AG, Waldkrich, Germany) (**Figure 2A**), an RTK-GPS (Cruizer II, Raven Industries Inc., Sioux Falls, SD, USA) (**Figure 2B**), and a rugged laptop as a DAQ (data acquisition) and storage device. The LiDAR was mounted on a tractor (Spider DL, LeeAgra, Inc., Lubbock, TX, USA) platform at a height of 2.4 m to scan cotton plots from directly above (**Figure 2C**). The RTK-GPS was mounted on the roof of the tractor, which provided the spatial coordinates during data collection. The LiDAR and RTK-GPS receiver were aligned to the center of the tractor (**Figure 2D**).

The LiDAR was developed for outdoor use, and measured in 2D radial coordinates from -5 to 185° with a maximum range of 80 m. Line scans could be acquired at a rate of 25–100 Hz with an angular resolution of $0.1667-1^\circ$. The built-in filters eliminated interference from particles of dust, raindrops, and snowflakes. An enclosure was used to provide a controlled environment for data acquisition. The RTK-GPS provided coordinates with 1 cm accuracy with an update rate up to 10 Hz.

Data collections were conducted in the field from July 26 to September 30, 2016, i.e., from 43 to 109 days after planting (DAP) (**Table 1**). When scanning plants, the angular resolution of the LiDAR was configured to be 0.33° with a sampling frequency of 50 Hz, and the echo filter and particle filter were enabled. The update frequency of the RTK-GPS was 5 Hz. The tractor scanned the field row by row. For each row, the tractor traveled from column 1 to column 8 with an average speed of about 0.5 m/s. Since the first two sample dates—July 26 and 28—were close to each other, the data for July 26 were not presented except for growth rate analysis.

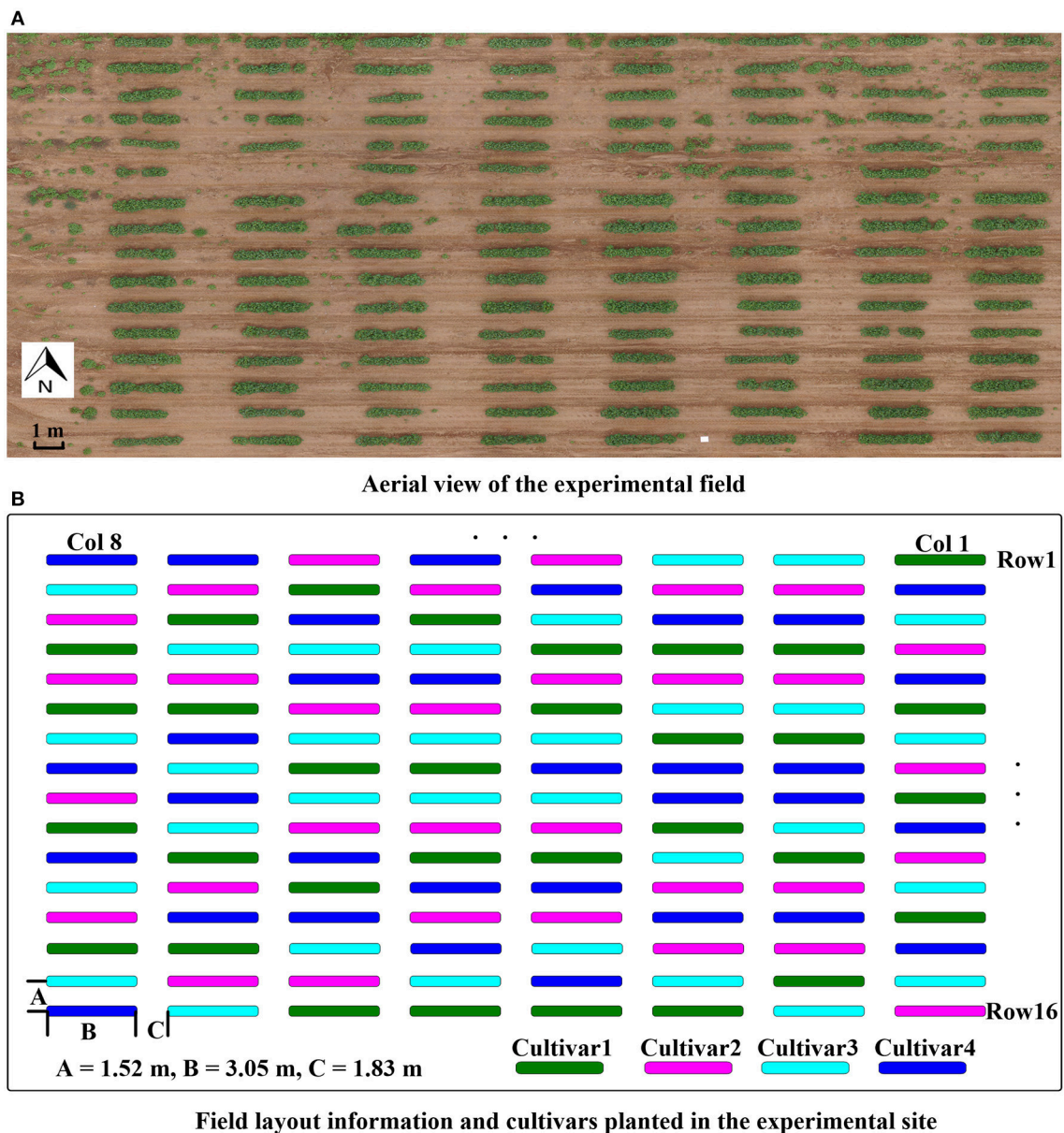


FIGURE 1 | Experimental field layout. **(A)** Aerial view of the experimental field; **(B)** illustration of experimental design with cultivar and field layout information.

Data Processing Algorithms

After raw data was collected in the field, further processing and analysis were performed in the lab. The data processing and analysis program was developed and implemented in MATLAB 2016b (The Math Works Inc., Natic, MA, USA) on a desktop equipped with an intel i7-6700 CPU 3.40 GHz with 16 GB RAM, running on a Windows 10 Enterprise operating system.

Two steps were executed to derive plant features (**Figure 3**): 3D plant surface model generation (section Generation of 3D Model) and morphologic plant parameter extraction (section Extraction of Morphologic Traits).

Generation of 3D Model

The 3D model for each row was reconstructed based on GPS and LiDAR data. The GPS and LiDAR dataset was depicted by Equation (1). The two kinds of data were synchronized using timestamps.

$$\begin{aligned} P_{GPS} &= \{P_{GPS0}, P_{GPS1}, \dots, P_{GPS(N-1)}\} \\ F_{LiDAR} &= \{F_{LiDAR0}, F_{LiDAR1}, \dots, F_{LiDAR(M-1)}\} \end{aligned} \quad (1)$$

P_{GPS} was the set of collected GPS data, and F_{LiDAR} was the set of scanned frames of LiDAR. The number of GPS points was N , and the number of LiDAR frames was M .

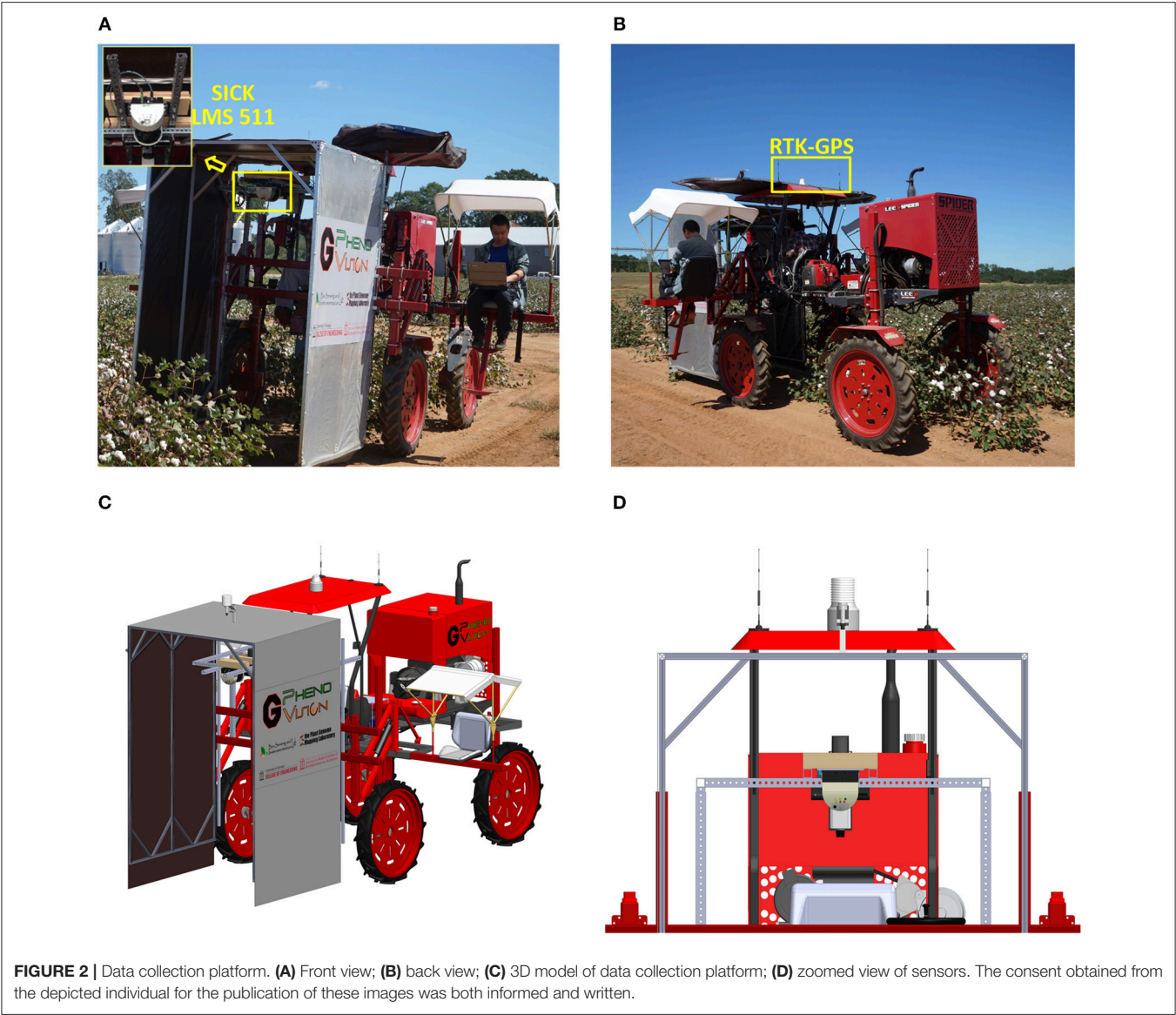


TABLE 1 | Summary of data collection dates (Year: 2016).

Period	P1	P2	P3	P4	P5	P6	P7	P8
Date	26 July–28 July	28 July –04 Aug	04 Aug–19 Aug	19 Aug–26 Aug	26 Aug–09 Sep	09 Sep–16 Sep	16 Sep–23 Sep	23 Sep–30 Sep
DAP	43–45	45–52	52–67	67–74	74–88	88–95	95–102	102–109

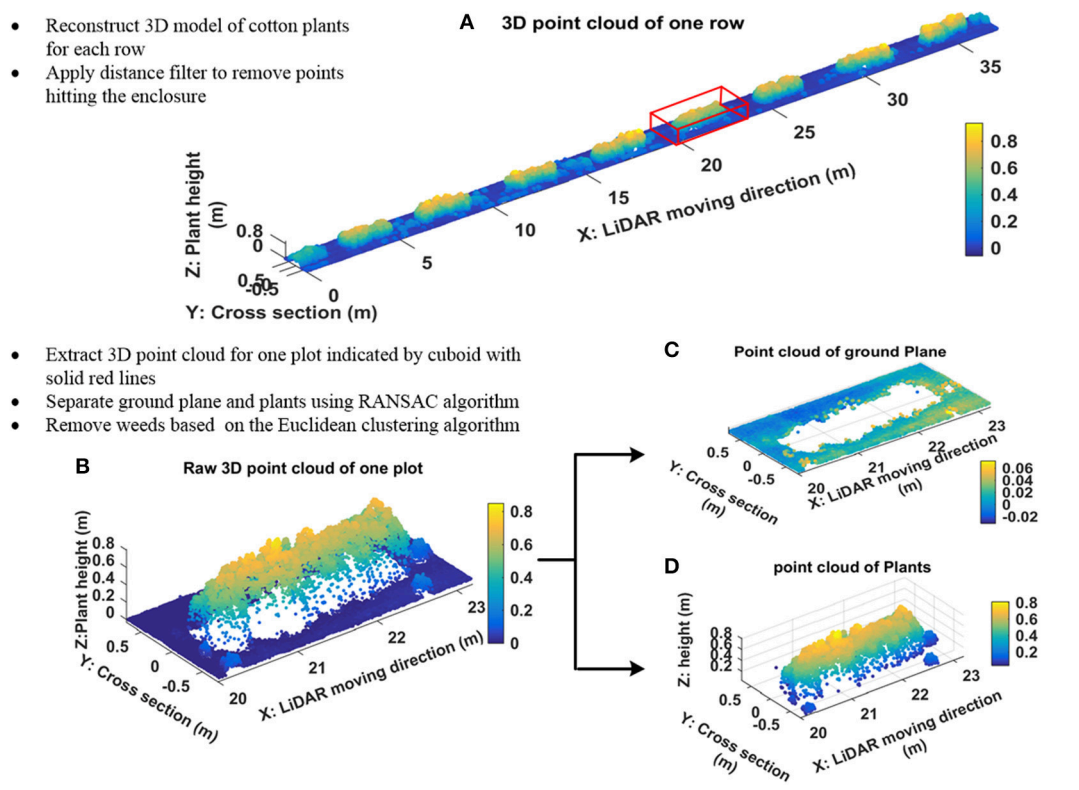
The distance between two adjacent GPS points, denoted by ΔP_{GPS} , was computed by Equation (2). f_{LiDAR} and f_{GPS} were the data acquisition frequency of LiDAR and GPS, respectively. In this study, the data acquisition frequency of GPS was $f_{GPS} = 5\text{ Hz}$, and LiDAR scanning frequency was $f_{LiDAR} = 50\text{ Hz}$. Therefore, there were 10 scanned frames, each containing 571 points (the aperture angle was 190° with angular resolution 0.33°) between every two adjacent GPS points (Equation 3). Assuming that the tractor was moving at a constant speed during the interval (200 ms) of two adjacent GPS points, the distance

of the two adjacent frames within two adjacent GPS points was computed using Equation (4). Therefore, the position of each LiDAR scanned frame was obtained using Equation (5). D_{offset} was the offset between LiDAR and GPS. In this study, D_{offset} was fixed during data collection in the field, and the measured point at 0° scanning angle was used to depict the frame position.

$$\begin{cases} \Delta \vec{P}_{GPS}(i) = P_{GPS}(i) - P_{GPS}(i - 1), & i = 1, 2, \dots, N - 1 \\ \Delta \vec{P}_{GPS}(0) = \vec{0}, & i = 0 \end{cases} \quad (2)$$

Step1: Reconstruct 3D model based on LiDAR and GPS data

- Reconstruct 3D model of cotton plants for each row
- Apply distance filter to remove points hitting the enclosure



Step2: Derive morphologic traits based on the clean point cloud of plants

- Canopy height (CH)
- Canopy size and projected canopy area (PCA)
- Plant volume (PV)

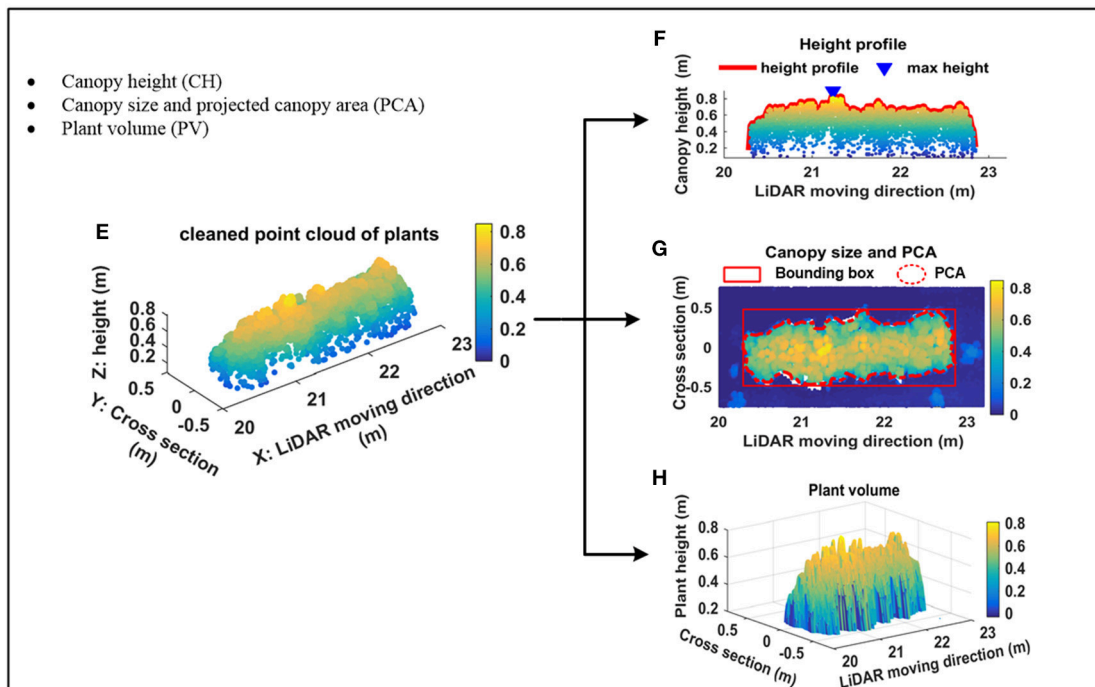


FIGURE 3 | Data processing pipeline. **(A)** Example of 3D point cloud of one scanned row; **(B)** segmented 3D point cloud of a plot indicated by cuboid with solid red lines in **(A)**; **(C)** point cloud of ground plane; **(D)** point cloud of plants; **(E)** denoised point cloud of plants; **(F)** height profile of plants and maximum canopy height; **(G)** projected canopy area; **(H)** plant volume.

$$\alpha = \frac{f_{LiDAR}}{f_{GPS}} \quad (3)$$

$$\vec{d}_i = \frac{\Delta \vec{P}_{GPSi}}{\alpha}, i = 0, 1, 2, \dots, N-1 \quad (4)$$

$$f_{LiDAR}(k) = f_{LiDAR}(i\alpha + j) = p_{GPS}(i) + j \times \vec{d}_i + \vec{D}_{offset} \quad (5)$$

$$i = 0, 1, 2, \dots, N-1, j = 0, 1, 2, \dots, 9, k = 0, 1, \dots, M-1$$

More details related to this processing can be found in Sun et al. (2017). **Figure 3A** shows an example of the reconstructed 3D model of one scanned row. A distance filter was used to remove points hitting the enclosure. A plot level 3D model was extracted from the 3D model of the row according to the proportions of the field layout (**Figure 3B**). The standard RANSAC algorithm was applied to cut-off points of the ground plane (**Figures 3C,D**). An Euclidean clustering algorithm (Rusu et al., 2008) was used to remove points generated by weeds that were not attached to cotton plants (**Figure 3E**).

Extraction of Morphologic Traits

The maximum canopy height (CH) was measured by calculating the distance from the ground plane to the apex of all measured points (**Figure 3F**). In addition, different percentiles canopy height—from 85th to maximum CH with steps of 3%—were calculated based on the histogram of the z coordinate for all measured points. The boundary points of the plant canopy were detected by projecting all points onto the ground plane, and the projected canopy area (PCA) and bounding box of canopy structure (representing maximum length and width occupied by the canopy) were extracted from boundary points (**Figure 3G**). A Trapezoidal rule based algorithm was used to calculate plant volume (PV) (**Figure 3H**) in order to provide an indication of the 3D space occupied by each plot.

S denoted the line scan set for a plot (Equation 6) which contained k line scans. For a scan s_i which contained n measured points, the dashed line was the surface profile and the red spots were measured points by LiDAR (**Figure 4**). The area denoted as A_i under the line scan s_i could be estimated using the measured points (red spots in **Figure 4**) based on Trapezoidal rule (Equation 7), where (x_i, y_i) were the coordinates of the i th measured point.

$$S = \{s_1, s_2, \dots, s_k\} \quad (6)$$

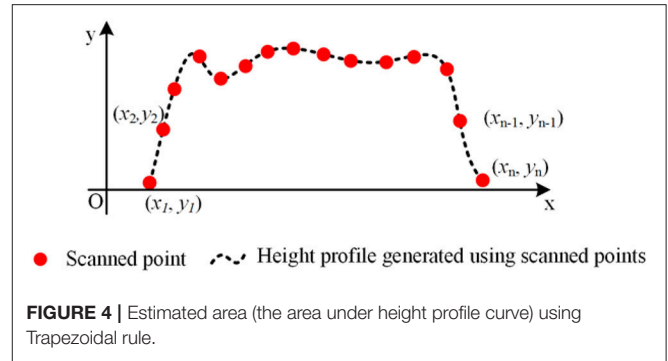
$$A_i = \sum_{j=1}^{n-1} (x_{j+1} - x_j) \left(\frac{y_{j+1} + y_j}{2} \right) \quad (7)$$

Therefore, the area set denoted as A for the plot was obtained using Equation (8).

$$A = \{A_1, A_2, \dots, A_k\} \quad (8)$$

Similar to the area calculation process, PV was obtained with Equation (9).

$$PV = \sum_{i=1}^{k-1} (l_{i+1} - l_i) \left(\frac{A_{i+1} + A_i}{2} \right) \quad (9)$$



l_i was the position along the tractor moving direction of the i th line scan.

Validation Experiments

To verify the accuracy of canopy height measurements using point cloud data, 96 samples of maximum CH—the perpendicular distance from the highest point to the ground plane—were manually measured using a tape measure during data collection in the field. The samples were measured on four different days: July 28, August 04, August 26, and September 09, 2016. On those days, the average wind speeds were around 2.7, 3.4, 2.3, and 2.5 m/s, respectively.

For PCA validation experiments, eight “dummy plants” were made with different canopy shapes using printed paper leaves and metal wires. The dummy leaves were in three different sizes—10, 41, and 92 cm²—and there were six of each size. The leaves of two dummy plants were arranged to be overlapped. Each dummy plant was scanned by LiDAR and imaged by a DSLR camera (FUJIFILM X-A10 mirrorless camera, FUJIFILM, Tokyo, Japan). A scale bar was used to calibrate the real size of the plant. The ground truth of PCA was computed by segmenting leaves based on color information.

For PV validation experiments, eight plants at different growth stages were used, among which five were real cotton plants, one was a shrub plant, and two were dummy plants. The six real plants were in the leaf and canopy development growth stage, whereas the two dummy plants were used to mimic flowering and boll development growth stages. Each plant canopy was divided into 50 mm cylindrical discs from the top to the bottom and the diameter of each disc was manually measured using a tape measure. The diameter and height of each cylindrical disc were used to estimate the volume of each disc, and then the volume of the whole plant was obtained by summing the volumes of each disc. Manual measures of the three traits were plotted vs. point cloud estimates, and regression analysis was used to compute the root mean square error (RMSE) and coefficient of determination (R^2). More information about PCA and PV validation experiments can be found in supplementary materials.

Plant Growth Analysis

Plant growth dynamics were obtained by monitoring plants over the growing season when plant morphologic traits were

extracted. Growth trends and growth rates for each morphologic trait noted above (CH, PCA, PV) were computed and analyzed over the monitoring period. Growth trends were determined as the variation in measured traits over the monitoring period. The results were given as mean values and standard deviations of all morphologic parameters. A three-parameter logistic model (3PLM) was used to fit growth curves of the three detected traits for each cultivar (Tessmer et al., 2013). The model is a function of time t as shown in Equation (10).

$$y(t) = \frac{x_0 x_n}{x_0 + (x_n - x_0)e^{-\tau(t-T)}} \quad (10)$$

where t was the time which was denoted by days after planting in this study, τ was a coefficient, $y(t)$ was detected traits at time t , x_0 , x_n were the initial value and the upper horizontal asymptote of the detected trait, respectively, and T was the first day of data collection which was also denoted by days after planting (in this study $T = 45$). x_0 , x_n , and τ can be estimated using

non-linear least squares based on the observations of the detected traits.

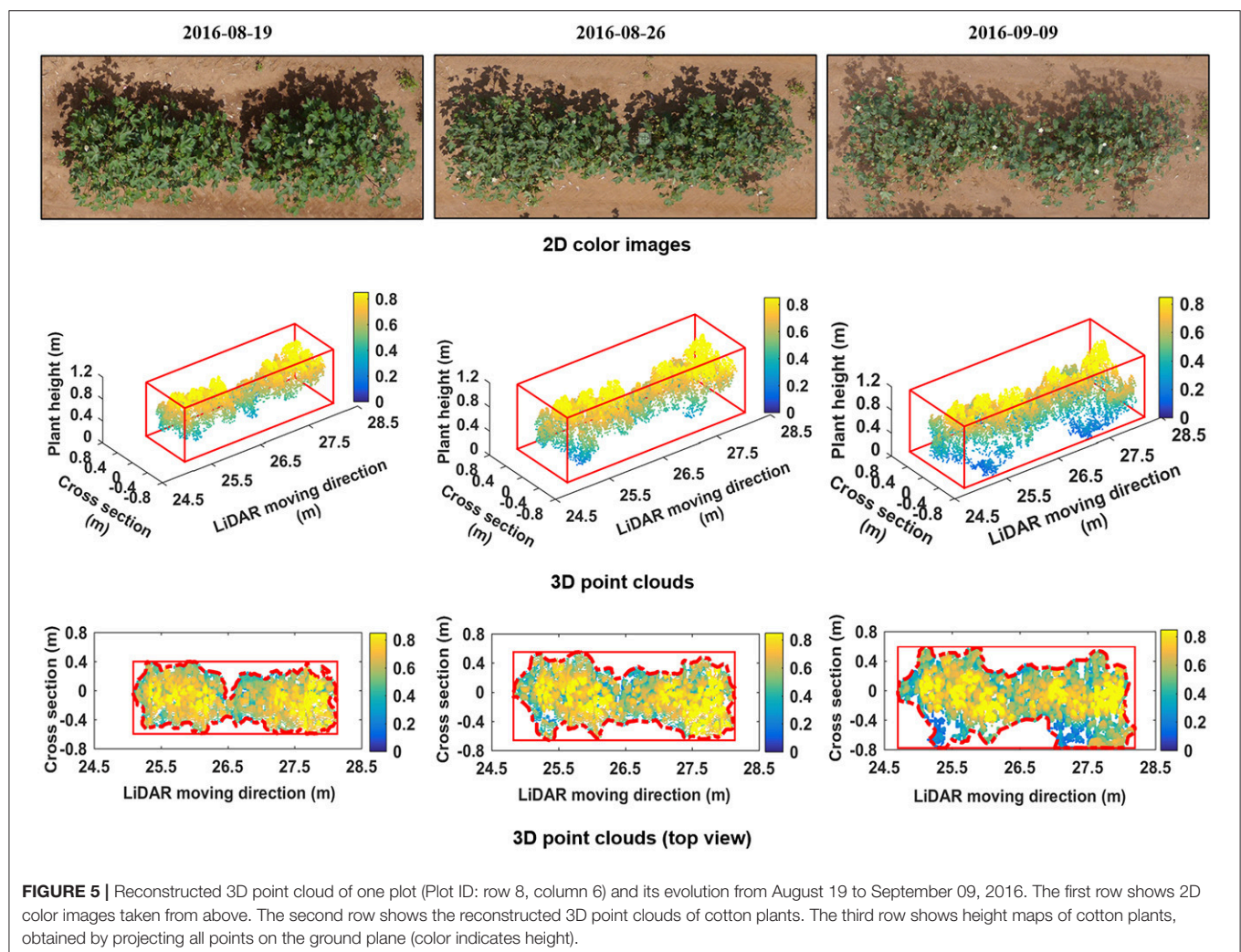
Growth rate was determined as the average change in measured traits over a time interval. Growth rate was calculated by Equation (11).

$$GR = \frac{P_t - P_{t-\Delta t}}{\Delta t} \quad (11)$$

where GR was the growth rate, Δt was the time interval, and P_t was the measured plant trait at time t .

Correlation Analysis between Morphologic Traits and Yield

Seed cotton (mature fiber plus seeds to which it was attached) was harvested manually on November 4, 2016, and yield was expressed as g/plot. In order to explore the relationship between derived morphologic traits and final yields, linear regression analysis was conducted for each cultivar over the monitoring period. The coefficient of determination (R^2) was computed.



RESULTS

Morphologic Traits Extraction and Validation

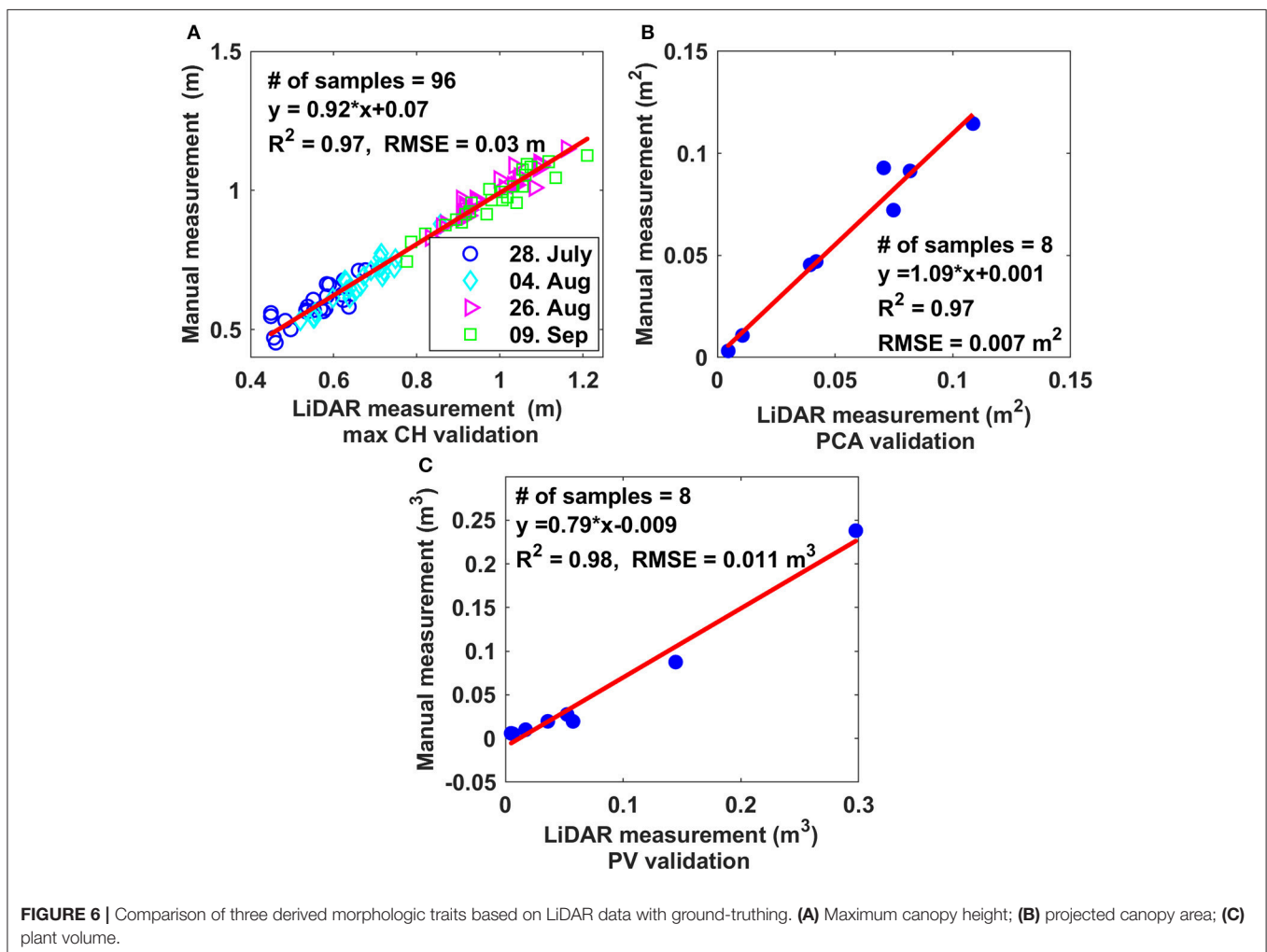
Morphologic trait information of cotton plants was derived from the reconstructed 3D point clouds. **Figure 5** shows the 3D point clouds over time compared to the 2D color images of the same plot. The reconstructed 3D point clouds on August 19, August 26, and September 9, respectively, were obtained using 207, 256, and 263 scans and contained 11,403, 14,379, and 14,485 points. The cuboids indicated by solid red lines for graphs in row 2 of **Figure 5** are 3D bounding boxes and points within the cuboids belong to plants. The rectangles indicated by red solid lines for graphs in row 3 of **Figure 5** are 2D bounding boxes, and the dashed red lines are the detected boundaries from which PCA was derived. Morphologic traits including CH, PCA, and PV were extracted from the 3D point clouds.

CH-, PCA-, and PV-values derived from LiDAR data were highly correlated with manually measured values, with R^2 -values of 0.97, 0.97, and 0.98 and RMSE values of 0.03 m, 0.007 m², and 0.011 m³, respectively (**Figure 6**).

Figure 7A shows the side view of the point cloud of the plot presented in **Figure 5**. The data was collected on August 19 and consisted of 207 line scans. The number of measured points for each line scan varied depending on the width of plants. A total of 66 points were measured for the 144th scan (**Figure 7B**). Points near the center were denser than points located at two terminal areas because the inter-distance between two adjacent points increased with increasing distance between the LiDAR and measured points. A minimum threshold of 0.05 m was applied to remove points. Based on the proposed PV calculation method, the area under the profile and PV were estimated: $A_{144} = 0.52$ m², $PV = 2.33$ m³ (**Figure 7C**).

Plant Growth Analysis

Overall, the three measured morphologic traits—CH, PCA and PV—showed similar growth trends over the monitoring period based on the measurements from the sensor, but the actual value was cultivar specific for each trait (**Figure 8**). For maximum CH, all cultivars reached the maximum height on around day 88. Cultivar 1 and 2 had similar average maximum heights of 1.08 m, while cultivars 3 and 4 reached the peak values of 0.96 and 0.88 m,

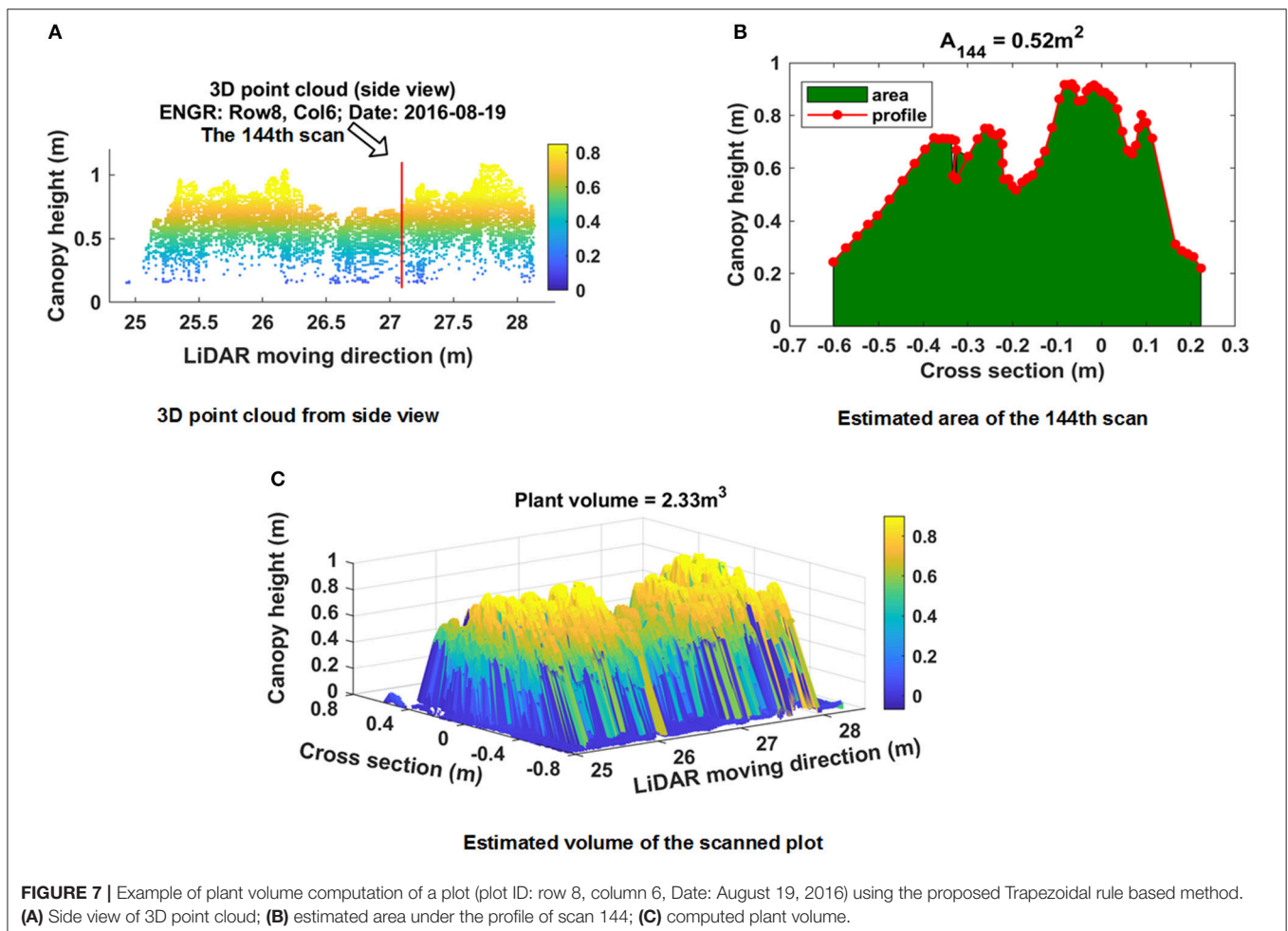


respectively. Cultivar 1 and 2 were around 22.7% higher than cultivar 4. PCA continued to increase until around 95 DAP, which was 7 days longer compared to maximum CH. The maximum PCAs of cultivars 1, 2, 3, and 4 were 2.73, 2.23, 2.47, and 2.34 m², respectively. Cultivar 1 had much larger PCA-value than the other three cultivars which had similar PCA. It is around 22.4% larger than cultivar 2 which had the minimum area. Cultivar 2 showed the highest CH, but lowest PCA over the whole monitoring period due to less horizontal canopy expansion in the horizontal direction. For the growth curve of PV, the maximum points for all four cultivars were reached on day 88, they were 2.75, 2.17, 2.11, and 1.59 m³ for cultivar 1, 2, 3, and 4. Larger differences among cultivars were observed for PV than PCA. The maximum volume was around 73.0% larger than the minimum one. In summary, cultivar 1 had the largest canopy (highest PV, CH, and PCA) while cultivar 4 had the lowest maximum CH and PV. Although cultivar 2 had high maximum CH, its PCA was the lowest, making its PV in the middle range among the cultivars. The derived values and variation of morphologic traits over time were in good agreement with the developmental phases of cotton (Ritchie et al., 2007).

Growth curves generated for each cultivar of all three detected traits showed good correlation with the sensor measurements.

The date when the traits reached their upper horizontal asymptotes and the values of their upper horizontal asymptotes were similar to the results obtained from measured data. The growth curve varied little after reaching the upper horizontal asymptotes, while the sensor measurements showed a decreasing trend which was mainly due to defoliation. Supplementary Table 1 presented estimated parameters of the 3 PLM and their 95% confidence interval (CI).

Figure 9 shows the growth rates of maximum CH, PCA, and PV at different periods of time. For maximum CH, cultivars 1 and 2 had similar growth rates, and grew faster than cultivars 3 and 4 during the period from P1 to P5. The peak of GR of cultivar 1 and 2 was observed during P3 (between 52 and 67 DAP), which was around 0.018 m per day (m/d); However, it was during P2 for cultivar 3 and 4, which was around 0.015 m/d (**Figure 9A**). For PCA, all four cultivars had similar GR during P2 and P3; However, large differences emerged during the period from P3 to P6. The peak of GR for cultivar 1 and 3 was in P4, with the values of 0.065 m²/d; while it was in P2 for cultivar 2 and 4, with the same value of 0.054 m²/d. PCA increased until P6 (95 DAP) and then started to decrease (**Figure 9B**), 21 days later than the period when the maximum CH started to decrease. During the PCA increasing period, cultivar 1 grew faster than the other three.



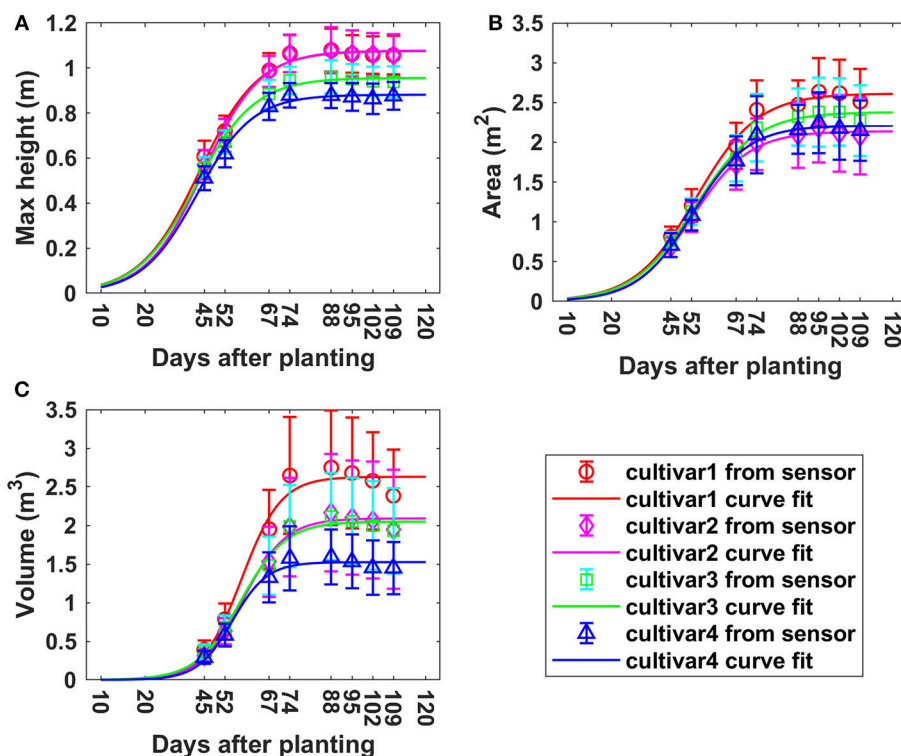


FIGURE 8 | Growth curves for derived morphologic parameters over the monitoring period. **(A)** Maximum canopy height; **(B)** projected canopy area; **(C)** plant volume.

PV had a faster GR than the other three cultivars during P1 to P4—it reached the peak in P4 with the value of around $0.1 \text{ m}^3/\text{d}$. Cultivar 2, 3, and 4 reached the peak of GR in P4 ($0.06 \text{ m}^3/\text{d}$), P4 ($0.07 \text{ m}^3/\text{d}$), and P3 ($0.05 \text{ m}^3/\text{d}$), respectively. In summary, the fastest GRs for all three traits were during P2 to P4, that was from 45 to 74 DAP, which was also indicated by the growth curves in **Figure 8**.

ANOVA-tests showed that cultivars had a significant influence on the derived morphologic traits over the monitoring period (**Table 2**). This allowed us to utilize regression analysis to determine which morphologic traits and measurement times were most closely associated with final yield for each cultivar.

Relationship between Morphologic Traits and Yield

A significant difference in the final yield was observed between the four cultivars (**Table 3**). Cultivar 2 produced significantly lower yields than cultivar 3 and 4.

The relationship between CH and final yield showed similar trends using data from various CH percentiles, especially between the 85th and 94th percentiles. The R^2 -values for cultivar 1, 3, and 4 increased over the monitoring period although there existed only slight variation, whereas cultivar 2 exhibited a decreasing trend (**Figure 10**). Among the cultivars studied, cultivar 4 had the highest R^2 -values (up to 0.84) from day 67 on, based on the 85th CH to maximum CH. The variation curve of cultivar 1 was analogous to the curve for cultivar 3, both reaching the

highest R^2 -values between 88 and 95 DAP. In contrast, cultivar 2 demonstrated the highest correlation with yield at an early growth stage around 52 DAP. Based on the results from this study, percentiles from 85 to 94% of CH during the period from 67 to 109 DAP are recommended for yield estimation application due to not only high but also stable R^2 -values for all four cultivars.

Overall, the R^2 -values between PCA and final yield increased over the monitoring period for all four cultivars (**Figure 11A**). The R^2 -values for cultivars 2, 3, and 4 on each data collection day were similar, while cultivar 1 exhibited lower R^2 -values. Compared to CH parameters, a major difference was that cultivar 2 had an opposite variation trend. The highest R^2 -values for all four cultivars were reached between 88 and 109 DAP, and they were 0.65, 0.83, 0.87, and 0.88 for cultivars 1, 2, 3, and 4, respectively. This indicated that PCA was more closely related to final yield than CH. For both CH and PCA, cultivar 4 showed the strongest correlation with final yield among the four cultivars.

For PV, the difference in R^2 -values between the four cultivars became smaller over the monitoring period compared to CH and PCA, especially between 88 and 102 DAP (**Figure 11B**). The maximum R^2 -values were 0.77, 0.85, 0.84, and 0.83 on 95, 67, 74, and 74 DAP for cultivars 1, 2, 3, and 4, respectively, which were similar to the values for PCA, and better than CH. This indicated that PV was a more stable trait than CH and PCA and could feasibly be used to predict cotton final yield.

PCA and PV had higher R^2 -values than CH over the monitoring period (**Figure 12**) when combining all cultivars. The R^2 -values increased steadily to a final value of 0.72 for PCA. PV

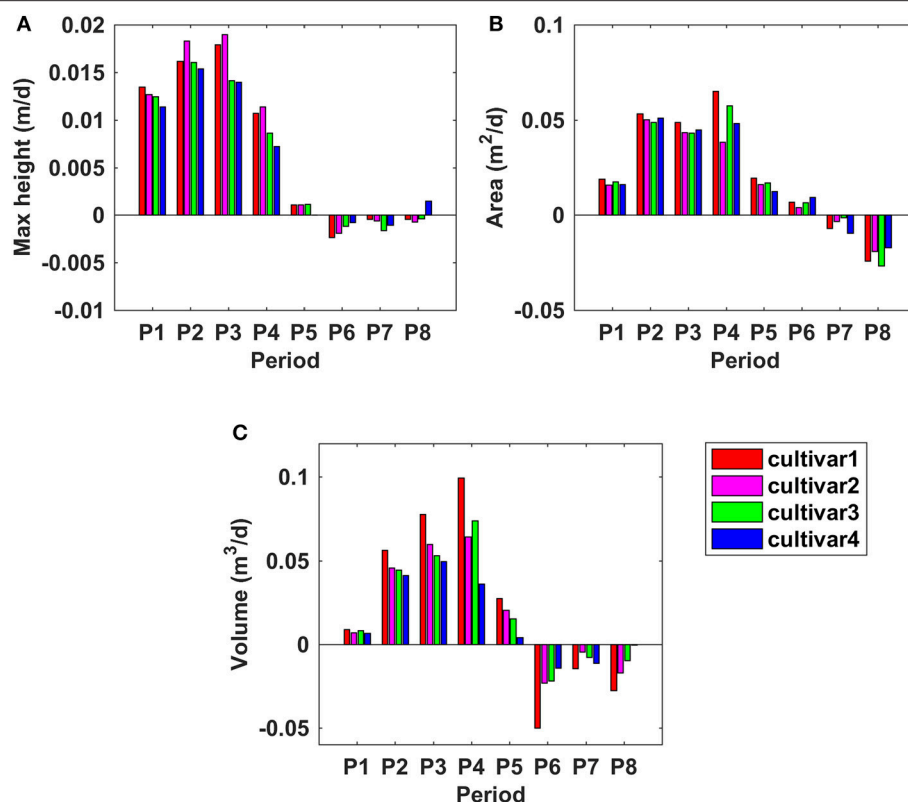


FIGURE 9 | Growth rates for derived morphologic parameters at different time frames during the monitoring period. **(A)** Maximum canopy height; **(B)** projected canopy area; **(C)** plant volume.

TABLE 2 | Effects of cultivars on derived parameters over the monitoring period (Year: 2016).

	July 28 (DAP 45)	Aug 04 (DAP 52)	Aug 19 (DAP 67)	Aug 26 (DAP 74)	Sep 09 (DAP 88)	Sep 16 (DAP 95)	Sep 23 (DAP 102)	Sep 30 (DAP 109)
Max CH	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**
PCA	0.003**	0.003**	0.034*	0.005**	0.001**	0.001**	0.002**	0.006**
PV	0.002**	0.001**	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**	<0.001**

*Significant at the 0.05 probability level, **Significant at the 0.01 probability level.

TABLE 3 | Differences in final yield between four cultivars.

Cultivar	Mean yield (g/plot)
Cultivar1	928.51 ^{AB}
Cultivar2	781.58 ^B
Cultivar3	954.18 ^A
Cultivar4	937.22 ^A

$p < 0.05$. Different letters indicate significant differences between cultivars.

reached the highest R^2 -value (0.56) on 67 DAP—although the PV R^2 -values were less than those of PCA from 67 to 109 DAP, they were more stable than those of PCA. The lowest correlation was found between CH and final yield. A decreasing trend was observed from 45 to 109 DAP which was mainly due to cultivar 2.

DISCUSSION

We have demonstrated that 3D point clouds reconstructed based on data collected by a 2D LiDAR and an RTK-GPS, and associated data processing methodology, accurately estimated specific morphologic traits from cotton plants under field conditions. The precise 3D point cloud, which was the basic dataset for analysis, was critical for the successful extraction of plant morphologic traits (Paprocki et al., 2012; Paulus et al., 2014b; Duan et al., 2016). The reconstructed 3D model visually represented plant canopy structures (Figure 5). The CH, PCA, and PV derived by our system were highly correlated (the slopes were close to one) with those measured manually (Figure 6).

The system demonstrated great potential for high-throughput phenotypic analysis. The number of points for each plot varied from 2,000 to 20,000 over the monitoring period, depending on the plant size. The tractor was driven in the field at a speed of

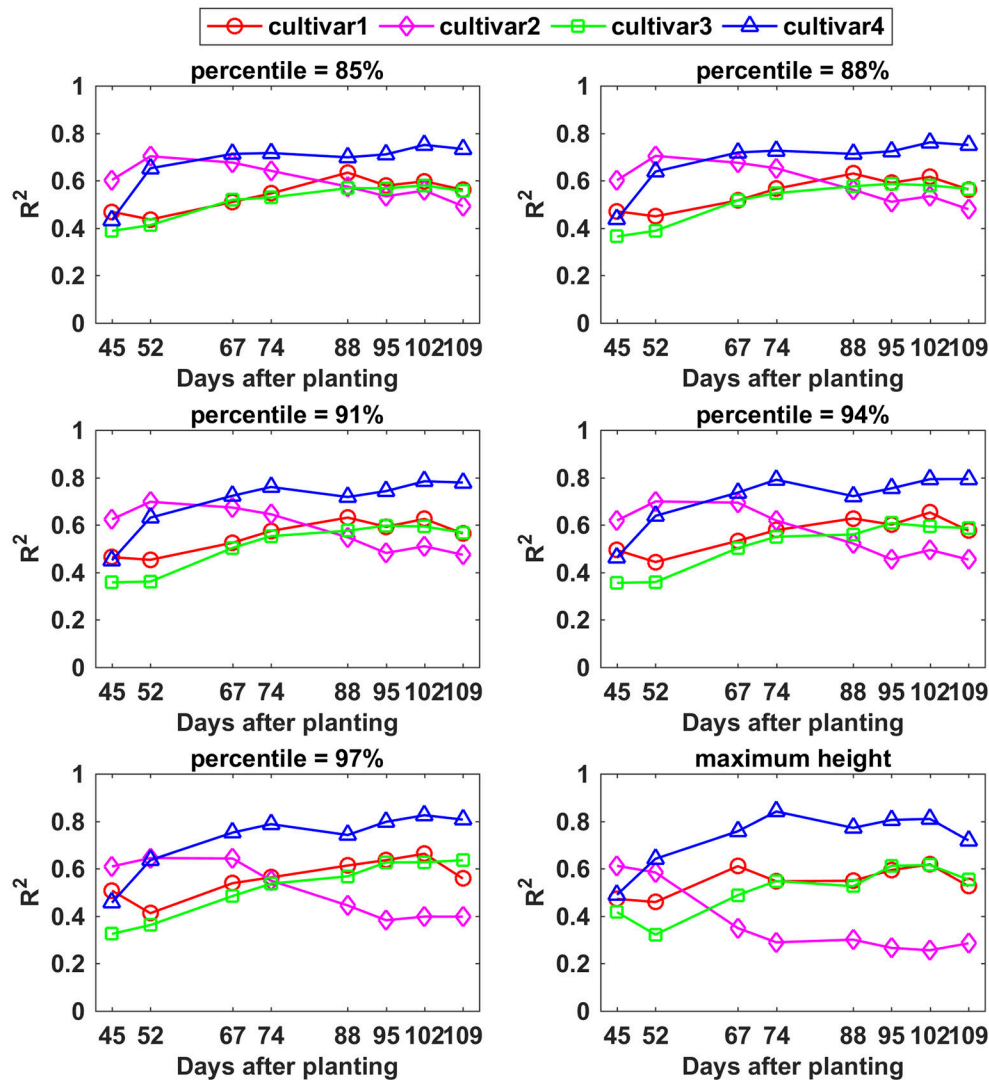


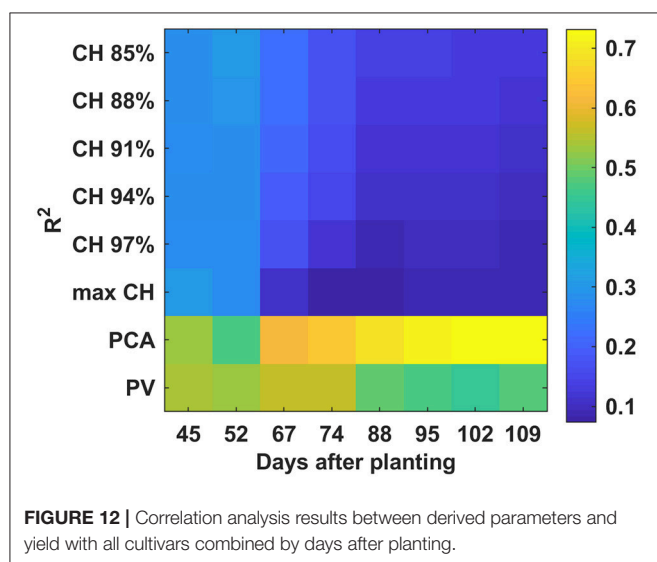
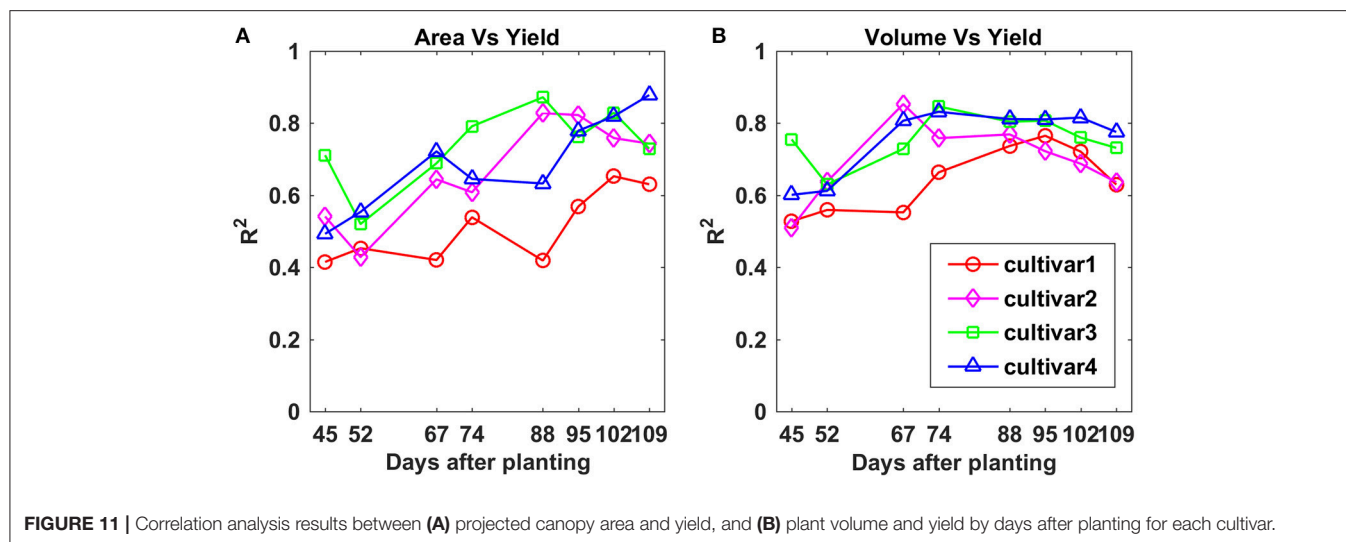
FIGURE 10 | Correlation analysis results between different percentiles of canopy height and yield by days after planting for each cultivar.

around 0.5 m/s for data collection, taking about 6 s to scan one plot that contained 15 plants. Thus, one could anticipate scanning about 600 such plots per hour. The average time consumed for 3D reconstruction and parameter extraction was around 5.27 s per plot so about an hour for the aforementioned 600-plot field. Therefore, both data acquisition and data analysis speeds are suitable for application to field experiments of the sizes used in many breeding programs. The system also has great potential to implement online data analysis with faster computing power in the future, which is particularly useful for large field applications.

The high throughput phenotypic analysis by a LiDAR-based method has several advantages compared to image-based methods. Paproki et al. (2012) conducted cotton plant morphologic trait analysis using an image-based 3D reconstruction method, requiring about 7 min to collect images for each plant and an average processing time of 15 min for 3D

reconstruction work of each plant. Additionally, image quality could be significantly affected by highly variable illumination conditions (Nuske et al., 2014), which limited its in-field applications. In contrast, LiDAR is more versatile in a wide range of light conditions since it is equipped with its own light source. For manual operation, ~30 min were required for a typical analysis per plant depending on the size and complexity, and destructive harvests were often required (Paproki et al., 2012).

The high throughput of the proposed system and its non-invasive features permit data acquisition repeatedly over the growing season, opening the door to acquisition of plant growth rate variation, which is valuable for a wide range of applications such as building plant growth models (Tessmer et al., 2013; Weraduwaage et al., 2015), exploring factors influencing growth processes (Rahaman et al., 2015; Awlia et al., 2016), genomics-assisted crop breeding (Watanabe et al., 2017), and QTL analysis



(Bac-Molenaar et al., 2015; Cui et al., 2017). In this study, all plants in the experimental field were scanned from 43 to 109 DAP. Plant growth curves were generated and correlation between derived parameters and the final yield were explored, which could be used for yield prediction (Sharma et al., 2015) and species identification (Remagnino et al., 2016). Thus, the system described herein not only provides an indication of which LiDAR-derived parameters are most closely associated with yield, but also allows us to define the period during the growing season in which a morphologic trait has the greatest correlation with final yield.

While CH in the current study was highly correlated with within-cultivar yield variation in some instances (i.e., cultivar 4 on later sample dates), it should be noted that correlations between CH and yield for all cultivars combined were substantially lower than those obtained for other derived parameters such as PCA and PV. This is not surprising since

the cotton crop exhibits a quadratic response of yield to CH (Sui et al., 2012), where internode elongation must often be controlled using exogenously applied plant growth regulators (PGRs) in real-world production scenarios (Dodds et al., 2010). While CH may not be inherently predictive of genotypic differences in yield, the deployment of high-throughput methods for estimating maximum CH and rate of change in CH could strongly influence cultivar-specific PGR management strategies. Specifically, cultivars with rapid plant height development typically require more aggressive management strategies (Collins, 2013).

PCA and PV were more strongly associated with yield than CH when considered across all cultivars, suggesting these parameters might have greater utility as high-throughput phenotyping tools to identify potential differences in productivity. The impact of canopy development on yield is easily understood when yield is expressed as the product of total intercepted photosynthetically active radiation, radiation use efficiency, and harvest index (Monteith and Moss, 1977; Monteith, 1994; Earl and Davis, 2003; Stöckle and Kemanian, 2009). Within this conceptual framework, the amount of intercepted photosynthetically active radiation during a growing season will be strongly impacted by the length of the growing season and the leaf area available to intercept solar radiation. Thus, PCA may serve as a suitable proxy for leaf area magnitude and persistence throughout the growing season. PV incorporates both CH and PCA, which could potentially be related to both the amount of leaf area available to intercept incoming solar radiation and the three dimensional space available for developing fruit. Both of these considerations are important given the indeterminate nature of cotton. As fruit development progresses, vegetative growth slows, and Constable and Bange (2015) have clearly illustrated that the ability of the cotton crop to attain high yield potential will be influenced by total leaf area and number of fruiting sites available prior to the development of a large fruiting load on the plant. Thus, it is not surprising that PCA and PV were correlated with yield for much of the

growing season. Future work should couple the aforementioned measures with high-throughput assessment of fruit development as fruiting dynamics can drastically impact final yield in cotton (Constable and Bange, 2015).

The top view scan setting of LiDAR provided precise 3D surface models of plants; however, organs under the canopy could not be reached due to plant occlusion. This does not reduce the accuracy of CH and PCA measurements, but hinders the estimation of PV by our system, especially when plant structure becomes complex. Multi-view scan is a commonly used method to reduce plant occlusion effects (Paulus et al., 2014b), and in our future studies two side view scans could be added. In addition, algorithms such as plant shape models (Pastrana and Rath, 2013) have been proposed to overcome plant occlusion effects. However, when plant structure complexity increases, especially at the mature stage with a greatly increased number and size of leaves and stems, occlusion effects would still be present. Plant occlusion remains a challenge for plant phenotyping, especially under field conditions with limited inter-row spacing (Paprocki et al., 2012; Paulus et al., 2014a). Wind is another factor which could affect the accuracy of the derived traits since it might result in blurred point clouds. Therefore, calm weather conditions are best for data collection. Based on the CH validation results, our system has a certain robustness against wind influence; wind speed under 3 m/s was feasible. Another limitation was that the tractor speed was restricted by the scanning frequency of the LiDAR (Sun et al., 2017), limiting throughput. While the present system is already compatible with iterative study of thousands of plots, 3D LiDAR technology may greatly reduce this limitation (Weiss and Biber, 2011).

CONCLUSION

Precise 3D surface models were reconstructed by the high-throughput phenotyping system developed in this study under field conditions. Multiple morphologic traits at the plot level including plant height, projected canopy area, and plant volume were extracted simultaneously. The system could be used to scan the field repeatedly due to its relatively high data collection and

processing capability, which was particularly useful for large field applications. The measured morphologic traits were most highly correlated with final yield during the period between around 67 and 109 DAP. Projected canopy area and plant volume were more closely correlated than plant height to final yield. Future work will focus on involving other sensor data to extract more phenotypic traits from the 3D point cloud. Although this system was solely tested on cotton plants, it is expected to be applicable for use with other crops such as wheat, rice, and soybeans.

AUTHOR CONTRIBUTIONS

SS, CL, and AP: conceived the idea and designed the experiments; AP and PC: contributed to the preparation of materials and equipment; SS, YJ, RX, and JR: conducted the experiments; SS, YJ, and RX: analyzed the data; SS, CL, AP, JS, and PC: interpreted results and wrote the paper; All authors read and approved the final manuscript.

FUNDING

This study was funded jointly by the Agricultural Sensing and Robotics Initiative of the College of Engineering, and the College of Agricultural and Environmental Sciences of the University of Georgia. The project was also partially supported by the National Robotics Initiative grant (NIFA grant No. 2017-67021-25928).

ACKNOWLEDGMENTS

The authors gratefully thank Tariq Shehzad, Jeevan Adhikari, Jamal Hunter, Mengyun Zhang, Zikai Wei, and Shuxiang Fan for their assistance in conducting the experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00016/full#supplementary-material>

REFERENCES

- Awlia, M., Nigro, A., Faikus, J., Schmoedel, S. M., Negrão, S., Santelia, D., et al. (2016). High-throughput non-destructive phenotyping of traits that contribute to salinity tolerance in *Arabidopsis thaliana*. *Front. Plant Sci.* 7:1414. doi: 10.3389/fpls.2016.01414
- Bac-Molenaar, J. A., Vreugdenhil, D., Granier, C., and Keurentjes, J. J. (2015). Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. *J. Exp. Bot.* 66, 5567–5580. doi: 10.1093/jxb/erv176
- Bietresato, M., Carabin, G., Vidoni, R., Gasparetto, A., and Mazzetto, F. (2016). Evaluation of a LiDAR-based 3D-stereoscopic vision system for crop-monitoring applications. *Comput. Electron. Agric.* 124, 1–13. doi: 10.1016/j.compag.2016.03.017
- Bussemeyer, L., Ruckelshausen, A., Möller, K., Melchinger, A. E., Alheit, K. V., Maurer, H. P., et al. (2013). Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. *Sci. Rep.* 3:2442. doi: 10.1038/srep02442
- Cobb, J. N., DeClerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theor. Appl. Genet.* 126, 867–887. doi: 10.1007/s00122-013-2066-0
- Collins, G. (2013). *Plant Growth Regulator (PGR) Management Considerations for the 2013 Crop*. Available online at: <http://www.ugacotton.com/2013/07/plant-growth-regulator-pgr-management-considerations-for-the-2013-crop-collins-and-whitaker/>
- Constable, G. A., and Bange, M. P. (2015). The yield potential of cotton (*Gossypium hirsutum* L.). *Field Crops Res.* 182(Suppl. C), 98–106. doi: 10.1016/j.fcr.2015.07.017
- Cui, T.-T., He, K.-H., Chang, L.-G., Zhang, X.-H., Xue, J.-Q., and Liu, J.-C. (2017). QTL mapping for leaf area in maize (*Zea mays* L.) under multi-environments. *J. Integr. Agric.* 16, 800–808. doi: 10.1016/S2095-3119(16)61524-1
- Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., and Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 4, 349–379. doi: 10.3390/agronomy4030349

- Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 433–444. doi: 10.1016/j.tplants.2013.04.008
- Dodds, D. M., Banks, J., Barber, L. T., Boman, R. K., Brown, S. M., Edmisten, K. L., et al. (2010). Beltwide evaluation of commercially available plant growth regulators. *J. Cotton Sci.* 14, 119–130. Available online at: <http://www.cotton.org/journal/2010-14/3/upload/JCS14-119.pdf>
- Duan, T., Chapman, S. C., Holland, E., Rebetzke, G. J., Guo, Y., and Zheng, B. (2016). Dynamic quantification of canopy structure to characterize early plant vigour in wheat genotypes. *J. Exp. Bot.* 67, 4523–4534. doi: 10.1093/jxb/erw227
- Earl, H. J., and Davis, R. F. (2003). Effect of drought stress on leaf and whole canopy radiation use efficiency and yield of maize. *Agron. J.* 95, 688–696. doi: 10.2134/agronj2003.0688
- Friedli, M., Kirchgessner, N., Grieder, C., Liebisch, F., Mannale, M., and Walter, A. (2016). Terrestrial 3D laser scanning to track the increase in canopy height of both monocot and dicot crop species under field conditions. *Plant Methods* 12:15. doi: 10.1186/s13007-016-0109-7
- Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005
- Gardner, M. P. (1985). Mood states and consumer behavior: a critical review. *J. Consum. Res.* 12, 281–300. doi: 10.1086/208516
- Gebbers, R., Ehler, D., and Adamek, R. (2011). Rapid mapping of the leaf area index in agricultural crops. *Agron. J.* 103, 1532–1541. doi: 10.2134/agronj2011.0201
- Gerland, P., Raftery, A. E., Sevciková, H., Li, N., Gu, D., Spoorenberg, T., et al. (2014). World population stabilization unlikely this century. *Science* 346, 234–237. doi: 10.1126/science.1257469
- Gibbs, J. A., Pound, M., French, A. P., Wells, D. M., Murchie, E., and Pridmore, T. (2017). Approaches to three-dimensional reconstruction of plant shoot topology and geometry. *Funct. Plant Biol.* 44, 62–75. doi: 10.1071/FP16167
- Goggin, F. L., Lorence, A., and Topp, C. N. (2015). Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops. *Curr. Opin. Insect Sci.* 9, 69–76. doi: 10.1016/j.cois.2015.03.002
- Großkinsky, D. K., Pieruschka, R., Svensgaard, J., Rascher, U., Christensen, S., Schurr, U., et al. (2015). Phenotyping in the fields: dissecting the genetics of quantitative traits and digital farming. *New Phytol.* 207, 950–952. doi: 10.1111/nph.13529
- Hosoi, F., and Omasa, K. (2012). Estimation of vertical plant area density profiles in a rice canopy at different growth stages by high-resolution portable scanning lidar with a lightweight mirror. *Isprs J. Photogramm. Remote Sens.* 74, 11–19. doi: 10.1016/j.isprsjprs.2012.08.001
- Jiang, Y., Li, C., and Paterson, A. H. (2016). High throughput phenotyping of cotton plant height using depth images under field conditions. *Comput. Electron. Agric.* 130, 57–68. doi: 10.1016/j.compag.2016.09.017
- Krieg, D. R., and Sung, J. (1986). “Source-sink relations as affected by water stress during boll development,” in *Cotton Physiology*, eds J. Mauney and J. Stewart (Memphis, TN: The Cotton Foundation), 73–78.
- Li, L., Zhang, Q., and Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors* 14, 20078–20111. doi: 10.3390/s141120078
- Li, Y., Cao, Z., Lu, H., Xiao, Y., Zhu, Y., and Cremers, A. B. (2016). In-field cotton detection via region-based semantic image segmentation. *Comput. Electron. Agric.* 127, 475–486. doi: 10.1016/j.compag.2016.07.006
- Lin, Y. (2015). LiDAR: an important tool for next-generation phenotyping technology of high potential for plant phenomics? *Comput. Electron. Agric.* 119, 61–73. doi: 10.1016/j.compag.2015.10.011
- Llorens, J., Gil, E., and Llop, J. (2011). Ultrasonic and LIDAR sensors for electronic canopy characterization in vineyards: advances to improve pesticide application methods. *Sensors* 11, 2177–2194. doi: 10.3390/s110202177
- Monteith, J. L. (1994). Validity of the correlation between intercepted radiation and biomass. *Agric. For. Meteorol.* 68, 213–220. doi: 10.1016/0168-1923(94)90037-X
- Monteith, J. L., and Moss, C. (1977). Climate and the efficiency of crop production in Britain. *Philos. Trans. R. Soc. Lond. B* 281, 277–294. doi: 10.1098/rstb.1977.0140
- Nuske, S., Wilshusen, K., Achar, S., Yoder, L., and Singh, S. (2014). Automated visual yield estimation in Vineyards. *J. Field Robot.* 31, 837–860. doi: 10.1002/rob.21541
- Paproki, A., Sirault, X., Berry, S., Furbank, R., and Fripp, J. (2012). A novel mesh processing based technique for 3D plant analysis. *BMC Plant Biol.* 12:63. doi: 10.1186/1471-2229-12-63
- Pastrana, J. C., and Rath, T. (2013). Novel image processing approach for solving the overlapping problem in agriculture. *Biosyst. Eng.* 115, 106–115. doi: 10.1016/j.biosystemseng.2012.12.006
- Paulus, S., Dupuis, J., Riedel, S., and Kuhlmann, H. (2014a). Automated analysis of barley organs using 3D laser scanning: an approach for high throughput phenotyping. *Sensors* 14, 12670–12686. doi: 10.3390/s140712670
- Paulus, S., Schumann, H., Kuhlmann, H., and Léon, J. (2014b). High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosyst. Eng.* 121, 1–11. doi: 10.1016/j.biosystemseng.2014.01.010
- Rahaman, M. M., Chen, D., Gillani, Z., Klukas, C., and Chen, M. (2015). Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Front. Plant Sci.* 6:619. doi: 10.3389/fpls.2015.00619
- Remagnino, P., Mayo, S., Wilkin, P., Cope, J., and Kirkup, D. (2016). *Computational Botany: Methods for Automated Species Identification*. Berlin: Springer.
- Ritchie, G. L., Bednars, C. W., Jost, P. H., and Brown, S. M. (2007). *Cotton Growth and Development*. Available online at: <http://cotton.tamu.edu/General%20Production/Georgia%20Cotton%20Growth%20and%20Development%20B1252-1.pdf>
- Rosell-Polo, J. R., Llorens, J., Sanz, R., Arnó, J., Ribes-Dasi, M., Masip, J., et al. (2009). Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning. *Agric. For. Meteorol.* 149, 1505–1515. doi: 10.1016/j.agrformet.2009.04.008
- Rusu, R. B., Marton, Z. C., Blodow, N., Dolha, M., and Beetz, M. (2008). Towards 3D point cloud based object maps for household environments. *Robot. Auton. Syst.* 56, 927–941. doi: 10.1016/j.robot.2008.08.005
- Sharma, B., Mills, C. I., Snowden, C., and Ritchie, G. L. (2015). Contribution of boll mass and boll number to irrigated cotton yield. *Agron. J.* 107, 1845–1853. doi: 10.2134/agronj15.0024
- Sharma, B., and Ritchie, G. L. (2015). High-Throughput Phenotyping of cotton in multiple irrigation environments. *Crop Sci.* 55, 958–969. doi: 10.2135/cropsci2014.04.0310
- Si, Y., Liu, G., and Feng, J. (2015). Location of apples in trees using stereoscopic vision. *Comput. Electron. Agric.* 112, 68–74. doi: 10.1016/j.compag.2015.01.010
- Simko, I., Jimenez-Berni, J. A., and Sirault, X. R. (2016). Phenomic approaches and tools for phytopathologists. *Phytopathology* 107, 6–17. doi: 10.1094/PHYTO-02-16-0082-RVW
- Stöckle, C., and Kemanian, A. (2009). “Crop radiation capture and use efficiency: a framework for crop growth analysis,” in *Crop Physiology: Applications for Genetic Improvement and Agronomy* (San Diego, CA: Academic Press), 145–170. doi: 10.1016/B978-0-12-374431-9.00007-4
- Sui, R., Fisher, D. K., and Reddy, K. N. (2012). Cotton yield assessment using plant height mapping system. *J. Agric. Sci.* 5:23. doi: 10.5539/jas.v5n1p23
- Sun, S., Li, C., and Paterson, A. H. (2017). In-field high-throughput phenotyping of cotton plant height using LiDAR. *Remote Sens.* 9:377. doi: 10.3390/rs9040377
- Taheriazad, L., Moghadas, H., and Sanchez-Azofeifa, A. (2016). “A new approach to calculate Plant Area Density (PAD) using 3D ground-based LiDAR,” in *Earth Resources and Environmental Remote Sensing/GIS Applications VII*, eds U. Michel, K. Schulz, M. Ehlers, K. G. Nikolakopoulos, and D. Civco (Bellingham, WA: Spie-Int Soc Optical Engineering), 1000518-1–1000518-10.
- Tessmer, O. L., Jiao, Y., Cruz, J. A., Kramer, D. M., and Chen, J. (2013). Functional approach to high-throughput plant growth analysis. *BMC Syst. Biol.* 7:S17. doi: 10.1186/1752-0509-7-S6-S17
- Nguyen T. T., Slaughter, D. C., Max, N., Maloof, J. N., and Sinha, N. (2015). Structured light-based 3D reconstruction system for plants. *Sensors* 15, 18587–18612. doi: 10.3390/s150818587
- Tilman, D., Balzer, C., Hill, J., and Bafort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20260–20264. doi: 10.1073/pnas.1116437108

- Vazquez-Arellano, M., Griepentrog, H. W., Reiser, D., and Paraforos, D. S. (2016). 3-D imaging systems for agricultural applications-a review. *Sensors* 16:24. doi: 10.3390/s16050618
- Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., et al. (2017). High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. *Front. Plant Sci.* 8:421. doi: 10.3389/fpls.2017.00421
- Weiss, U., and Biber, P. (2011). Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. *Rob. Auton. Syst.* 59, 265–273. doi: 10.1016/j.robot.2011.02.011
- Weraduwage, S. M., Chen, J., Anozie, F. C., Morales, A., Weise, S. E., and Sharkey, T. D. (2015). The relationship between leaf area growth and biomass accumulation in *Arabidopsis thaliana*. *Front. Plant Sci.* 6:167. doi: 10.3389/fpls.2015.00167
- Xiong, X., Yu, L., Yang, W. N., Liu, M., Jiang, N., Wu, D., et al. (2017). A high-throughput stereo-imaging system for quantifying rape leaf traits during the seedling stage. *Plant Methods* 13:17. doi: 10.1186/s13007-017-0157-7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sun, Li, Paterson, Jiang, Xu, Robertson, Snider and Chee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Considerations for Achieving Cross-Platform Point Cloud Data Fusion across Different Dryland Ecosystem Structural States

Tyson L. Swetnam^{1*}, Jeffrey K. Gillan², Temuulen T. Sankey³, Mitchel P. McClaran², Mary H. Nichols⁴, Philip Heilman⁴ and Jason McVay³

¹ BIO5 Institute, University of Arizona, Tucson, AZ, United States, ² School of Natural Resource and Environment, University of Arizona, Tucson, AZ, United States, ³ Informatics and Computing Program, Northern Arizona University, Flagstaff, AZ, United States, ⁴ USDA Agricultural Research Service, Southwest Watershed Research Center, Tucson, AZ, United States

OPEN ACCESS

Edited by:

Alexander Bucksch,
University of Georgia, United States

Reviewed by:

Jinhu Wang,
Delft University of Technology,
Netherlands

Andrew M. Cunliffe,
University of Edinburgh,
United Kingdom

*Correspondence:

Tyson L. Swetnam
tswetnam@email.arizona.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 09 September 2017

Accepted: 04 December 2017

Published: 10 January 2018

Citation:

Swetnam TL, Gillan JK, Sankey TT, McClaran MP, Nichols MH, Heilman P and McVay J (2018) Considerations for Achieving Cross-Platform Point Cloud Data Fusion across Different Dryland Ecosystem Structural States. *Front. Plant Sci.* 8:2144. doi: 10.3389/fpls.2017.02144

Remotely sensing recent growth, herbivory, or disturbance of herbaceous and woody vegetation in dryland ecosystems requires high spatial resolution and multi-temporal depth. Three dimensional (3D) remote sensing technologies like lidar, and techniques like structure from motion (SfM) photogrammetry, each have strengths and weaknesses at detecting vegetation volume and extent, given the instrument's ground sample distance and ease of acquisition. Yet, a combination of platforms and techniques might provide solutions that overcome the weakness of a single platform. To explore the potential for combining platforms, we compared detection bias amongst two 3D remote sensing techniques (lidar and SfM) using three different platforms [ground-based, small unmanned aerial systems (sUAS), and manned aircraft]. We found aerial lidar to be more accurate for characterizing the bare earth (ground) in dense herbaceous vegetation than either terrestrial lidar or aerial SfM photogrammetry. Conversely, the manned aerial lidar did not detect grass and fine woody vegetation while the terrestrial lidar and high resolution near-distance (ground and sUAS) SfM photogrammetry detected these and were accurate. UAS SfM photogrammetry at lower spatial resolution under-estimated maximum heights in grass and shrubs. UAS and handheld SfM photogrammetry in near-distance high resolution collections had similar accuracy to terrestrial lidar for vegetation, but difficulty at measuring bare earth elevation beneath dense herbaceous cover. Combining point cloud data and derivatives (i.e., meshes and rasters) from two or more platforms allowed for more accurate measurement of herbaceous and woody vegetation (height and canopy cover) than any single technique alone. Availability and costs of manned aircraft lidar collection preclude high frequency repeatability but this is less limiting for terrestrial lidar, sUAS and handheld SfM. The post-processing of SfM photogrammetry data became the limiting factor at larger spatial scale and temporal repetition. Despite the utility of sUAS and handheld SfM for monitoring vegetation phenology and structure, their spatial extents are small relative to manned aircraft.

Keywords: lidar, terrestrial laser scanning, structure from motion, sUAS

INTRODUCTION

Measurement and monitoring of ecological processes are limited by what Levin (1992) termed the ‘problem of pattern and scale’ where linking observations across cells, leaves, plants, community, and ecosystem require exponential amounts of information be transferred between fine and broad spatial scale, short and long temporal scale. Conventional ecological research studies require tens to hundreds of small quadrats or plots between 1 and 100 meters square (m^2) for enough observations to ensure statistically significant p -values for testing hypotheses (Huenneke et al., 2001; Kachamba et al., 2017). This change in scale from individual plot to ecosystem inventory is perhaps the most important concept of remote sensing in the life sciences (Woodcock and Strahler, 1987; Turner, 1989; Turner et al., 1989).

Dryland ecosystems, characterized by sparse or patchy vegetation and long periods of senescence, are punctuated by short periods of rapid growth following seasonal precipitation events. Therefore, characterization of both spatial and temporal patterns of vegetation abundance requires both high spatial resolution and temporal repetition over short periods of time. For instance, landscape-scale models of primary production and carbon-uptake should be more accurate when the vegetation abundance is frequently represented at a fine, plant-scale spatial resolution. Recent advances in sensor and platform technology appear to bring these resolutions within reach, but better understanding of detection bias and accuracy are needed, as well as examining the potential for combining multiple sources in a data-fusion approach that builds on strengths of one system to avoid the weakness of the other. For example, measuring the height and volume of an individual tree or shrub requires accurate minimum elevation (bare ground), apical leader, and crown diameter measurements. Given these requirements, the data may become unusable if bare ground is obstructed or incompletely illuminated. In these situations, multiple observations from different platforms may be required to generate a better estimate.

Light detection and ranging (lidar) is the predominant technology for measuring vegetation and earth surface phenomena in three dimensions (3D) (Glennie et al., 2013; Harpold et al., 2015). Manned aircraft equipped with lidar now survey local areas to entire ecosystems or countries (Higgins et al., 2014; Stoker et al., 2016), while terrestrial lidar is used to collect data at sub-centimeter resolution over hectare size areas. Critically, the spatial scale needed for managing at landscape scale can be provided by manned aerial lidar but the scale needed for monitoring ecosystem process often requires resolution that can only be provided by terrestrial lidar.

Advancements in computational processing power and machine vision also allow 3D object reconstruction from nadir and oblique 2D imagery by so-called structure from motion (SfM) photogrammetry and multi-view stereo (MVS) methods (Westoby et al., 2012; Carrivick et al., 2016; Wallace et al., 2016; Mlambo et al., 2017), hereafter referred to simply as “SfM.” Under the right conditions small unmanned aerial systems (sUAS) with SfM can generate sub-centimeter precision point clouds and create digital elevation models (DEM) and digital surface models

(DSM) (Dandois and Ellis, 2010; Dandios and Ellis, 2013; Cruzan et al., 2016). Cruzan et al. (2016) give excellent examples of the capability of small and micro UAS in plant ecology.

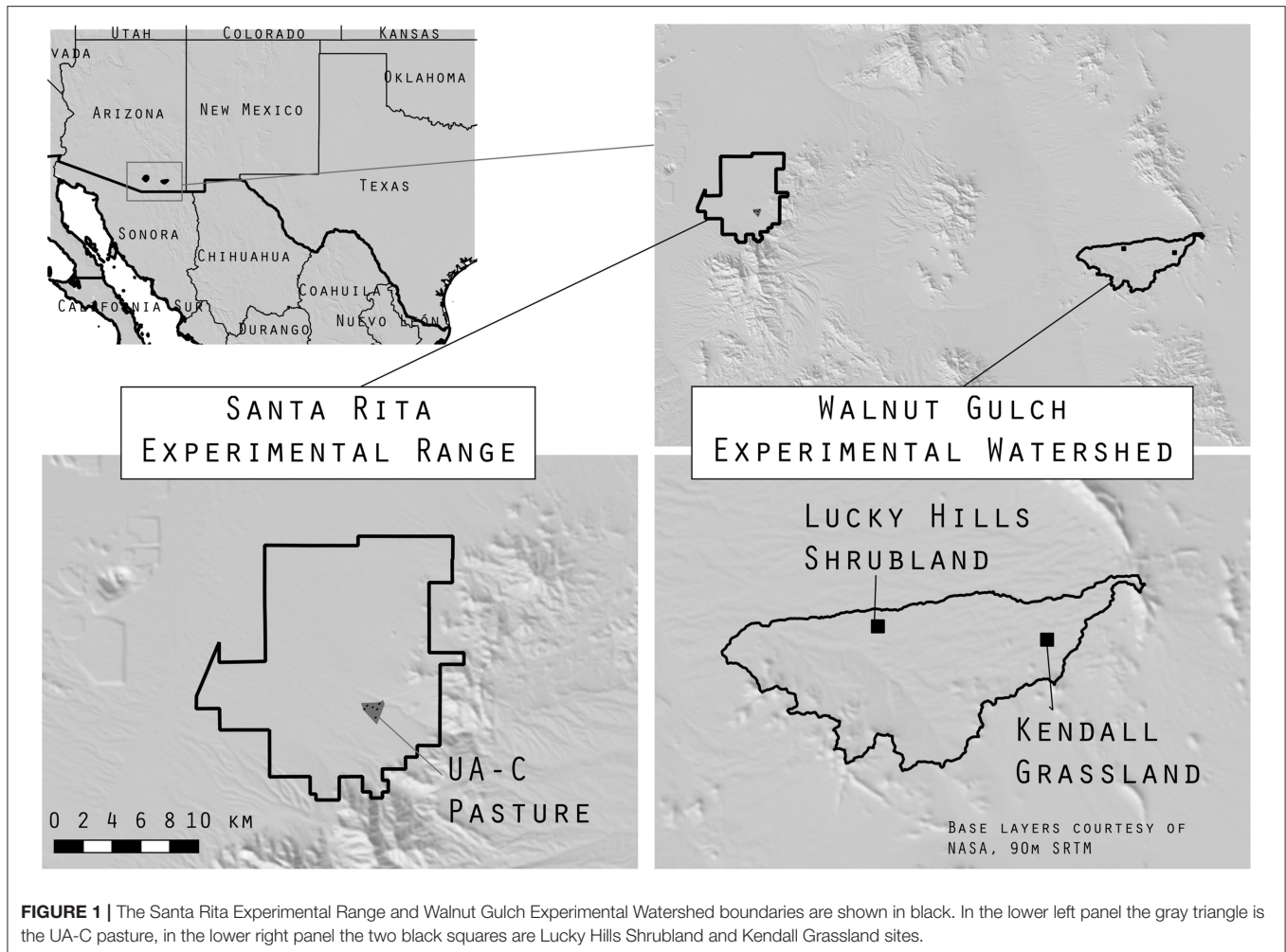
In the present study, we compare detection biases amongst different 3D remote sensing techniques (lidar and SfM photogrammetry), from three platforms (ground, sUAS, and manned aircraft) in various dryland ecosystem structural states. We combine the new and less expensive technologies, specifically sUAS SfM, with manned aerial lidar to produce more representative measurements of vegetation height and volume. The objectives of the paper are (1) to identify detection biases amongst technologies for four general feature classes: bare earth (both in barren areas and beneath vegetation), herbaceous plants, low woody shrubs, tall woody shrubs and trees, and (2) explore the potential for combining or fusing observations from any two platforms to overcome the detection bias of another. For Objective 1, we use point cloud-to-cloud inter-comparison and raster-based differencing to establish the detection bias. For Objective 2, we use raster-based differencing and also mesh-to-point cloud differencing to establish improved accuracy by combining data sources.

Current trends in research are moving toward collaborative open data and code sharing (Boulton et al., 2011; Kitchin, 2014; Hampton et al., 2015). Version control systems, like GitHub, offer common repositories for software and code development. In this paper, we provide only abbreviated descriptions of the technical details involving the image processing and point cloud alignment workflows in our main text, and instead provide those details in a supplemental public GitHub repository: <https://github.com/tyson-swetnam/srer-wgew>. We also provide a quick reference table for acronyms and abbreviations used in this text in the supplemental materials (Table S1).

METHODS

Study Areas and Species

The Walnut Gulch Experimental Watershed (WGEW) is a long-term research site surrounding the town of Tombstone, Arizona (30.74° N, -110.05° W) (Figure 1) administered by the United States Department of Agriculture, Agricultural Research Service Southwest Watershed Research Center. The soils in WGEW vary from high in carbonates on the western lower watershed to sandy gravel loams on the eastern upper watershed (Keefer et al., 2008; Osterkamp, 2008; Pelletier et al., 2016). Mean annual temperature in WGEW, measured in Tombstone at 1,384 m a.s.l., is 17.6°C and mean annual precipitation is 300 mm year^{-1} . The Santa Rita Experimental Range (SRER) near Tucson, Arizona (31.80° N, -110.84° W) (Figure 1) is administered by the University of Arizona (<http://cals.arizona.edu/srer>) (McClaran et al., 2003). The SRER soils are characterized as clay loams, sandy loams, and limey upland soils (Brekenfeld and Robinett, 1997). Mean annual temperature and precipitation are 19°C and 358 mm year^{-1} , respectively (McClaran and Wei, 2014). Both study sites are typical of dryland ecosystems, characterized as regions where evaporation is 2–5 times greater than precipitation (Safriel et al., 2006).



At WGEW, we worked in two areas: Lucky Hills Shrubland and Kendall Grassland (**Figure 1**). Lucky Hills is characterized as Chihuahuan desert scrub with shrubs as the dominant life form (Scott et al., 2006; Scott, 2010b). Common species are *Larrea tridentata* (creosote), *Vachellia vernicosa* (whitethorn acacia), *Flourensia cernua* (tarbush), *Parthenium incanum* (mariola), *Rhus microphylla* (little-leaf desert sumac), *Condalia warnockii* (Warnock's Snakewood), *Ephedra triferca* (Mormon tea). The Kendall Grassland site is characterized as semi-arid desert grassland with a mixture of shrub, grass and cactus life forms (Scott et al., 2006; Scott, 2010b). Common species are *Eragrostis lehmanniana* (Lehmann's lovegrass) a non-native grass, and *Prosopis velutina* (velvet mesquite), native grasses include *Hilaria belangeri* (curly mesquite), *Bouteloua eriopoda* (black grama), *B. hirsuta* (hairy grama), and *Aristida hamulosa* (threeawn) (Skirvin et al., 2008), other species include *Yucca baccata* (Banana yucca), *Y. elata* (Soap-tree yucca), and *Agave palmeri* (Palmer's agave).

At SRER, we worked in another semi-desert grassland area with mixture of shrubs, grass, and cactus (Pasture UA-C). Common species are *P. velutina*, *Opuntia engelmannii* (Engelmann's prickly pear cactus), *Acacia greggii* (catclaw

acacia), *Mimosa dysocarpa* (velvet pod mimosa), and *Calliandra eriophylla* (false-mesquite), and many species of perennial grass (McClaran et al., 2010; Scott, 2010a).

Software, Code, and Computing

Links to the aerial lidar vendor reports, code, and analysis workflows are maintained in a public GitHub repository: <https://github.com/tyson-swetnam/srer-wgew>. Watershed boundaries and major infrastructure features in WGEW are available from Heilman et al. (2008), Moran et al. (2008), and the SRER data portal: <https://cals.arizona.edu/srer/>.

For the SfM point cloud creation we used proprietary software packages including: Pix4D (SenseFly, Switzerland) and Photoscan (AgiSoft, 2017). Analysis of point clouds and derivative models were also done in open-source CloudCompare (Girardeau-Montaut, 2011), Point Data Abstraction Library (PDAL, Hobu Inc., 2017) with Docker Community Edition (Merkel, 2014; Boettiger, 2015; Docker Development Team, 2017), and QGIS (Quantum GIS Development Team, 2017). SfM generation of dense point clouds in Agisoft Photoscan took from 1 to 192 h, depending on the number of images used in the reconstruction. Point cloud analyses in CloudCompare were

typically less than 5 min per procedure, with large multi-scale model to model cloud comparison (M3C2, Lague et al., 2013) analyses required up to 20 min.

NSF Jetstream cloud computing service (Stewart et al., 2015) was used to filter point clouds for outliers and to translate projections for the aerial lidar and SfM data to common datums using the Docker version of PDAL (docker://pdal/pdal:1.5). PDAL and LAStools (Isenburg, 2014) were also used to remove outliers from the original *.las data files and to compress to *.laz format for online data hosting.

Instruments and Field Methods

Characteristics of lidar sensors varied among terrestrial, manned aerial, and sUAS platforms (Table 1). At the Lucky Hills site, the terrestrial lidar survey (TLS) included nine 360 degree scans on October 1, 2015; and at the Kendall grassland site, eight scan and six scan collections were made on September 23, 2015 and October 8, 2015, respectively. Eight terrestrial lidar scans were made at the SRER UA-C site on August 25th, and September 30th, 2016. The individual terrestrial lidar scans from positions located around the AOI insured >50% overlap between scans. Individual scans were mosaicked together during post-processing in Riegl's RiScan Pro software to comprise a total scanned area of ~10 ha with a focus area of interest of ~2 ha (Riegl, 2017). The individual terrestrial lidar scans were aligned using 10 cm diameter reflective cylinder targets mounted on 2 m tall poles. Reflective targets for each scan were centered on rebar driven into the ground and their locations were measured with a survey-grade Trimble GPS system in UTM coordinates tied to existing Coast and Geodetic Survey benchmarks on the watershed (Trimble, 2013). The Trimble R10 (Trimble, 2012) had a measurement error atop ground control points (GCP) of ~5–7 mm in the horizontal and ~6–16 mm in the vertical dimension. At the SRER, stable rebar GCPs were established at 13 locations and 17 cm diameter round plastic lids painted black & white in a cross pattern were temporarily fixed to the rebar rods. We surveyed the GCPs with a Real-Time Kinematic GNSS on October 7, 2016.

Manned aerial lidar scanning (ALS) data was collected by Woolpert Inc. over all of WGEW (~50 km²) on September 16–18, 2015. The lidar was georeferenced and geometrically corrected using a Real-Time Kinematic GNSS survey as well as a Rapid-Static GPS survey (see Supplemental Materials). Manned aerial lidar were gathered at the SRER in April 2011 by Sanborn Inc.. The SRER data was georeferenced and geometrically corrected using GNSS Survey (see Supplemental Materials). The positional accuracy of the manned aerial lidar given by the

vendors was 6.6–8.0 cm RMSE_z for bare earth, and 36–50 cm RMSE_z for vegetation (Table 1), meeting the USGS quality level 1 standard (Heidemann, 2012).

An octocopter sUAS (Service-Drone, Germany) equipped with Velodyne-32 lidar sensor was deployed at the WGEW in October, 2016 (Velodyne Acoustics, Inc., Morgan Hill, CA) (Table 1). The octocopter sUAS weighs 5.5 kg and was developed to carry an additional heavy payload of up to 6.5 kg, for a maximum takeoff weight of 12 kg (Sankey et al., 2017a,b). The lidar data were mosaicked together in ENVI v. 5.3 (Exelis Visual Information Solutions, Boulder CO, 2010).

Characteristics of the cameras used for SfM varied (Table 2). Collections from hand held photography focused on either specific individual plants (grasses, cacti, mesquite) or along 30 m transect lines within the SRER UA-C pasture plots. The DJI Osmo was held on a 1 m range pole overhead with the camera pointed approximately nadir and 20–40° oblique. Both still photos and UHD (4k) video were recorded along transects. Individual frames were extracted from the videos using ffmpeg (FFmpeg Developers, 2016). The rate at which frames were extracted from video varied, given the forward speed of the collection (~1–2 m/s). Typically, 3–6 frames per second were enough to ensure 80% overlap between frames (4k videos were shot at 24 and 30 frames per second). Quadcopter sUAS were flown over the SRER in summer 2016 using the DJI Phantom 3 and Phantom 4, each weighing ~1.5 kg (Gillan et al., 2017). A third party software Altizure for DJI (Everest Innovation Technology, 2017), was used

TABLE 2 | Structure from Motion, platforms, sensors, image size [megapixels (M)], estimated Ground Sample Distance (GSD) pixel size, and SfM photogrammetry points per square meter (ppsm), and the sites of deployments: Santa Rita Experimental Range (SRER) and Walnut Gulch Experimental Watershed (WGEW).

Platform	Sensor	Image size	GSD (mm)	SfM ppsm	Site
Handheld Sony	Sony Exmor a6000	24.3 M	<4	>100,000	SRER
Handheld DJI Osmo	DJI Micro 4/3	16 M	<4	>100,000	SRER
sUAS DJI Phantom 3/4	Sony Exmor	12.4 M	10	5,000–34,000	SRER
sUAS FireFLY6	Sony Exmor a6000	24.3 M	20	1,000–4,000	SRER
sUAS Ebee	SensFly multiSPEC 4C	1.2 M (4 bands)	150	5–140	WGEW

TABLE 1 | Lidar platforms (TLS-terrestrial lidar scanner, ALS-Aerial lidar scanner, sUAS small Unmanned Airborne System), sensors, collection pulse rate in kilohertz (kHz), laser divergence (mrad), height above ground level (AGL), pulse size (centimeters [cm]), root mean square error (RMSE) in the vertical (z) and horizontal (h) plane (cm), laser pulse return density per meter square (ppsm), and the sites of deployments: Santa Rita Experimental Range (SRER) and Walnut Gulch Experimental Watershed (WGEW).

Platform	Sensor	Pulse Rate (kHz)	mrad (@1/e)	AGL (m)	Pulse size (cm)	RMSE _z (cm)	RMSE _h (cm)	ppsm	Sites
TLS	Riegl VZ-400	1200	0.3	2	~0.2	±1	±1	1 to 110,000	WGEW, SRER
ALS	Leica ALS-50	48–104	0.15	900	~13.5	±6.4	±50	8 to 25	SRER
ALS	Leica ALS-70	500	0.15	900	~13.5	±8.0	±36	8 to 34	WGEW
sUAS	Velodyne-32	700	3.0 (h), 1.2 (v)	80	~3.0	±89–169	±55	35 to 115	WGEW

to program the flight missions with 80% forward and side overlap of the imagery.

The fixed-wing FireFLY6 flights were performed at the SRER UA-C site on March 17, 2016 and June 28, 2016 at an average flight altitude of 110 m AGL with 65 and 75% forward and side overlap between image tiles. Most of the aerial images were collected facing approximately nadir because the camera is fixed to the belly of the FireFLY6. Pitch, yaw, and roll during the flights resulting in images 0°–15° degrees off nadir.

The eBee is a fixed wing (140 cm wingspan, 750 g mission weight), electric platform with a single pusher propeller at the rear (SenseFly, Switzerland). Flights were performed at an average altitude of 110 m AGL with 80% forward overlap between successive image and 70% side overlap respectively between adjacent flight lines at Kendall grassland site on October 8, 2016. The multispectral Sensfly multiSpec 4C camera records four bands: green (520–580 nm), red (630–690 nm), red edge (720–750 nm), and NIR (760–820 nm). The geo-located, spatially paired, and radiometrically calibrated images were combined to generate an orthomosaic and DSM of the entire flight area in Pix4D (SenseFly, Switzerland). Separately a 3D point cloud was generated for each band and then merged together to generate a single point cloud dataset. The eBee sUAS multispectral images were processed in eMotion and Pix4D (SenseFly, Switzerland) and in Photoscan (AgiSoft, 2017).

Post-hoc Point Cloud Registration

Comparing two or more point clouds required alignment using at least three GCPs shared between datasets. In the case of our sUAS plots, 10–13 GCPs were needed to ensure high enough accuracy (<5 cm) to successfully detect vegetation change over time (Gillan et al., 2017).

We did not establish GCPs for the eBee SfM photogrammetry in WGEW or the FireFLY6 in SRER. Features that were clearly identifiable in the terrestrial lidar or manned aerial lidar data, e.g., eddy-covariance flux towers, water sampling flumes, fence posts, road features, and small boulders were used *post-hoc* to align point clouds in CloudCompare (Girardeau-Montaut, 2011) with the translation/rotation tool and three-point-picking tool.

Detection Bias

For bare earth analyses locations with barren ground that were clearly observable by the terrestrial lidar were used as references, i.e., roadways and bare patches with no herbaceous component. For the vegetation analyses locations scanned with the terrestrial lidar with no obstruction of the bare earth at the base of the reference plant up to its apical leader were used as references. Terrestrial lidar reference point cloud data were not used in areas with high incidence of occlusion and shadowing due to increasing distance from the scanner.

Raster Differencing

We produced DEM of Difference (DoD) (Lane et al., 2003; Pelletier and Orem, 2014) for the minimum elevation DEMs and maxima elevation DSMs. DEMs were generated by “cloth simulation filter (CSF)” (Zhang et al., 2016) and “Rasterization” tools in CloudCompare (Girardeau-Montaut, 2011). DEMs were

generated using the minimum height function, or bare earth classified points (Class 2, Heidemann, 2012), in the case of the aerial lidar. The DSMs use a maximum height function after outlier points are removed. DEM and DSM rasters were produced at 0.5 m resolution; this was the finest possible resolution for comparison to the aerial manned lidar given its ground return point density. DEM and DSM were exported to QGIS as geotiff (.tif) rasters in their original datum (Table S2). In QGIS we used the raster calculator to generate (a) bare earth DoD₁: DEM₁–DEM₂ (Equation 1); and (b) vegetation DoD₂: DSM₁–DSM₂ (Equation 2) for the different point cloud types. The DoD assess bias within the AOI by establishing the absolute difference between each platform; however DoD differences alone did not intrinsically identify which platform as the source of the bias.

Cloud-to-Cloud Differencing

Lague et al. (2013) created the direct point cloud-to-cloud “multi-scale model to model cloud comparison” (M3C2) tool to compare change across time in point cloud data. The M3C2 is successful at measuring surface and volume differences in geomorphological applications (James et al., 2017; Midgley and Tonkin, 2017). M3C2 computes local distances between “core points” from a “reference” cloud (Lague et al., 2013), we used the terrestrial lidar as the reference cloud and the other clouds as the “test” for differencing. M3C2 uses a “normal” flat surface to estimate the confidence interval based on cloud roughness and registration error (Lague et al., 2013). The difference between two points can be measured in the vertical (z) or horizontal (x and y) axis, or as a linear multi-scale distance based on the normal. In our analysis we compared the vertical (z) change. The significance of the measure is based on the registration error for each core point. Our registration error used the instrument precision of the lidar (Table 1) and the uncertainty of the RTK GNSS (Table S3).

Mesh Creation and Object Height Measurement

To improve the estimate of vegetation height in a fused data comparison (Objective 2), we used a cloth simulation filter (CSF) technique (Zhang et al., 2016) in CloudCompare (Girardeau-Montaut, 2011), to calculate a minimum elevation (bare earth) model as a mesh. One benefit of the mesh was that it fills space beneath large areas of vegetation where no bare earth surface is remotely sensed. We also used CloudCompare’s rasterization tool to generate maximum (DSM) models for object heights. To measure the height of points above the mesh bare earth level we used CloudCompare’s Mesh-to-Cloud function.

RESULTS

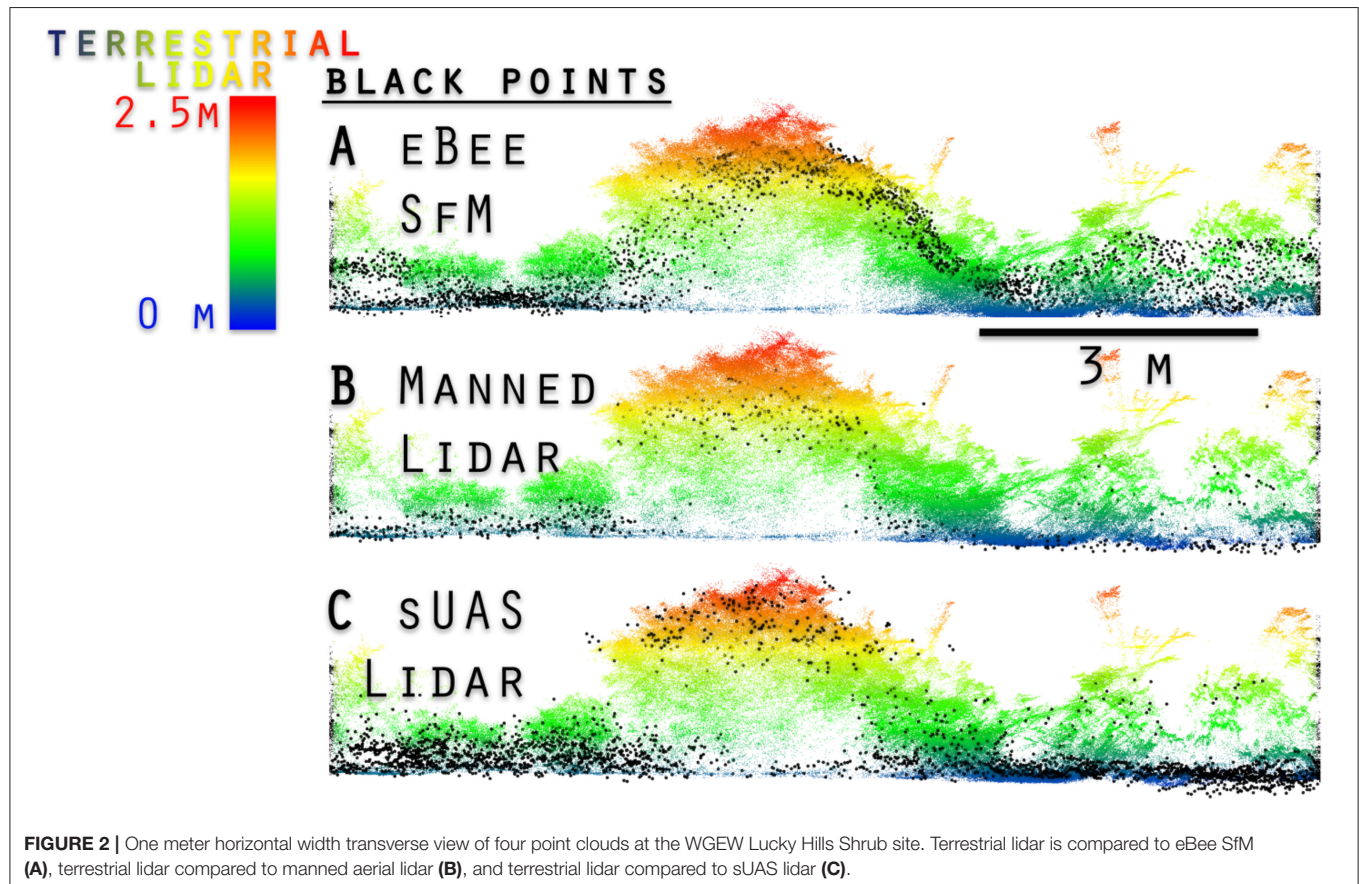
Feature Class Detection Bias

Bare earth elevation is only clearly resolved by the terrestrial lidar in barren areas or along roads. The propagated uncertainty (Supplementary Materials) of any observation of bare earth derived from manned aerial lidar to a terrestrial lidar fused measurement was ±8.24 to 9.4 cm RMSE_z (Table 3). Aerial lidar (manned and sUAS) was found to be best at detecting bare ground, in part because of its unique point of view, penetrating

TABLE 3 | Vertical (z) differences in point clouds by platform using M3C2 with reference cloud uncertainty of 5 cm; mean \pm 99% CI observed difference (cm).

	Kendall grassland			Santa Rita mesquite		
	ALS	sUAS lidar	sUAS SfM [†]	ALS	sUAS SfM [‡]	Hand-held SfM [§]
Bare ~0 m	2 \pm 22	2 \pm 73	6 \pm 24	5 \pm 22	2 \pm 16	0 \pm 6
Grasses < 1 m	-25 \pm 51	0 \pm 25	-4 \pm 52	n/a	-12 \pm 25	-15 \pm 15
Shrubs < 2 m	-15 \pm 26	-12 \pm 25	-16 \pm 25	-16 \pm 21	-1 \pm 32	0 \pm 5
Shrubs > 2 m	-31 \pm 169	-10 \pm 14	-25 \pm 13	-18 \pm 25	-2 \pm 52	-4 \pm 13

Negative mean values represent systematic under measurement by the test cloud relative to the terrestrial lidar reference cloud. Positive values are associated with the reference cloud failing to observe elevations at depth, e.g., no returns beneath dense vegetation. n/a value are for no temporal alignment for comparison (Santa Rita). The SfM point clouds derived from: [†]Sensefly eBee, [‡]DJI Phantom 4, [§]DJI Osmo.



down through the vegetation. Conversely, the sUAS fails to detect bare ground beneath grass and shrub vegetation (Figure 2), in part due to the GSD of the imagery used to generate it. However, ALS did not always penetrate the canopies of dense leaf-on woody shrubs > 2 m tall, nor does it detect sense the fine branches of the shrubs or trees (Table 3). The ALS does resolve larger features on tall woody vegetation (> 2 m) with typical underestimates of 5–25 cm (Table 3).

The octocopter sUAS lidar had lower positional certainty than the manned aerial lidar, due in part to the sUAS lidar having lower quality GNSS (50+ cm RMSE₂), and unresolved GPS-IMU drift aboard the octocopter platform (Sankey et al., 2017a,b). When the sUAS lidar point cloud was adjusted using fine alignment tools in

CloudCompare its features were comparable to the manned aerial lidar (Figure 2, Table 3). The sUAS lidar resolved both grasses and bare surface in herbaceous and shrubs in the Kendall site, while the manned aerial lidar does not. Despite the increased point density relative to the manned aerial lidar (8–24 ppsm) the sUAS lidar (35–115 ppsm) was still too sparse to identify individual herbaceous plants or shrubs based on only the point cloud.

The Ebee SfM failed at sensing bare earth elevation for areas with dense vegetation on the WGEW. This was not surprising at Kendall Grassland where the herbaceous cover was high. The eBee SfM also tended to under-predict height of woody vegetation in the same way as the manned aerial lidar: it failed

to discriminate diffuse outer branches and maximum tree height (**Figure 2, Table 3**). The initial GNSS referenced eBee SfM point cloud relative to the RTK referenced terrestrial lidar was not accurate and required realignment using the 3-point-picking and fine alignment tools in CloudCompare.

The Phantom 3 and Phantom 4 sUAS SfM photogrammetry also produced point clouds suitable for discriminating large woody vegetation or herbaceous vegetation. However, the resultant SfM point clouds were less accurate than the terrestrial lidar, under-estimating grass height on average by 15 and 4 cm for the larger woody shrubs (**Table 3**). SfM photogrammetry based on the FireFLY6 imagery was not able to accurately resolve tall woody vegetation (mesquite trees taller than 2 m) on the SRER during the spring leaf-off period, but was improved during the leaf on period in late summer (Supplementary Materials).

DEM/DSMs of Difference for Detection Bias

Our objective #1 was to establish detection bias between point cloud data and derivatives. In the case of raster data generated from point clouds we establish the difference between the same type of measurement for two different platforms, i.e., $DEM_1 - DEM_2$ (DoD₁) & $DSM_1 - DSM_2$ (DoD₂).

The DoD₁ for manned aerial lidar (**Figure 3A**) vs. terrestrial lidar (**Figure 3B**) are ± 8.24 cm RMSE in areas immediately surrounding the terrestrial laser scan locations in the AOI (**Figure 3D**). Bare earth elevations are >100 cm different, where bare ground was not visible to the terrestrial lidar sensor (**Figure 3D**). In these locations the manned aerial lidar and SfM

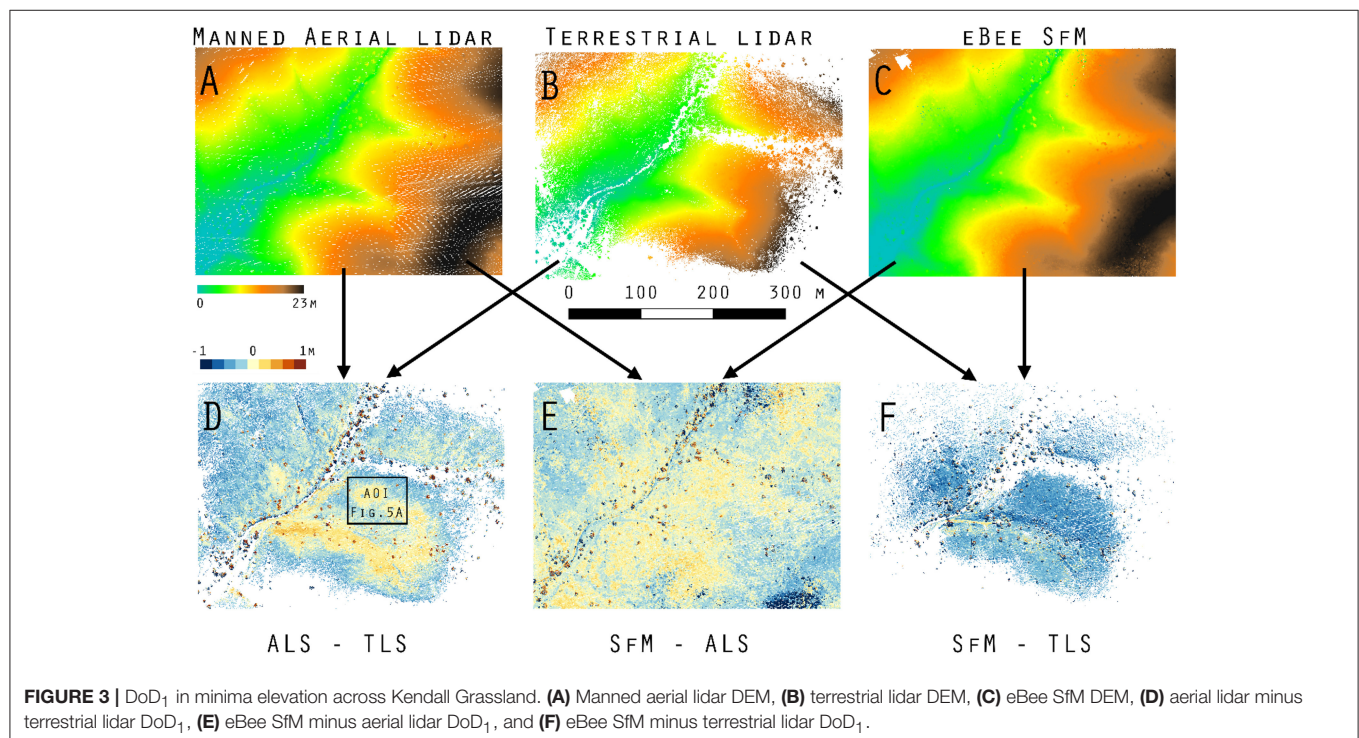
(**Figure 3E**) are lower (cooler colors) than the terrestrial lidar estimate of bare ground (**Figure 3F**).

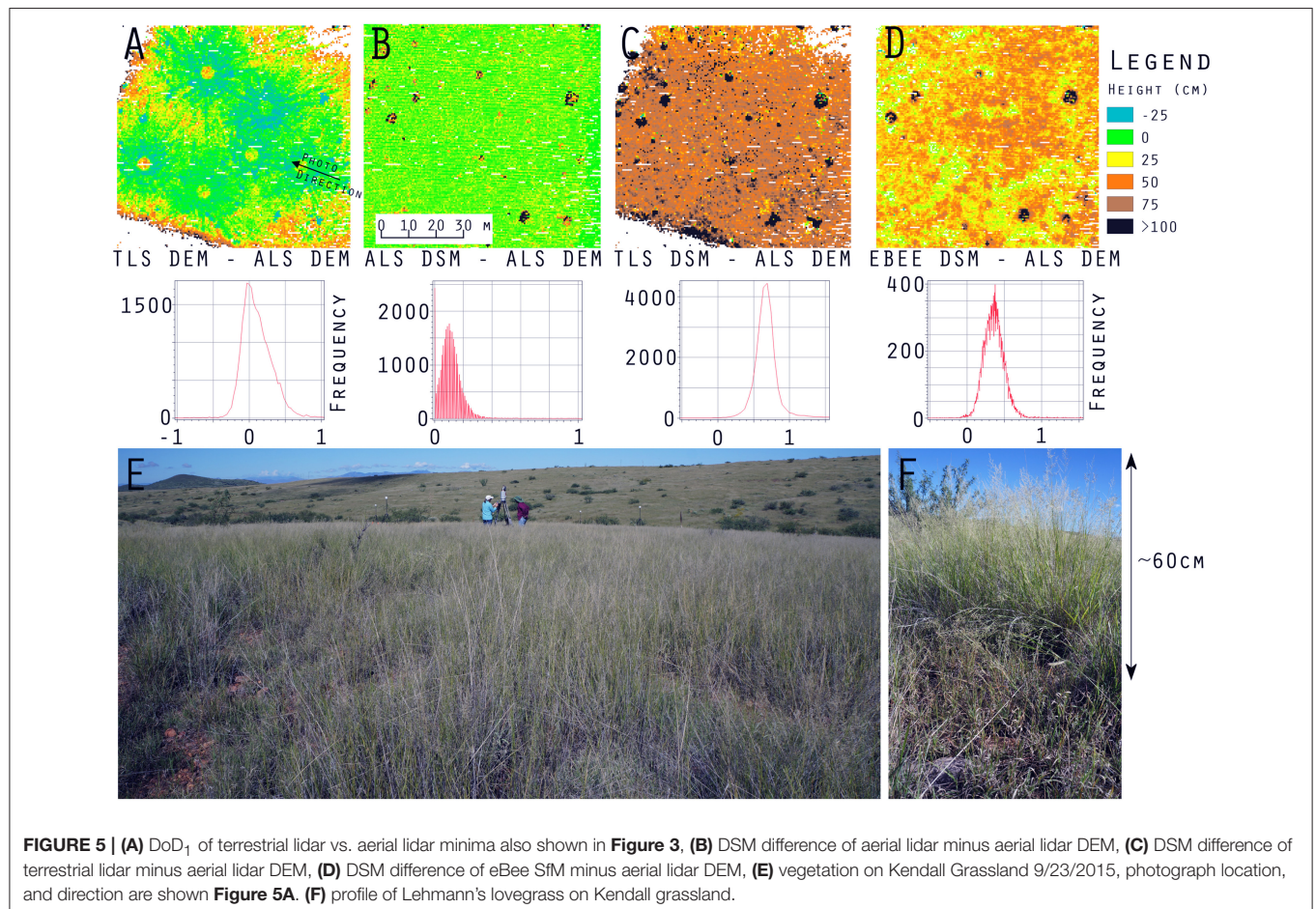
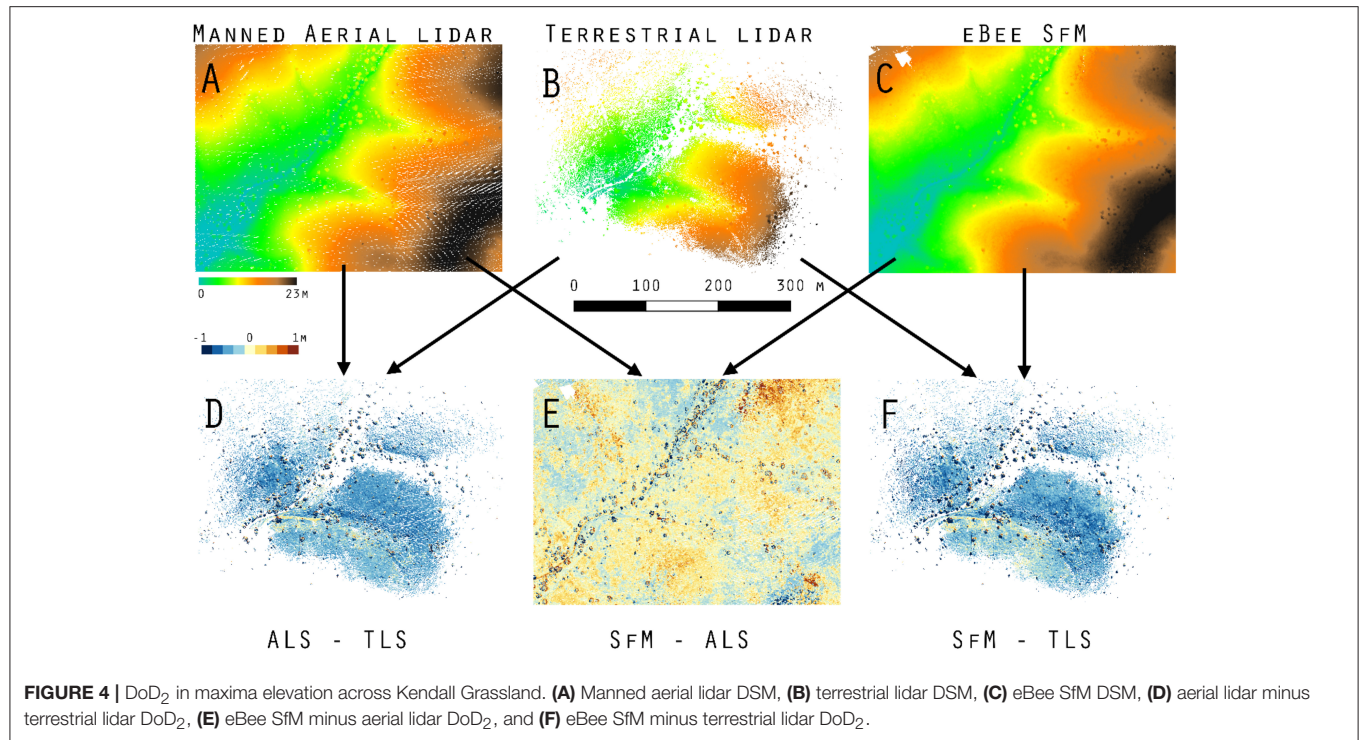
The aerial lidar DSM (**Figure 4A**) under estimate height of grass and woody shrubs by 10 to 120 cm when subtracted from the terrestrial lidar DSM (**Figure 4D**). There was also a consistent difference across both the SfM-ALS DoD₂ (**Figure 4E**), due to the failure of the aerial lidar to sense the herbaceous component. There was an under prediction of height for SfM-TLS DoD₂ (**Figure 4F**), with the eBee SfM DSM typically 4 to 52 cm lower than the terrestrial lidar DSM.

Data Fusion Accuracy

The area immediately around the terrestrial lidar scanner had the most similar bare elevation as aerial lidar DEM (**Figure 5A**). The associated histogram below **Figure 5A** shows a mean of approximately zero with standard deviation of 25 cm—similar to the results shown in **Table 3** for the M3C2. The weakness of the aerial lidar in measuring vegetation is apparent in **Figure 5B** where the average height of the point cloud is only 10 cm in grass—an approximate 55 cm under measurement relative to the terrestrial lidar in the same locations (**Table 4**), visible in **Figure 5C**. The eBee SfM produced DSM was a more accurate representation of grass height (**Figure 5D**) when using the aerial lidar derived DEM, and was only 24 cm lower on average than the terrestrial lidar scans (**Table 4**).

Using the mesh-to-cloud height measure in CloudCompare we produced a point cloud based height value for the terrestrial lidar (**Figure 6**). The bare earth mesh was derived from the CSF (Zhang et al., 2016) (Supplementary Materials). Our anecdotal example, shown in **Figure 6**, demonstrates one potential method





for improving the characterization of vegetation through two platform data fusion: terrestrial lidar and SfM point clouds measure vegetation height above ground but fail to sense bare earth. When they incorporate an aerial lidar point cloud derived bare earth mesh surface the height measure improves. Note in **Figure 6** the Ebee SfM points are almost entirely along the top of the grasses and woody shrub (*Yucca elata*, soaptree yucca) as scanned by the terrestrial lidar, while the larger diameter black points from the aerial lidar have few returns in the vertical profile except for the flowering stalk of the yucca. Also note, terrestrial lidar has few or no points low in the vertical profile near ground level. This was due to the flat angle of the horizontal scan profile being obscured by dense herbaceous cover.

DISCUSSION

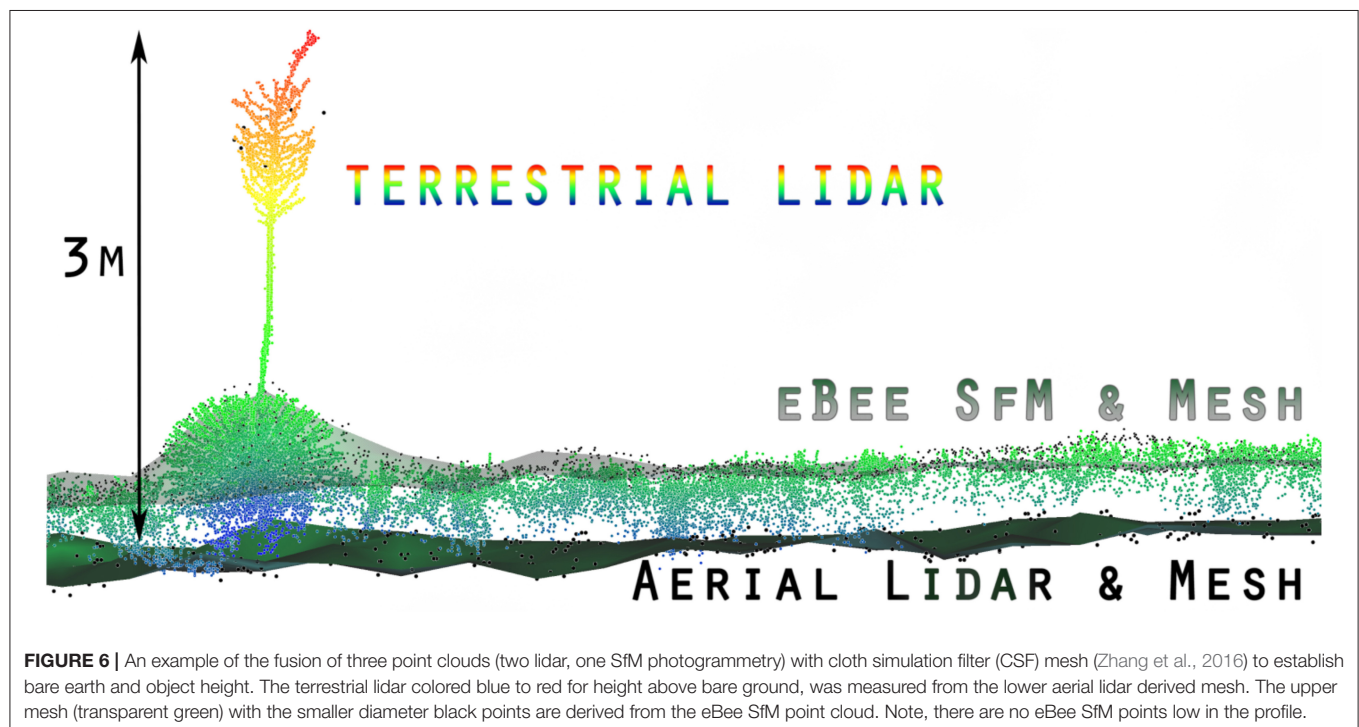
No single observation platform solved Levin's "problem of pattern and scale" because each had its own limitations in

(1) detecting patterns of important objects in the dryland ecosystem (bare ground, grass, and shrubs), (2) the accuracy of representing the heights of grass and shrubs, and (3) the extent of area represented in a typical data collection campaign. Nonetheless, the combination or fusion of information from the different platforms appears to make greater progress toward solving "problems of pattern and scale" than any single platform alone. Specifically, we suggest the combination of sUAS SfM and either manned or sUAS aerial lidar as providing the best solution at large spatial scales amongst the platforms studied.

To represent spatial and temporal patterns of vegetation, sUAS SfM provides a high quality representation of the presence of grass and shrub vegetation at centimeter scale, but fails to reliably detect bare earth or DEM beneath that vegetation, thus limiting the ability to represent the height and volume of those plants. Aerial lidar on the other hand, provides that needed high quality representation of the pattern of bare earth or DEM. For scale, manned aerial lidar provides spatial resolutions of 0.25–1.0 m², and sUAS Velodyne lidar spatial resolution of 0.1–0.25 m² which is slightly coarser than the terrestrial lidar or sUAS SfM. However, from the perspective of scale of extent or coverage, most aerial lidar data collection campaigns represent hundreds to thousands of km², and there is less need to frequently repeat those campaigns because DEM change in dryland ecosystems are less frequent than the height and volume of grass and shrub vegetation. On the other hand, the extent of a sUAS SfM campaign is at least two orders of magnitude smaller than an aerial lidar campaign, but it is easier and cheaper to repeat several times per year in specific areas of interest, e.g., eddy-covariance flux tower footprints.

TABLE 4 | Average (mean \pm 95% CI) object height measurements (cm) taken on the Kendall Grassland site; examples of the aerial lidar (ALS) height model is shown in **Figure 5B**, the terrestrial lidar DSM minus the aerial lidar DEM (TLS-ALS) in **Figure 5C**, and eBee DSM minus aerial lidar DEM (SfM-ALS) in **Figure 5D**.

Feature classes	SfM	ALS	TLS-ALS	SfM-ALS
Bare ~0m	5 \pm 14	2 \pm 16	-1 \pm 8	6 \pm 8
Grasses < 1 m	20 \pm 16	10 \pm 12	65 \pm 26	41 \pm 42
Shrubs < 2 m	114 \pm 112	55 \pm 124	178 \pm 50	87 \pm 45
Shrubs > 2 m	27 \pm 278	225 \pm 24	252 \pm 46	177 \pm 215



Point Cloud Bias amongst Platforms

The availability of high density lidar point clouds, particularly terrestrial lidar, allowed us to carefully examine differences amongst sUAS lidar, manned aerial lidar, and SfM photogrammetry point clouds. We found the manned aerial lidar, sUAS lidar, and sUAS SfM data all tended to underestimate herbaceous vegetation height in the grassland and shrub areas.

The accuracy of the terrestrial lidar measurements are over an order of magnitude finer than the manned aerial lidar. When comparing the various point clouds to the terrestrial lidar data it was apparent large biases (positional shifts) had been introduced by the various GNSS for both the aerial lidar and SfM photogrammetry. It was also apparent that terrestrial lidar had difficulty resolving bare earth elevations along flat trajectories further away from the scan locations when dense herbaceous vegetation was present (Figures 5, 6). The sUAS lidar was able to penetrate the grass to bare earth elevations and characterize the grass vertical profile, similar to that observed in the terrestrial lidar, but was unable to segment individual plants due to lower density. The manned aerial lidar was not suitable for measuring or monitoring herbaceous grass and small shrubs, as highlighted in Figure 5B. Most of the WGEW manned aerial lidar data are single return at the near ground level in herbaceous grasses (Figure 6). It is likely the ~13.5 cm footprint of the aerial lidar meant smaller features such as grass panicles and small diameter woody stems were not large enough to be detected or were not at the threshold level for discretization from the full waveform data. The largest woody plants, e.g., leaf-on mesquites were detected, but their diffuse branches were not, resulting in an overall 10 to 25 cm height under-estimate of tree height relative to the terrestrial lidar (Figure 5B, Table 4). These results are similar to Luscombe et al. (2014), who compared manned aerial lidar and terrestrial lidar and found similar under-estimates in height for taller vegetation.

In bare ground areas, sUAS SfM data were equivalent to aerial lidar and terrestrial lidar, which is consistent with other recent studies (Lucieer et al., 2012; Westoby et al., 2012; Anderson and Gaston, 2013; Fonstad et al., 2013; James et al., 2017). In dense herbaceous cover SfM photogrammetry point clouds failed to accurately resolve bare ground elevations beneath vegetation accurately. Error in sUAS SfM point clouds are attributed in parts to (1) barrel or pin cushion lens distortion (James and Robson, 2014), and (2) algorithm error when interpreting between bare earth and vegetation (James et al., 2017).

Influence of Canopy Cover on DEM generation

We found the sUAS SfM point cloud M3C2 values (Table 3) were all within the uncertainty of the aerial lidar for bare ground (± 9.6 cm RMSE_z) and large woody vegetation (± 18 cm RMSE_z). In the case of the sUAS Phantoms and Osmo SfM, values are within the range of uncertainty of the terrestrial lidar for bare ground (Table S3). Our results were essentially the same as Nouwakpo et al. (2015), who compared soil texture and elevation differences between ground-based SfM and terrestrial lidar and

were able to stay within 5 mm RMSE for bare earth patches but not in areas of vegetation with >50% canopy cover.

The sUAS SfM (eBee, Phantoms, Firefly6) data could not be used to generate bare earth elevations beneath herbaceous vegetation (grasses and low shrubs) with any confidence. Another problem in densely vegetated herbaceous sites such as the WGEW Kendall Grassland and the SRER UA-C, occurred when attempting to generate bare earth models with imagery from time periods of vegetation senescence. The resultant models were more likely to be biased because the senescent vegetation cannot be differentiated from soil, such as during the dry season (e.g., the March 2016 Firefly6 flights at UA-C). Standard attempts at removing vegetation using a filter or masking technique, such as NDVI (Cunliffe et al., 2016; James et al., 2017), can only be executed during the leaf-on rainy season in these areas when canopy cover is highest and is spectrally responsive to NDVI.

Influence of Canopy Cover on DSM Generation

Many recent SfM research papers (Cruzan et al., 2016; Cunliffe et al., 2016; James et al., 2017) report on how SfM does very well in areas with low vegetation cover. In our study, the sUAS SfM was not as accurate at sensing the maximum herbaceous vegetation height as the terrestrial lidar, rather the SfM tended to under predict max heights of the fine herbaceous and woody component. However, SfM at high resolution in hand-held DJI Osmo or low flying Phantoms, was capable of resolving fine herbaceous features and woody stems (Supplementary Materials). Kato et al. (2015) compared SfM vs. terrestrial lidar for tree canopy structure and reported a strong correlation between canopy shape for both technologies. We found similar patterns where the sUAS and handheld SfM reconstruct the shape of large woody canopy shrubs, but are unable to penetrate deeply into the canopy of individual trees to reveal their internal structure.

Raster Differencing (DoD) vs. Cloud-to-Cloud Comparison (M3C2)

The DoD show vertical change in both the bare elevation and maximum height of vegetation in the two surface models (Figures 4, 5); meanwhile the M3C2 was able to diagnose changes in multiscale and horizontal distance both beneath vegetation and across canopy (Table S3). M3C2 was able to more clearly diagnose errors of horizontal change than the DoD which must be inferred by looking at rasters for distinct vertical changes around the vegetation.

Both methods have utility: DoD rasters are useful in eco-hydrology applications, while the M3C2 maybe more useful for geomorphology (bare elevation change), and monitoring of vegetation dynamics (growth, mortality, abundance and composition). Notably, the computational overhead required for M3C2 is orders of magnitude greater than for raster based modeling but is reduced by the use of GPU computing.

Data Management

Our large data sets created a computing challenge when comparing the largest point clouds (>50 GB) in a single computing instance. Cyberinfrastructure storage, such as the

CyVerse.org DataCommons (<http://datacommons.cyverse.org>, Walls, 2017), the NSF lidar repository OpenTopography.org, and USGS National Map are available to users who wish to publically archive processed point clouds. The original images (.JPG or .TIF files) generated by sUAS or handheld collection should be archived at the highest resolution possible, with associated time of flight mission metadata, i.e., GNSS location and camera specifications. One benefit of retaining the imagery for SfM photogrammetry over delivered (i.e., processed discrete) lidar data are their availability for reanalysis with new versions of SfM software which have improved optimization techniques as well as camera and lens correction algorithms. Lidar data, both discrete and full waveform, available in their raw (time-of-flight) format should similarly be stored for future reanalysis.

To help with processing we used CyVerse Atmosphere and NSF Jetstream (Stewart et al., 2015) virtual machine instances. We found the largest instances were ideal for large distributed jobs running point cloud analyses (without GPU). For users who are not able to purchase workstation class PCs, access to cloud or HPC solutions are ideal alternatives at a fraction of the initial start-up cost. We expect the availability of cloud computing and cyberinfrastructure to become more common as computing technology further develops. At the same time, we expect the volume of high-resolution 3D data to increase, as sUAS technology becomes more widely used by geoscientists and ecologists.

Fusion of point cloud data from different sensors and formats presents a problem for ensuring data standards and metadata attribution continue through the life cycle of data. Currently, lidar point cloud data are managed under the LAS standard developed from the American Society for Photogrammetry and Remote Sensing (ASPRS). SfM point clouds, stored as LAS/LAZ files, have fundamental differences from lidar LAS/LAZ, including the lack of scan angle, intensity, or return number.

Costs (Time, Human, Computing)

The SfM point cloud generation became prohibitive with very large image collections (250–5,000 high quality images) requiring from 12 to 192 h to generate the ultra-high resolution point clouds, also maxing out the available RAM on the workstations without chunking the data into tiles and further increasing processing time. In addition, the photogrammetrist typically had to work with individual images to geo-locate ground control points from anywhere between 4 and 16 h for each model. Conversely, the terrestrial lidar required only a few hours of post-processing by the technicians to resolve the ground-control, typically being completed within 24 h of the initial collection. Aerial lidar data require a significant investment, however that step was shouldered by the contracted vendor and was reflected in the cost of the aerial lidar collection.

FUTURE OUTLOOK

The scale of remotely sensed 3D data from hand held platforms, sUAS, and manned aircraft allows us to ask new questions which

25 years ago were unassailable for Levin (1992). However, we must recognize the strengths and weaknesses inherent to the technology and the data when applying them to the measurement of ecological features. We found aerial lidar from both sUAS and manned aircraft to be more accurate at sensing bare ground elevation than terrestrial lidar or sUAS SfM in dense vegetation, but less accurate at measuring fine and woody herbaceous vegetation. By using the aerial lidar to derive an accurate bare earth model and a second technology (terrestrial lidar, sUAS SfM photogrammetry) to measure herbaceous vegetation height and structure, we were able to better characterize vegetation than by any single technology alone. Quantifying propagation of uncertainty when fusing or combining point cloud (Supplementary Materials) or derivative data is likely to be one of the most important steps researchers take when attempting to make ecologically important measurements.

Newer constellations of “dove” satellites which form “flocks” and record the entire global surface at weekly to daily time scale solve a temporal limitation to remote sensing at scale, with nominal spatial resolution (1–3 m) similar to aerial orthophotography (Butler, 2014; McCabe et al., 2016; Zimmerman et al., 2017). In the United States, nationally available lidar data sets (Stoker et al., 2016) provide bare earth models at 1–2 m resolution. Our methods show how these data could be used for more accurately measuring dense vegetation cover, height, and volume from sUAS derived lidar or SfM photogrammetry data and scaling out spatially with satellite products. Despite the utility of sUAS for monitoring vegetation phenology and structure at orders of magnitude higher resolution, the overhead of computational processing of these imagery can become a limiting factor at progressively larger spatial and temporal resolution, suggesting the need for distributed HPC and cloud based solutions like those provided by Stewart et al. (2015).

Keeping pace with innovation and the breadth of remotely sensed information available for conducting ecological research can be challenging for any single field ecologist, particularly for those looking to add new skills or maintain technical proficiency. Establishing where the limitations of these new remote sensing technology are was one goal of this research. We hope to have revealed a few useful, scalable solutions to these problems through data fusion, and provided insight into improved monitoring programs in dryland ecosystems.

AUTHOR CONTRIBUTIONS

TLS, TTS, PH, MN, and MM conceptualized and designed the research ideas. TLS, TTS, JG, MN, and JM acquired and analyzed the data. TLS wrote the draft with contributions from TTS, PH, MN, MM, and JG.

ACKNOWLEDGMENTS

The Southwest Watershed Research Center USDA-ARS, provided the terrestrial lidar equipment used in this study. We also thank two reviewers for improving an earlier versions of the manuscript.

This material is based upon work supported by the U.S. Department of Agriculture, Agricultural Research Service, under Agreement No. 58-2022-5-13. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

This material is based upon work supported by the National Science Foundation under Award Numbers DBI-0735191 and

DBI-1265383, www.cyverse.org; and Jetstream (Award number ACI-144506) cloud resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2017.02144/full#supplementary-material>

REFERENCES

- AgiSoft (2017). *AgiSoft PhotoScan. Professional Edn.* AgiSoft LLC, St. Petersburg, Russia.
- Anderson, K., and Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Ecol. Environ.* 11, 138–146. doi: 10.1890/120150
- Brekenfeld, D. J., and Robinett, D. (1997). *Soil and Range Resource Inventory of the Santa Rita Experimental Range*. Tucson, AZ: Natural Resources Conservation Service. Available online at: <https://cals.arizona.edu/srer/maps/BR1997.pdf>
- Boettiger, C. (2015). An introduction to docker for reproducible research. *ACM SIGOPS Operat. Syst. Rev.* 49, 71–79. doi: 10.1145/2723872.2723882
- Boulton, G., Rawlins, M., Vallance, P., and Walport, M. (2011). Science as a public enterprise: the case for open data. *Lancet* 377, 1633–1635. doi: 10.1016/S0140-6736(11)60647-8
- Butler, D. (2014). Many eyes on earth. *Nature* 505, 143–144. doi: 10.1038/505143a
- Carrivick, J. L., Smith, M. W., and Quincey, D. J. (2016). *Structure from Motion in the Geosciences*. Hoboken, NJ: John Wiley & Sons.
- Cruzan, M. B., Weinstein, B. G., Grasty, M. R., Kohn, B. F., Hendrickson, E. C., Arredondo, T. M., et al. (2016). Small unmanned aerial vehicles (micro-UAVs, Drones) in plant ecology. *Appl. Plant Sci.* 4:1600041. doi: 10.3732/apps.1600041
- Cunliffe, A. M., Brazier, R. E., and Anderson, K. (2016). Ultra-Fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry. *Remote Sens. Environ.* 183, 129–143. doi: 10.1016/j.rse.2016.05.019
- Dandois, J. P., and Ellis, E. C. (2010). Remote sensing of vegetation structure using computer vision. *Remote Sens.* 2, 1157–1176. doi: 10.3390/rs2041157
- Dandios, J. P., and Ellis, E. C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sens. Environ.* 136, 259–276. doi: 10.1016/j.rse.2013.04.005
- Docker Development Team (2017). *Version 17.03*. Available online at: <https://www.docker.com/>
- Everest Innovation Technology (2017). *Altizure for DJI App*. Kowloon: Everest Innovation Technology. Available online at: <https://next.altizure.com/>
- Exelis Visual Information Solutions (2010). *ENVI v. 5.3*. Boulder, CO.
- FFmpeg Developers (2016). *ffmpeg Tool (Version be1d324) [Software]*. Available online at: <http://ffmpeg.org/>
- Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L., and Carboneau, P. E. (2013). Topographic structure from motion: a new development in photogrammetric measurement. *Earth Surf. Processes Landforms* 38, 421–430. doi: 10.1002/esp.3366
- Gillan, J. K., McClaran, M. P., Swetnam, T. L., Heilman, P., and Turner, R. (2017). “Estimating forage utilization with unmanned aerial imagery,” in *Poster Presented at Research Insights in Semi-arid Ecosystems (RISE)* (Tucson, AZ). Available online at: <http://www.tucson.ars.ag.gov/RISE/Posters/GillanPoster.pdf>
- Girardeau-Montaut, D. (2011). *CloudCompare-Open Source Project*. Grenoble: OpenSource Project. Available online at: <http://www.cloudcompare.org/>
- Glennie, C. L., Carter, W. E., Shrestha, R. L., and Dietrich, W. E. (2013). Geodetic imaging with airborne LiDAR: the earth's surface revealed. *Rep. Prog. Phys.* 76:086801. doi: 10.1088/0034-4885/76/8/086801
- Hampton, S. E., Anderson, S. S., Bagby, S. C., Gries, C., Han, X., Hart, E. M., et al. (2015). The tao of open science for ecology. *Ecosphere* 6, 1–13. doi: 10.1890/ES14-00402.1
- Harpold, A. A., Marshall, J. A., Lyon, S. W., Barnhart, T. B., Fisher, B. A., Donovan, M., et al. (2015). Laser vision: lidar as a transformative tool to advance critical zone science. *Hydrol. Earth Syst. Sci. Discuss.* 12, 1017–1058. doi: 10.5194/hessd-12-1017-2015
- Heidemann, H. K. (2012). *Lidar Base Specification*. Techniques and Methods.
- Heilman, P., Nichols, M. H., Goodrich, D. C., Miller, S. N., and Guertin, D. P. (2008). Geographic information systems database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44:W05S11. doi: 10.1029/2006WR005777
- Higgins, M. A., Asner, G. P., Martin, R. E., Knapp, D. E., Anderson, C., Kennedy-Bowdoin, T., et al. (2014). Linking Imaging Spectroscopy and LiDAR with Floristic Composition and Forest Structure in Panama. *Remote Sens. Environ.* 154, (Suppl. C), 358–367. doi: 10.1016/j.rse.2013.09.032
- Hobu Inc. (2017). *PDAL: Point cloud Data Abstraction Library. Release 1.5. Copywrite (c) 2017*. Available online at: <http://www.pdal.io>
- Huenneke, L. F., Clason, D., and Muldavin, E. (2001). Spatial heterogeneity in chihuahuan desert vegetation: implications for sampling methods in semi-arid ecosystems. *J. Arid Environ.* 47, 257–270. doi: 10.1006/jare.2000.0678
- Isenburg, M. (2014). *LAStools - Efficient LiDAR Processing Software (version 141017, unlicensed)*. Available online at: <http://rapidlasso.com/LAStools>
- James, M. R., and Robson, S. (2014). Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surf. Processes Landforms* 39, 1413–1420. doi: 10.1002/esp.3609
- James, M. R., Robson, S., and Smith, M. W. (2017). 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys. *Earth Surf. Processes Landforms* 42, 1769–1788. doi: 10.1002/esp.4125
- Kachamba, D. J., Orka, H. O., Næsset, E., Eid, T., and Gobakken, T. (2017). Influence of plot size on efficiency of biomass estimates in inventories of dry tropical forests assisted by photogrammetric data from an unmanned aircraft system. *Remote Sens.* 9:610. doi: 10.3390/rs9060610
- Kato, A., Obana, H., Hayakawa, Y., Watanabe, M., Yamaguchi, Y., and Enoki, T. (2015). “Fusion between UAV-SFM and terrestrial laser scanner for field validation of satellite remote sensing,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (Milan).
- Keefer, T. O., Moran, M. S., and Paige, G. B. (2008). Long-term meteorological and soil hydrology database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44:W05S07. doi: 10.1029/2006WR005702
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences*. London: SAGE Publishing.
- Lague, D., Brodu, N., and Leroux, J. (2013). Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the rangitikei canyon (N-Z). *ISPRS J. Photogramm. Remote Sens.* 82, 10–26. doi: 10.1016/j.isprsjprs.2013.04.009
- Lane, S. N., Westaway, R. M., and Murray Hicks, D. (2003). Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surf. Processes Landforms* 28, 249–271. doi: 10.1002/esp.483
- Levin, S. A. (1992). The problem of pattern and scale in ecology: the robert H. MacArthur Award Lecture. *Ecology* 73, 1943–1967. doi: 10.2307/1941447
- Lucieer, A., Robinson, S., Turner, D., Harwin, S., and Kelcey, J. (2012). “Using a micro-UAV for ultra-high resolution multi-sensor observations of Antarctic moss beds,” in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIX-B1* (Melbourne), 429–433.
- Luscombe, D. J., Anderson, K., Gatis, N., Wetherelt, A., Grand-Clement, E., and Brazier, R. E. (2014). What does airborne LiDAR really measure in upland ecosystems? *Ecohydrology* 8, 584–594. doi: 10.1002/eco.1527

- McCabe, M. F., Houborg, R., and Lucieer, A. (2016). "High-resolution sensing for precision agriculture: from earth-observing satellites to unmanned aerial vehicles," in *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII* (Edinburgh).
- McClaran, M. P., Browning, D. M., and Huang, C. (2010). "Temporal dynamics and spatial variability in desert grassland vegetation," in *Repeat Photography: Methods and Applications in the Natural Sciences*, eds R. H. Webb, D. E. Boyer, and R. M. Turner (Washington, DC: Island Press), 145–166.
- McClaran, M. P., Ffolliott, P. F., and Edminster, C. B. (2003). "Santa Rita experimental range—100 years (1903 to 2003) of accomplishments and contributions," in *Conference Proceedings* (Tucson, AZ).
- McClaran, M. P., and Wei, H. (2014). Recent drought phase in a 73-year record at two spatial scales: implications for livestock production on rangelands in the Southwestern United States. *Agric. For. Meteorol.* 197, 40–51. doi: 10.1016/j.agrformet.2014.06.004
- Merkel, D. (2014). Docker: lightweight Linux containers for consistent development and deployment. *Linux J.* 2014:239.
- Midgley, N. G., and Tonkin, T. N. (2017). Reconstruction of former glacier surface topography from archive oblique aerial images. *Geomorphology* 282, 18–26. doi: 10.1016/j.geomorph.2017.01.008
- Mlambo, R., Woodhouse, I. H., Gerard, F., and Anderson, K. (2017). Structure from Motion (SfM) photogrammetry with drone data: a low cost method for monitoring greenhouse gas emissions from forests in developing countries. *Forests* 8:68. doi: 10.3390/f8030068
- Moran, S. M., Emmerich, W. E., Goodrich, D. C., Heilman, P., Holifield Collins, C. D., Keefer, T. O., et al. (2008). Preface to special section on Fifty Years of Research and Data Collection: U.S. Department of Agriculture Walnut Gulch Experimental Watershed. *Water Resour. Res.* 44:W05SW01. doi: 10.1029/2007WR006083
- Nouwakpo, S. K., Weltz, M. A., and McGwire, K. (2015). Assessing the performance of structure-from-motion photogrammetry and terrestrial lidar for reconstructing soil surface microtopography of naturally vegetated plots. *Earth Surf. Processes Landforms* 41, 308–322. doi: 10.1002/esp.3787
- Osterkamp, W. R. (2008). Geology, soils, and geomorphology of the walnut gulch experimental watershed, tombstone, Arizona. *J. Arizona Nevada Acad. Sci.* 40, 136–154. doi: 10.2181/1533-6085-40.2.136
- Pelletier, J. D., Nichols, M. H., and Nearing, M. A. (2016). The influence of holocene vegetation changes on topography and erosion rates: a case study at walnut gulch experimental watershed, Arizona. *Earth Surf. Dyn. Discuss.* 4, 471–488. doi: 10.5194/esurf-4-471-2016
- Pelletier, J. D., and Orem, C. A. (2014). How do sediment yields from post-wildfire debris-laden flows depend on terrain slope, soil burn severity class, and drainage basin area? Insights from airborne-LiDAR change detection. *Earth Surf. Processes Landforms* 39, 1822–1832. doi: 10.1002/esp.3570
- Quantum GIS Development Team (2017). *Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project*. Available online at: <https://www.qgis.org>
- Riegl (2017). *Reigl VZ-400 Data Sheet*. Available online at: http://www.riegl.com/uploads/tx_pxpriegl/downloads/10_DataSheet_VZ-400_2017-06-14.pdf
- Safriel, U., Adeel, Z., Niemeijer, D., Puigdefabregas, J., White, R., Lal, R., et al. (2006). "Dryland systems," in *Ecosystems and Human Well-being. Current State and Trends*, Vol. 1, (Washington, DC: Island Press), 625–656.
- Sankey, T., Donager, J., McVay, J., and Sankey, J. B. (2017a). UAV Lidar and hyperspectral fusion for forest monitoring in the southwestern USA. *Remote Sens. Environ.* 195, 30–43. doi: 10.1016/j.rse.2017.04.007
- Sankey, T. T., McVay, J., Swetnam, T. L., McClaran, M. P., Heilman, P., and Nichols, M. (2017b). UAV hyperspectral and lidar data and their fusion for arid and semi-arid land vegetation monitoring. *Remote Sens. Ecol. Conserv.* 1–14. doi: 10.1002/rse2.44. [Epub ahead of print].
- Scott, R. (2010a). "AmeriFlux radiological and meteorological data for santa Rita mesquite savanna site," *Carbon Dioxide Information Analysis Center (CDIAC) Datasets* (Tucson, AZ).
- Scott, R. L. (2010b). Using watershed water balance to evaluate the accuracy of eddy covariance evaporation measurements for three semiarid ecosystems. *Agric. For. Meteorol.* 150, 219–225. doi: 10.1016/j.agrformet.2009.11.002
- Scott, R. L., Huxman, T. E., Cable, W. L., and Emmerich, W. E. (2006). Partitioning of evapotranspiration and its relation to carbon dioxide exchange in a chihuahuan desert shrubland. *Hydrol. Processes* 20, 3227–3243. doi: 10.1002/hyp.6329
- Skirvin, S., Kidwell, M., Biedenbender, S., Henley, J. P., King, D., Collins, C. H., et al. (2008). Vegetation data, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resour. Res.* 44:W05S08. doi: 10.1029/2006WR005724
- Stewart, C. A., Cockerill, T. M., Foster, I., Hancock, D., Merchant, N., Skidmore, E., et al. (2015). "Jetstream: a self-provisioned, scalable science and engineering cloud environment," in *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure* (St. Louis, MI: ACM), 1–8. doi: 10.1145/2792745.2792774
- Stoker, J. M., Brock, J. C., Soular, C. E., Ries, K. G., Sugarbaker, L. J., Newton, W. E., et al. (2016). *USGS Lidar Science Strategy—Mapping the Technology to the Science*: U.S. Geological Survey Open-File Report 2015–1209, 33.
- Trimble (2012) *Trimble R10 GNSS Dat Sheet. Trimble Navigation Limited 10355 Westmoor Dr Westminster CO 80021 USA*. Available online at: https://solutions.seilerinst.com/Portals/1/Trimble/TrimbleR10_DS_1012_DataSheetLR.pdf
- Trimble (2013) *Trimble S6 Robotic Total Station Data Sheet. Trimble Navigation Limited 10355 Westmoor Dr Westminster CO 80021 USA*. Available online at: http://trl.trimble.com/docshare/dsweb/Get/Document-208580/022543-098L_TrimbleS6_DS_0613_LR.pdf
- Turner, M. G. (1989). Landscape Ecology: the effect of pattern on process. *Annu. Rev. Ecol. Syst.* 20, 171–197. doi: 10.1146/annurev.es.20.110189.001131
- Turner, M. G., O'Neill, R. V., Gardner, R. H., and Milne, B. T. (1989). Effects of changing spatial scale on the analysis of landscape pattern. *Landsc. Ecol.* 3, 153–162.
- Wallace, L., Lucieer, A., Malenovsky, Z., Turner, D., and Vopěnka, P. (2016). Assessment of forest structure using two UAV techniques: a comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests* 7:62. doi: 10.3390/f7030062
- Walls, R. (2017). "CyVerse data commons," in *Plant and Animal Genome XXV Conference. Plant and Animal Genome, 2017* (San Diego, CA).
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M. (2012). 'Structure-from-motion' photogrammetry: a low-cost, effective tool for geoscience applications. *Geomorphology* 179, 300–314. doi: 10.1016/j.geomorph.2012.08.021
- Woodcock, C. E., and Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sens. Environ.* 21, 311–332. doi: 10.1016/0034-4257(87)90015-0
- Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., et al. (2016). An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sens.* 8:501. doi: 10.3390/rs8060501
- Zimmerman, R., Doan, D., Leung, L., Mason, J., Parsons, N., and Shahid, K. (2017). "Commissioning the world's largest satellite constellation," in *Conference on Small Satellites*. Available online at: <http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=3669&context=smallsat>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Swetnam, Gillan, Sankey, McClaran, Nichols, Heilman and McVay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

[@frontiersin](https://twitter.com/frontiersin)



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership