Title
Semi-direct tree reconstruction using terrestrial LiDAR point cloud data

Authors
Bailey, Brian N
Ochoa, Miguel H

Publication Date
2018-04-01

DOI
10.1016/j.rse.2018.02.013

Peer reviewed

Semi-direct tree reconstruction using terrestrial LiDAR point cloud data

Brian N. Bailey[a,*], Miguel H. Ochoa[a]

[a]*Department of Plant Sciences, University of California, Davis, Davis, CA USA*

Abstract

A new method was developed for reconstructing the geometric structure of large plants such as trees at the leaf-scale by utilizing terrestrial LiDAR data. The primary goal of the work was to develop a feasible means for accurately and rapidly reconstructing or "digitizing" entire trees in order to specify the position, orientation, and size of every leaf in digital tree models that provide geometric inputs for high-resolution biophysical models or analyses. As with any optical measurement technique, a primary challenge is accurately accounting for plant matter that is occluded from view of the sensor. The present method is termed "semi-direct" because it uses a triangulation procedure to approximately directly reconstruct as many leaves as possible that are in view of the scanner. For plant matter obstructed from view, a statistical backfilling procedure was used to add additional leaves such that the three-dimensional distribution of leaf area and orientation of the reconstructed plant matched that of the actual plant on average. In a best case scenario such as when leaf density is low, nearly all leaf area is directly reconstructed from the scan and the branch and clumping structure is preserved within the reconstruction. In the worst case scenario such as when the leaf density is very high and nearly all leaves are occluded from view of the scanner, only a small fraction of leaves can be directly reconstructed,

*Corresponding author
Email address: `bnbailey@ucdavis.edu` (Brian N. Bailey)

but at a minimum the distribution of leaf area and the leaf angle distribution across the reconstructed plant will be consistent with that of the actual plant. Unlike many other approaches, the present method does not rely on the woody matter of the plant to provide a skeleton for reconstruction, and can be used in dense plants where little woody matter is visible from the scanner.

Keywords: Leaf angle distribution function, Plant architecture, Plant reconstruction, Terrestrial LiDAR

---

1. Introduction

Leaf-level measurements of many biophysical processes (e.g., exchange of water vapor, CO_2, and heat) have become routine, yet scaling these processes up to entire plants and canopies remains a considerable challenge, as performing direct measurements of biophysical processes at these scales if often not possible (Amthor, 1994; Ehleringer, 2000). Instead, our understanding of whole-plant and -canopy biophysical processes typically relies on models that attempt to aggregate information originating at the leaf scale into plant communities. Such models make simplifying assumptions that focus on bulk canopy behavior, such as “big leaf” or “multilayer” models (Sinclair et al., 1976; Amthor, 1994; DePury and Farquhar, 1997). Given the scale of canopy representation in these models, inputs are also typically bulk values specified at or near the canopy scale.

With the continued exponential increase in computational performance (Moore, 1965), we are now in a position where direct scaling from leaves to canopies (i.e., representing every leaf in a canopy) is within reach. High-resolution, three-dimensional models are becoming increasingly common, and are able to represent an incredibly wide range of scales (e.g., Bailey et al., 2014, 2016; Bailey, 2018). The next generation of biophysical models are likely to shed new light on how processes at various scales interact to determine plant behavior over plant communities.

A considerable challenge in the utilization of such models is the accurate specification of geometric inputs. As the goal of these models is to explicitly represent heterogeneity at various scales and its impact on canopy-level processes, we must be able to accurately measure and input this geometry into the models (Vos et al., 2010; Sarlikioti et al., 2011). Manual measurement of canopy geometry is far too time consuming to be useful at providing canopy-level inputs at the leaf scale.

Remote sensing techniques have provided a means for rapidly measuring and recording the full three-dimensional geometry of plants for use in computer models (i.e., "digitizing"). These techniques make a compromise between level of detail and the size of system that can be represented. Various methods are available to extract plant-scale structural parameters such as crown diameter, height, etc. from remote measurements (e.g., Morsdorf et al., 2004; Henning and Radtke, 2006; Rosell et al., 2009; Yang et al., 2013). The clear advantage of these approaches is that they can be used to rapidly measure large spatial scales, but they do not provide detailed information at the sub-plant scale that may be needed for high-resolution modeling. At the opposite end of the spectrum, methods are also available to measure the full plant structure at the leaf scale. Early work by Sinoquet et al. (1998) used an electromagnetic instrument to directly record the position and orientation of individual foliage elements, which is limited by the need to manually place the instrument next to each leaf. Previous workers have also been relatively successful in using photographic methods to directly reconstruct small plants where nearly all foliage is in direct view of cameras placed on the perimeter of the plant (e.g., Delagrange and Rochon, 2011; Li et al., 2013; Pound et al., 2014). However, these methods cannot be used directly with large plants where a significant portion of plant area is occluded from view.

For large plants such as trees, the problem of measuring the full vegetative structure is complicated by the sheer size of the plants, number of leaves, and potentially large fraction of leaves occluded from view of a remote sensor. If only the woody structure of the tree is of interest, the occlusion problem becomes much less substantial. Numerous methods have been developed based on laser scanning that use the woody structure of the plant as a road map through laser scanning point clouds (e.g., Binney and Sukhatme, 2009; Xu et al., 2007; Côté et al., 2009; Raumonen et al., 2013; Hackenberg et al., 2015; Mèndez et al., 2016). Starting at the trunk, branches can be traced throughout the tree using point connectivity information, which can then be used to generate a reconstruction of the woody tree structure.

If reconstructions of trees at the leaf scale are desired, the occlusion problem must be somehow confronted. Often this involves measurement of the overall tree structure and making reasonable guesses as to where individual leaves should be placed. For example, Shlyakhter et al. (2001) used an aggregate approach which utilized photographic methods to determine the general shape of tree crowns, and then used a structural model to create a simulated tree that fit within the measured crown shape. In cases where vegetation is sparse or leaf-off measurements are available, a reconstruction of the woody structure can be used as a "skeleton" to guide the placement of individual leaves (e.g., Xu et al., 2007; Côté et al., 2009, 2011). Delagrange and Rochon (2011) demonstrated the possibility of adding leaves to the branch skeleton using allometric relations, but this method relies on empirical relations that may or may not be generally applicable.

Evaluations of plant reconstruction methods are most commonly performed using visual comparisons, as it is difficult to quantitatively evaluate their accuracy given that measurements of the true plant structure is typically not available. While many reconstruction methods produce tree models that appear visually reasonable, it is unclear whether the reconstructions are accurate enough for use in detailed model simulations. Côté et al. (2009) noted that reconstructed plants should be "radiatively consistent" with the actual plants, meaning that radiative transport through the reconstructed plants should be approximately equivalent to that of the actual plants. Côté et al. (2009) were able to produce tree reconstructions for *Pinus* species that demonstrated radiative consistency based on measurements of radiation reflection and transmission.

In this work, we develop a "semi-direct" method that uses terrestrial LiDAR data to reconstruct large plants such as trees that match the three-dimensional leaf area and angle distribution of the actual plant being reconstructed. The method is semi-direct in that it directly reconstructs the majority of leaves that are in direct view of the LiDAR scanner. The method then uses a statistical backfilling approach to recreate occluded leaves in a manner that ensures the overall leaf area and angle distribution matches that of the actual plant. Since the reconstructed leaf area and angle distributions are consistent with the actual trees, the reconstructions are applicable for use in model simulations of processes such as light interception.

2. Method description

2.1. Terrestrial LiDAR scanning

Typical terrestrial LiDAR scanning instruments are compact units that can be mounted on a tripod, and are used to measure the distance to surrounding objects. The instrument emits a large number of concentrated pulses or beams of radiation into the surrounding spherical space. In the event that a beam intersects solid matter, some fraction of the radiation beam is scattered back to the instrument. Using various methods such as time of flight, the instrument can calculate and record the distance to beam-object intersection points. The direction in which the pulse was sent is also known by the instrument, which allows calculation of the Cartesian (x, y, z) position of beam-object intersection points (Fig. 1a). By emitting millions of beams into the surrounding space, the instrument effectively maps the three-dimensional geometry surrounding the scan location.

Terrestrial LiDAR instruments generally do not emit beams at random, rather they perform a systematic scan of the surrounding spherical space. Most commonly, instruments discretely scan a certain range of zenithal angles while continuously rotating between a range of discrete azimuthal angles (Fig. 1b). This creates an approximately uniform two-dimensional grid of points in spherical space. The scan resolution is given by the number of discrete scan zenithal directions N_θ (# rows), and the number of discrete scan azimuthal directions N_φ (#columns), with $N_\theta \times N_\varphi$ being the total number of points in the scan.

2.2. Scan point triangulation

The basic idea behind the plant reconstruction methodology presented in this work is to connect adjacent scan hit points to form triangles, then identify continuous triangle groups that reconstruct individual leaves. The triangulation methodology is described in detail by Bailey and Mahaffee (2017b), and a brief description is repeated below.

The triangulation algorithm first seeks to construct a two-dimensional grid of scan points in spherical space. This grid consists of a (θ, φ) coordinate for each ray sent by the scanner (Fig. 1). This creates a two-dimensional plane of points that can be triangulated (Fig. 2). Bailey and Mahaffee (2017b) suggested an efficient triangulation algorithm that can be used when the indices of the scan points in the 2D spherical grid are recorded by the scanner. This allows for the construction of a "scan table" in which rows correspond to each scan zenithal

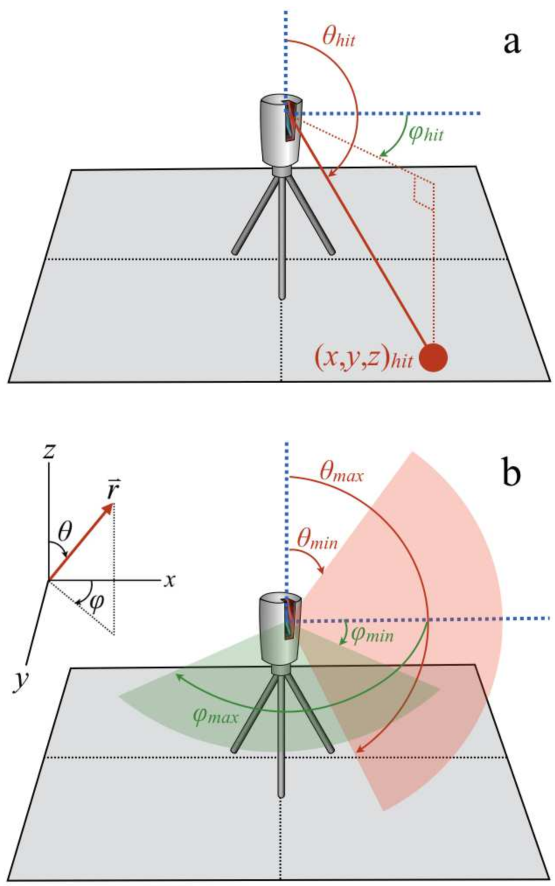

Figure 1: Schematic depiction of terrestrial LIDAR scanning. (a) scanning pattern in spherical coordinates, illustrating the range of scan zenithal angles (θ_{min} through θ_{max}) and azimuthal angles (φ_{min} through φ_{max}). (b) Cartesian coordinate $(x,\, y,\, z)_{hit}$ of hit point, and corresponding spherical coordinate $(\theta,\, \varphi)$.

angle, and columns correspond to each scan azimuthal angle. Given this table, it is relatively straightforward to form triangles between adjacent points in the uniform grid since scan point connectivity is already known. For instruments that do not directly record this information (such as the instrument used in this work), standard 2D Delaunay triangulation can be used (Press et al., 2007), which has the trade-off that it requires more computational effort since point connectivity is not initially known. Triangles exceeding a size or aspect ratio threshold are rejected to prevent erroneous triangles from being formed, such as triangles that connect adjacent leaves. Since each triangle vertex corresponds to a laser hit point, the (x, y, z) coordinates of the vertices are also known. The resulting triangulation gives a set of triangles that follow the surfaces of individual leaves that are in view of the scanner.

2.3. Direct leaf surface reconstruction

Neighboring triangles are connected to form continuous groups, where each group presumably corresponds to all or a portion of an individual leaf's surface. To accomplish this, an algorithm is applied that is similar to a traditional "flood-fill" algorithm (e.g., Lee, 1987), except that it connects adjacent triangles instead of adjacent pixels (Fig. 2). For each triangle, any neighboring connected triangles are identified, where a "connected" triangle is defined as a triangle that shares two vertices with the current triangle being examined. By requiring that two vertices are shared rather than one, this reduces the likelihood that adjacent leaves or branches will inadvertently be merged into a common group. The algorithm begins by iterating over each triangle in the triangulated set. The first triangle is assigned a fill group identifier of "0". For each triangle, any neighboring connected triangles are determined. If any connected triangles exist, each connected triangle is added to the current fill group by assigning it the current group identifier, and the neighbors of each connected triangle are examined in a recursive manner. The recursion halts when there are no connected triangles that have not yet been added to the current fill group. In this case, the current fill group has been completed, and the fill group identifier is increased by one. The original iteration over triangles proceeds, where triangles that have already been assigned to a fill group are skipped. Once the iteration is completed, all possible triangle groups have been formed (Fig. 2).

Triangle groups are filtered by their area to exclude very small or large groups. If only one to a few small triangles are identified in a single group, it is typically not desirable to allocate an entire leaf to this group. These small groups are filtered by specifying a threshold value for the minimum group surface area, below which groups are rejected. Similarly a threshold value is specified for the maximum group surface area, which is typically set to be much larger than the expected area of a single leaf. The primary purpose of filtering large leaf groups is to remove outliers when calculating the characteristic leaf dimension (see below).

Each continuous fill group is then replaced by a "prototype" leaf. Although there are many ways a prototype leaf could be specified (e.g., a rectangle, a triangular mesh), this work used a PNG image to define the leaf shape (Fig. 3). A leaf is specified by a planar rectangle, but a portion of that rectangle is removed according to the transparency channel of the PNG image (Bailey, 2018). The length and width of the prototype are denoted by l and w, and the fraction of the total rectangular area that is not transparent is the solid fraction s (Fig. 3).

There are three quantities that must be specified for each leaf: its (x, y, z) position, size, and orientation. The position and average orientation are readily available from the triangulation; the leaf is placed at the location of the triangle group centroid and oriented in the direction of the average triangle group normal. However, the size is more difficult to determine, because only a relatively

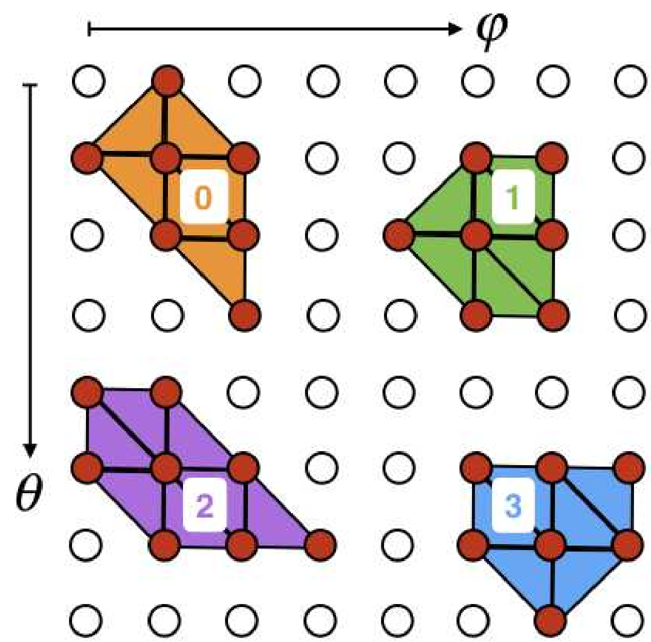

Figure 2: "Flood-fill" grouping of triangles. A two-dimensional grid of scan points in $\theta - \varphi$ space is shown, with "misses" denoted by open circles and "hits" denoted by filled circles. Connected triangle groups are identified and assigned a group identifier. In the example shown, four continuous triangle groups are formed, which are given identifiers of 0, 1, 2, and 3.

few number of leaves on the outside of the plant in full view of the scanner will be completely reconstructed by the triangulation. Most of the leaves are occluded to some degree and will only be partially triangulated, and thus the area of the fill groups will be less than the actual leaf area. One could perform manual measurements of leaf size using a ruler to obtain representative values for leaf sizes. The drawback of this method, aside from having to perform manual measurements, is that leaf size can change with position in the plant and thus specifying a single size value may not be representative. The method used here involved considering only the largest triangulated groups (e.g., 10 largest groups), and taking the characteristic leaf length L to be the average of the square root of the group areas. The spatial distribution of leaf size can be approximately represented by dividing the plant into sub-volumes, and the largest triangulation groups in each volume can be used to determine the representative leaf size for that particular volume. In order to specify the dimension of a leaf from the characteristic leaf size L, we must specify a leaf aspect ratio, which is the ratio r of the length of the leaf parallel (l) to perpendicular (w) to the midrib. Given that $L \equiv \sqrt{a} = \sqrt{wls}$ and $r \equiv l/w$, the leaf length l is equal to $L\sqrt{r/s}$, and $w = l/r$.

2.4. Backfilling occluded leaves

Direct leaf reconstruction based on the triangulation only represents a subset of the total leaf area. The leaf area that is not triangulated because it is occluded or because the triangulation failed must be represented through other means. In the present method, the remaining leaf area is reconstructed by backfilling leaves until the leaf area density of the reconstructed plant matches that of the actual plant. The plant is discretized into a grid of rectangular sub-volumes called voxels (see Bailey and Mahaffee, 2017a), and LiDAR points are grouped by the voxel in which they reside. The method described in detail by Bailey and

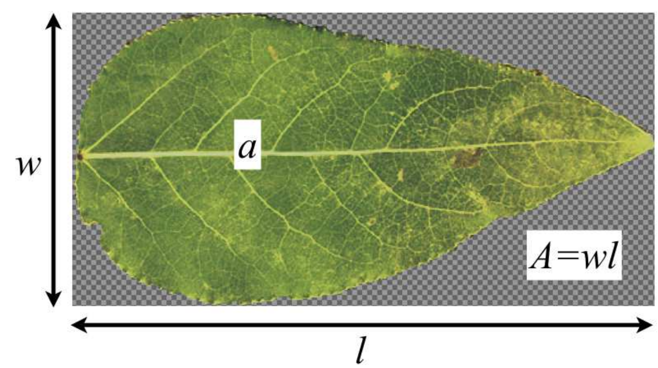

Figure 3: Example of leaf prototype image. The solid portion of the image is colored, while the checkered portion of the image is considered transparent. The area of the solid portion is a, and the area of the total image is $A = wl$, where w and l are respectively the width and length of the prototype. The fraction of the total image that is solid is $s = a/A$.

Mahaffee (2017a) can then be used to calculate the leaf area density and leaf angle distribution of the actual plant for each voxel using the LiDAR scan data. More precisely, it should be noted that the method actually measures the area density of all plant matter including branches. This method gives a relatively accurate measure of the total surface area of plant matter within each voxel for the actual plant. It is also straightforward to use the directly reconstructed leaves from Sect. 2.3 to determine the amount of leaf area in each voxel resulting from the direct reconstruction, as the area of each reconstructed leaf is known. The difference between the total and directly reconstructed area is the amount of leaf area that remains to be added through backfilling.

The backfilling process begins by randomly choosing a directly reconstructed leaf within a given voxel, which is duplicated and placed at a random, uniformly distributed position within the voxel. This process continues for each voxel until the reconstructed leaf area in the voxel matches the "actual" leaf area. It is possible that too much leaf area could have been added during the direct reconstruction, in which case leaf area can be removed by randomly deleting leaves which we term "thinning". Based on this process, the resulting reconstructed leaf area and leaf angle distribution should be consistent with that of the actual plant for each voxel. This method is dispersive in that it tends to spread out leaves in space. The larger the fraction of leaves that are directly reconstructed, the less dispersive the reconstruction method becomes, and the better the reconstructed tree will match the structure of the actual tree.

2.5. Woody plant material

Several methods have been suggested by previous authors for reconstruction of woody plant material (e.g., Xu et al., 2007; Binney and Sukhatme, 2009; Mèndez et al., 2016; Li et al., 2016). In this work, we focus only on reconstructing leaves within the crown volume, and present a simple method for reconstructing the main trunk similar to that of Xu et al. (2007). The primary purpose of representing the main trunk is simply to provide a visual reference for qualitative evaluation of the reconstruction. A voxel is specified that contains the portion of trunk to be reconstructed. Hit points within this voxel are triangulated, and the flood-fill algorithm of Sect. 2.3 is applied. The largest fill group is identified, which is assumed to correspond to the trunk. This produces a triangular mesh that approximately reconstructs the portion of the trunk visible from the scanner.

It should also be noted that it is possible that the reconstruction algorithm for leaves could inadvertently identify branches as a leaf group. Rather than attempting to filter out these relatively rare instances, the algorithm is simply applied in the same way as for leaves, and it is assumed that a reconstructed branch is a reasonable location to place a leaf. This work focuses on trees in which the (visible) leaf area is much larger than the woody area. For trees where the woody area is substantial compared to the leaf area, LiDAR hit points corresponding to woody material could be separated within the scan (Béland et al., 2014), and a branch reconstruction algorithm could be applied separate from the leaf reconstruction method presented in this work.

2.6. Multiple scan positions

To reconstruct an entire tree, scans from multiple locations surrounding the tree are typically required and must be combined. Generally, the the instrumentation on-board the scanner for measuring geographic position is not accurate enough to be used to merge multiple scans (it provides only an estimate). Standard methods are available to register multiple scans to a common global coordinate system, such as the iterative closest point (ICP) method (Zhang, 1994), or methods that use reflectors, checkerboards, spheres, or other common targets placed within the scan. Many instruments also come with software developed by the manufacturer that use proprietary algorithms.

The method for calculating the leaf area contained within each voxel (Bailey and Mahaffee, 2017a) does not distinguish between different scan positions, thus aggregating multiple scans is straightforward. For any given ray direction, the probability that a ray intersects vegetation, the leaf normal vector, and path length through the voxel are simply added to running totals for all scans. The totals for all scan points from all scan locations are used along with Beer's law to solve for leaf area density within the voxel (Bailey and Mahaffee, 2017a). For the leaf reconstruction procedure, the algorithm is applied on a scan-by-scan basis, and reconstructed leaves from each scan are simply aggregated together to form the reconstructed plant.

3. Evaluation of method

3.1. Data collection and processing details

Scanning data was collected for a 5 m tall Emerald Sunshine Elm (*Ulmus propinqua*) located in Davis, California USA to demonstrate application of the method and evaluate its performance. The tree was scanned using a full-waveform Riegl VZ-1000 terrestrial LiDAR scanner (RIEGL Laser Measurement Systems GmbH; Horn, Austria). The scanner sends concentrated beams of radiation with a wavelength of 1550 nm in a uniformly gridded pattern in spherical space, covering a range from 30°-130° in the zenithal direction and 0-360° in the azimuthal direction. The maximum scan resolution is about 41,000×150,000 points in the zenithal×azimuthal directions. The beam diameter as it leaves the instrument is approximately 7 mm, which diverges at an angle of approximately 0.3 mrad, meaning that at 10 m range the beam diameter is roughly 8.5 mm. The instrument can scan up to 122,000 points per second, with a range from 2.5 m up to approximately 350-450 m at this scanning rate. The full-waveform

LiDAR instrument used can record multiple hit points per pulse, but the point cloud was filtered to consider only the closest hit per pulse. The instrument was equipped with an on-board digital camera (Nikon D810 36 Mega Pixel) that was used to assign RGB color values to each scan point and obtain images for visual comparison with reconstructions.

Four scans were performed at equally spaced intervals surrounding the tree, which were automatically registered to a common coordinate system using Riegl's proprietary RiSCAN Pro software. The scanner was positioned on a tripod approximately 1.25 m above the ground, and approximately 5.5 m from the trunk of the tree. This distance was chosen because it was as close as possible to the tree such that the entire tree was in view of the scanner and digital camera. A modest scan resolution of 2500×4500 points (zenith$\times$azimuth) was chosen. At 10 m range, this meant that adjacent points on a surface orthogonal to the beam direction were separated by roughly 3.5-7 mm and 7-14 mm in the zenithal and azimuthal directions, respectively, depending on beam zenithal angle. Given the chosen resolution, the scans took roughly 2 minutes to complete, with an additional 2-3 minutes for GPS location and collection of digital photographs. Scans were performed under very low wind speed conditions to minimize leaf disturbances. The above scanning configuration worked well for the particular application of interest, but in general configurations are expected to be application-dependent. Since point density effectively decreases with distance, trees that are larger or further away will require a higher scanning density. Additionally, very large or dense trees could require more scans, potentially at multiple heights to ensure that all portions of the tree are in view of the scanner.

Additionally, the size of 40 random leaves were measured to evaluate the performance of the method for determining the leaf dimensions from the LiDAR data. The lengths of the leaves parallel and perpendicular to the midrib were measured and recorded for each of the 40 leaves. Admittedly, a robust sampling strategy was not used, and only leaves within reach of the ground were measured. This is because only a rough estimate of leaf size was desired in order to assess whether results of the LiDAR method were at least reasonable. Alternatively, a more robust quantification of errors in leaf dimension is presented in Sect. 4.3 using synthetic data.

For processing the data, a uniformly spaced 3D grid of voxels was overlaid on the tree, within which leaf area was calculated using the method described above and by Bailey and Mahaffee (2017a). The tree crown was divided into a $10\times10\times10$ grid of rectangular voxels, each of size $0.5\times0.5\times0.4$ m^3. In the triangulation methodology, triangles were rejected if the length of any of their sides exceeded 5 cm, or if their aspect ratio was greater than 10. In the flood-fill algorithm, triangle groups were rejected if their total area was less than 1 cm^2 or greater than 200 cm^2, which were chosen because they are much smaller or larger than the expected area of a leaf. The maximum leaf area threshold is relatively easy to specify since it is straightforward to estimate the maximum expected leaf area. Understanding the minimum leaf area threshold is slightly less straightforward. It may be undesirable to specify a minimum area threshold that is too small because we typically want at least a few connected triangles for each leaf in order to have confidence that the triangle group uniquely corresponds to a leaf. We recommend a minimum threshold that is roughly an order of magnitude smaller than the maximum area threshold. However, we varied the minimum area threshold between 0.1 and 50 cm^2 and found very little impact on the resulting tree reconstructions. Using tighter area thresholds generally results in slightly less directly reconstructed leaf area, but the overall distribution of leaf area and orientation remains the same.

3.2. Generation of synthetic scanning data

Quantitative evaluation of LiDAR data processing methods is extremely difficult when applied to large, dense trees, since there is typically no "gold standard" measurement against which to compare. Before proceeding to the application of the method under field conditions, an alternative approach is presented that uses simulated or "synthetic" LiDAR data in which the exact vegetation structure is known (see also Côté et al., 2009; Mèndez et al., 2013; Raumonen et al., 2013; Bailey and Mahaffee, 2017a,b). This approach was adopted to test the plant reconstruction method's ability to reproduce the distribution of leaf area, orientation, and characteristic size. Admittedly, this method also has its drawbacks, namely that it is for an idealized case. Thus, it clearly does not replace the need to perform some type of field validation, but represents a powerful tool for algorithm testing and evaluation.

The synthetic LiDAR data was produced by performing a ray-tracing simulation that mimics the actual LiDAR scanning procedure described above in Sect. 3.1. In short, a model or "reference" tree was created based on the architectural model of Weber and Penn (1995), which specifies the position of the trunk, branches, and leaves. The trunk and branches were made up of a mesh of triangular elements, and the leaves were rectangular transparency masks with zero thickness (see Fig. 3) of size $6{\times}20$ cm^2 and a solid fraction $s = 0.62$. The overall tree was roughly 7.5 m tall with a crown diameter of about 5.5 m, and had branches with a diameter ranging from 0.36 m at the trunk base to zero at the branch tips. The woody structure of the tree was made up of about 77,000 triangles, and the tree had about 30,000 leaves. Leaf orientations were specified as described in Weber and Penn (1995), where leaves tend to rotate around the axial direction of the branches, which leads to interesting non-uniform angle distributions (see Figs. 8 and 9). Rays were launched from each of the four simulated scanner locations in a spherical pattern approximately matching that of an actual LiDAR scan. Ray-object intersection tests were performed to determine the (x, y, z) location of the closest intersection point (Suffern, 2007). Note that for simplicity it was assumed that a ray had an infinitely small diameter that maintains 100% of the emitted intensity, which is not true for an actual LiDAR beam. The resulting field of (x, y, z) intersection points was taken to be an approximation of an actual LiDAR scan, and was used to run the reconstruction methodology. For the simulated tree case, the voxel grid size was slightly different than that of the real tree because the tree crowns were slightly different sizes (but still consisted of 10×10×10 total voxels). For this case, the voxels had a size of 0.55×0.55×0.65 m^3. On average, each voxel contained about 30 leaves.

3.3. Error quantification

Errors between exact and simulated data were quantified using three standard metrics: the index of agreement (Willmott, 1981, 1982), root-mean-squared error (RMSE), and mean bias. The index of agreement is defined as

$$\mathrm{d} = 1 - \frac{\sum\limits_{i=1}^{N} \left(M_i - L_i\right)^2}{\sum\limits_{i=1}^{N} \left(\left|M_i - \overline{M}\right| + \left|L_i - \overline{L}\right|\right)^2}, \tag{1}$$

where M_i and L_i are respectively the i^{th} estimated and exact values for each voxel, with N total values, and an overbar denotes an average over all voxels. The RMSE is defined as

$$RMSE = \left(\sum_i \left(L_i - M_i\right)^2\right)^{1/2}, \tag{2}$$

and the mean bias is defined as

$$\text{bias} = \frac{1}{N}\sum_{i=1}^{N}\left(M_i - L_i\right). \tag{3}$$

4. Evaluation using synthetic scanning data

4.1. Visualization

The visualizations shown in Fig. 4 provide a means for performing a qualitative evaluation of the reconstruction method using the synthetic scanning data. Overall, the reconstruction (Fig. 4b,d) appears visually reasonable in comparison with the reference tree (Fig. 4a,c), and reproduces the general tree structure. Clearly, the reconstruction does not produce an exact replica of the reference tree nor is it intended to do so. As mentioned previously, the reconstruction method is dispersive, meaning that it tends to spread out leaves and diminish structure. As a result, the reconstructed tree has lost some branch and clumping structure compared to the reference tree. The sub-voxel-scale structure that is present is primarily due to directly reconstructed leaves, which are shown in Fig. 5.

4.2. Leaf area

A more quantitative evaluation of the reconstruction methodology can be conducted by performing a voxel-by-voxel comparison of leaf area between the reconstructed and reference trees (Fig. 6a). Since the exact amount of leaf area in each voxel is known from the reference tree, this provides a means for quantifying the error in measured leaf area. It should be noted that this exercise is primarily a test of the leaf area measurement method of Bailey and Mahaffee (2017a), as this is what determines how much total leaf area should be produced within each voxel.

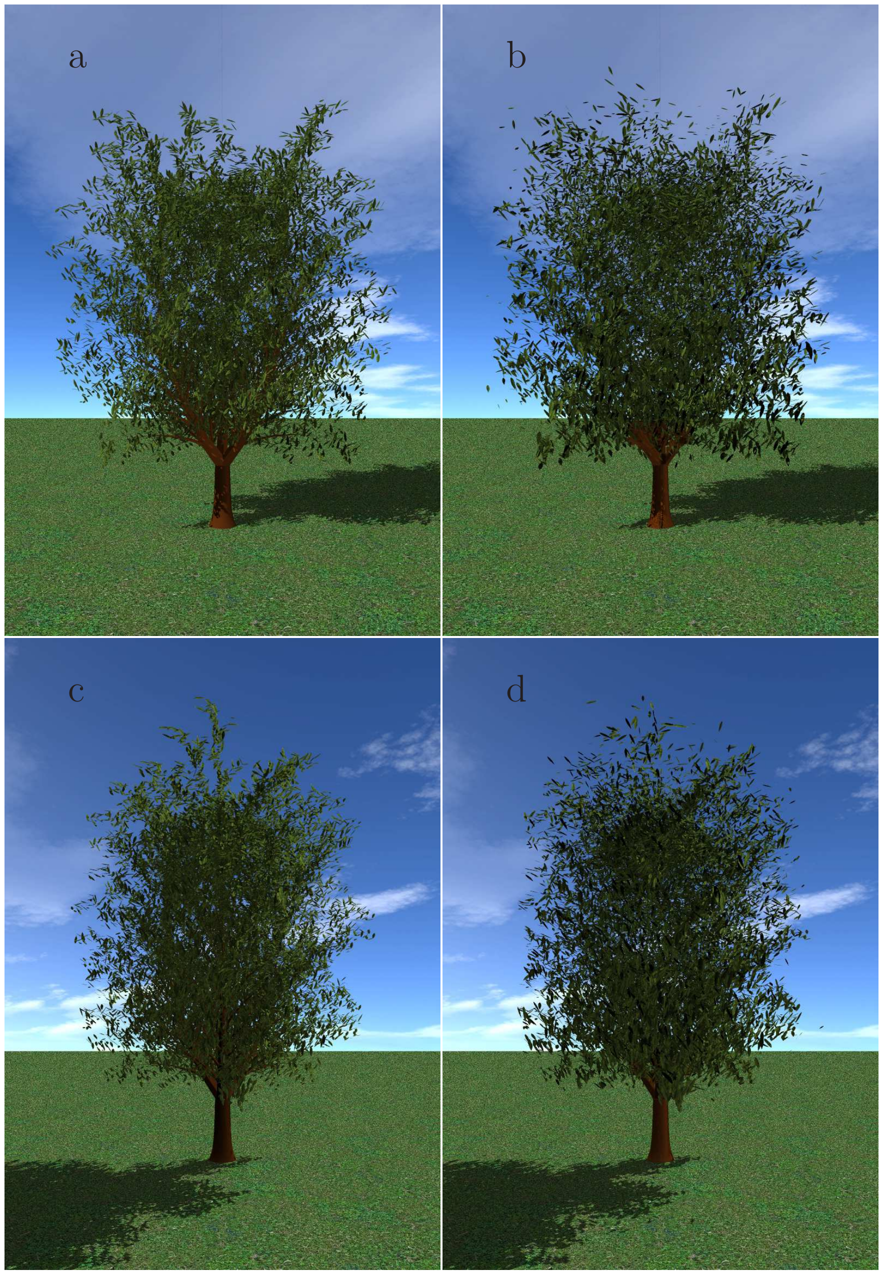

Figure 4: Visualization of (a,c) computer-generated or “reference” tree, and (b,d) reconstruction of the reference tree based on simulated LiDAR scanning data for two opposing viewing angles.

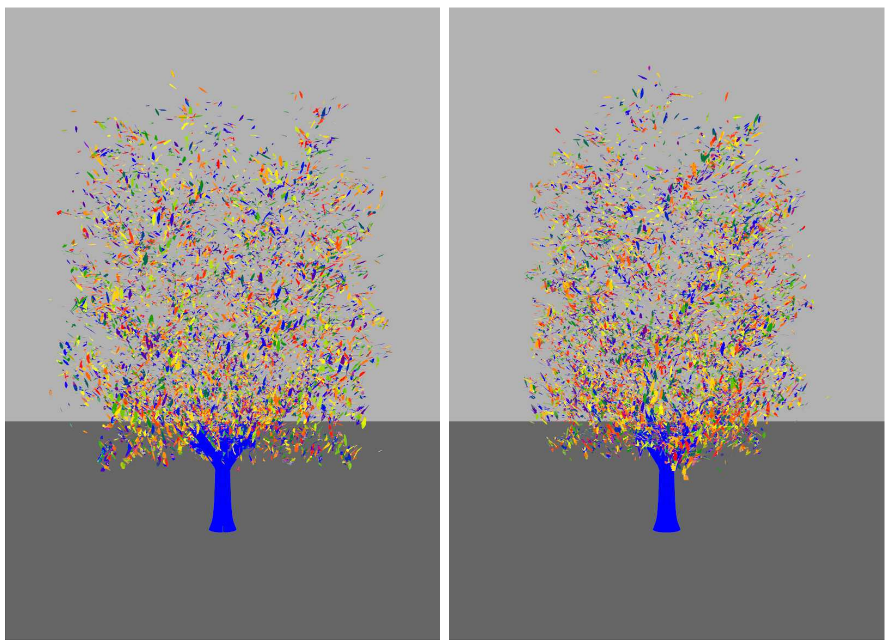

Figure 5: Visualization of the triangulated leaf groups used to determine the locations of directly reconstructed leaves in the reconstruction shown in Fig. 4b,d for two opposing viewing angles. Each independent fill group is given a unique color.

The index of agreement between the reference and reconstructed total leaf area within the 1000 voxels was 94.7%, and the RMSE was 0.169 m^2 (Fig. 6a), indicating reasonably good overall agreement. There is a notable amount of scatter in the LiDAR measurements, particularly as leaf area density becomes large. There is a small overall negative bias in the estimated leaf area (-0.056 m^2), meaning that the LiDAR methodology tended to slightly underestimate the actual amount of total leaf area. Above roughly 1 m^2 of leaves per voxel the scatter becomes increasingly apparent and there is more consistent under prediction. This is likely because the LiDAR inversion methodology used to measure leaf area loses sensitivity as leaf area index along the beam path becomes large (which occurs when either leaf area density or voxel size becomes large). The inversion for leaf area is based on the LiDAR's measurement of the probability that a beam is intercepted by leaves within a given voxel, and as leaf area index along the beam's path becomes large there is little difference in this probability as leaf area varies. There was no clear location in the tree where the relative error in leaf area tended to be largest, but the absolute error was largest wherever leaf area happened to be largest.

Figure 7 indicates the amount of leaf area that was directly reconstructed on average. The majority of voxels required backfilling to reach the measured leaf area. Some voxels required that more than 100% of the directly reconstructed leaf area be removed via thinning to match the measured leaf area.

4.3. Characteristic leaf dimension

The ability of the reconstruction method to determine the characteristic leaf size within a given voxel was evaluated in Fig. 6b. The leaf dimension in the reference tree was constant at 8.7 cm. The reconstruction method slightly skews to the left of the actual leaf dimension, which is expected since the leaf is rarely 100% triangulated. However, the majority of the reconstructed leaves are near

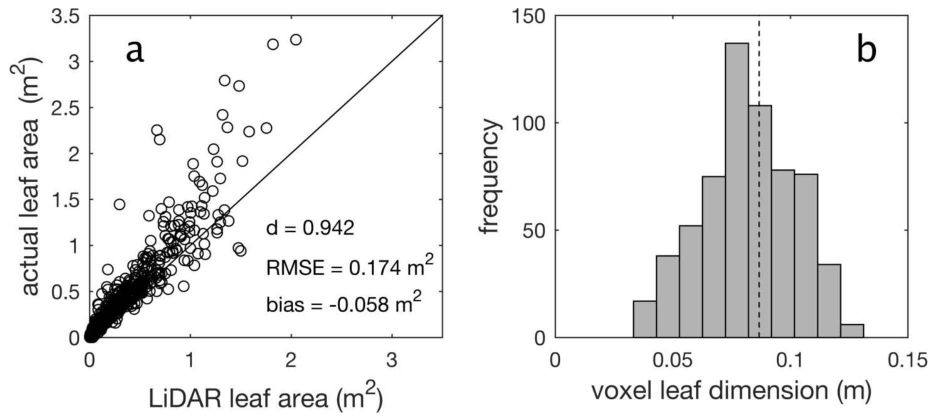

Figure 6: Comparison of exact values of leaf area (a) and leaf dimension (b) with values obtained from the synthetic LiDAR reconstruction for each voxel. In (a), the diagonal line denotes perfect agreement, and overall agreement is quantified by the index of agreement d, the root-mean-squared error (RMSE), and the mean bias. In (b), the dashed vertical line denotes the (constant) exact value, and bars give a histogram of predicted values over all voxels. Note that the characteristic leaf dimension L was defined as $\sqrt{a}$, where a is the leaf surface area.

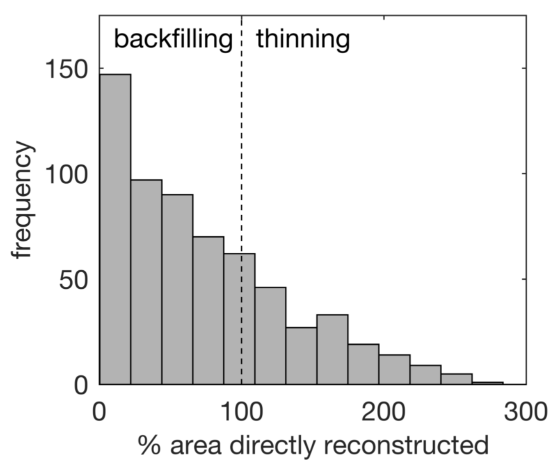

Figure 7: Histogram of the fraction of leaf area within each voxel that was directly reconstructed. Bars to the left of the vertical dotted line correspond to voxels that had less reconstructed leaf area than actual leaf area, and thus required backfilling. Bars to the right of the vertical dotted line correspond to voxels that had more reconstructed leaf area than actual leaf area, and thus required thinning.

the actual leaf dimension, and the actual mean bias is small at -4.8 mm. The overall RMSE for all reconstructed leaves was 2.0 cm.

4.4. Leaf orientation

To make it feasible to plot voxel leaf angle probability density functions (PDFs), the 10×10×10 voxel grid was downsampled to a 2×2×2 grid by simply aggregating neighboring voxels together. Probability density functions are plotted for the leaf inclination (Fig. 8) and azimuthal (Fig. 9) angles within each of these 8 total grid voxels. The exact PDFs from the reference tree are compared against PDFs for the reconstructed tree. PDFs were calculated following the procedure used in Bailey and Mahaffee (2017b), which can be consulted for further details. Overall, the reconstruction is able to qualitatively reproduce the general trends in the inclination and azimuthal angle PDFs. There are some deviations between the reference and reconstructed PDFs due to inadequate sampling of the true PDF, but overall agreement appears visually reasonable. A two sample Kolmogorov-Smirnov test was performed to quantitatively compare the exact and reconstructed leaf angle distributions for each voxel. The distributions for every voxel passed the Kolmogorov-Smirnov test at a 5% confidence interval for both the leaf inclination and azimuthal angle PDFs.

5. Evaluation using field data

5.1. Visualization

Unfortunately, the type of data used to perform quantitative evaluation of the method is not readily available in the field. Therefore, agreement between the actual (field) and reconstructed trees was assessed based on visual comparisons. In order to do so, the reconstructed trees must be visualized in a manner that is consistent with the way in which the scanner's digital camera perceives the actual tree, which was not an issue in the previous section since identical

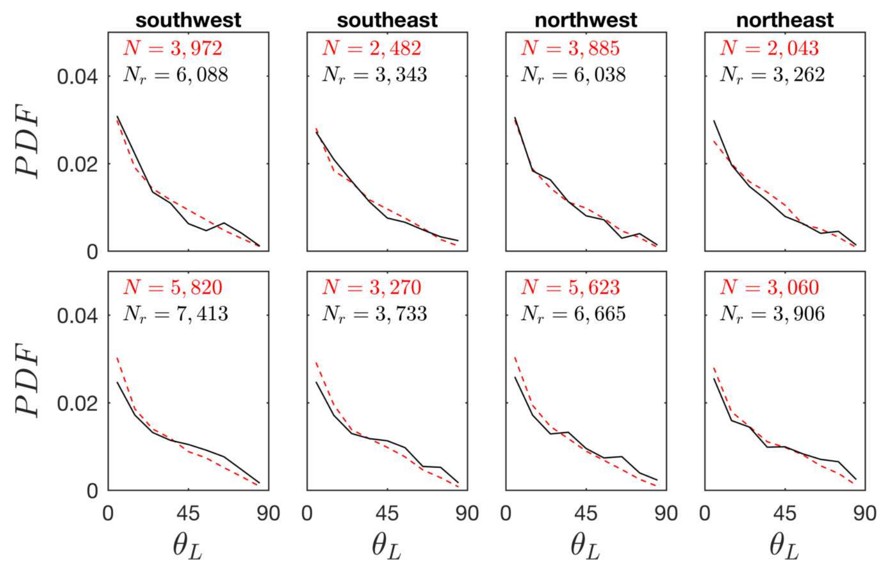

Figure 8: Probability density functions (PDFs) of leaf inclination angle (θ_L) with a discrete bin size of 10° for eight different leaf zones. The solid black lines correspond to the inclination angle of N total leaves from the tree reconstruction, and the dashed red lines correspond to the inclination angle of N_r total leaves from the reference tree (exact). The leaf zones were determined by downsampling the $10 \times 10 \times 10$ voxel grid to a grid of $2 \times 2 \times 2$ voxels. The top and bottom rows of plots correspond to the top and bottom half of the tree crown, respectively, and each column of plots corresponds to a different azimuthal zone of the tree.

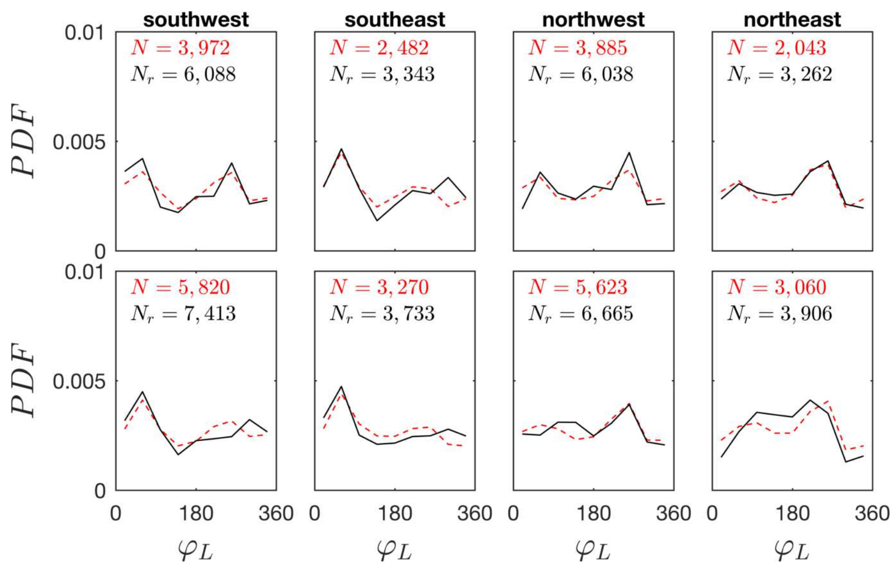

Figure 9: Probability density functions (PDFs) of leaf azimuthal angle (φ_L) with a discrete bin size of 40° for eight different leaf zones. The solid black lines correspond to the azimuthal angle of N total leaves from the tree reconstruction, and the dashed red lines correspond to the azimuthal angle of N_r total leaves from the reference tree (exact). The leaf zones were determined by downsampling the $10 \times 10 \times 10$ voxel grid to a grid of $2 \times 2 \times 2$ voxels. The top and bottom rows of plots correspond to the top and bottom half of the tree crown, respectively, and each column of plots corresponds to a different azimuthal zone of the tree.

visualization techniques could be applied for the actual and reconstructed trees. In plotting geometric elements associated with the reconstructed trees, a standard rectangular perspective transformation was applied to the geometry that approximately matched that of the camera lens (Shirley and Morley, 2003). The appropriate field of view for the camera lens was determined through trial-and-error by comparing visualizations of the LiDAR point cloud and photographs. As a result, there is some error in the visualization comparisons due to the camera model used to visualize the reconstructed trees.

Figure 10 shows a visualization of the tree triangulation, with each fill group given a unique color. Based on visual inspection, the method appears to perform reasonably well in terms of identifying individual leaves. Because of the limited number of distinct colors in the pseudocolor mapping, it can be difficult in some instances to determine whether neighboring leaves are in the same fill group or are actually slightly different colors. There appear to be instances in which neighboring leaves that are very close together are inadvertently placed into the same triangle group. However, these occurrences seem to be relatively minimal and still offer reasonable guesses as to where leaves should be placed.

A visualization of the resulting reconstruction as compared with actual photograph and point cloud data is shown in Fig. 11. Qualitative comparison between the actual and reconstructed trees shows close agreement. Individual shoot structures are clearly replicated by the reconstruction. Many individual leaves are closely represented by the reconstructed leaves. Figure 10 shows which leaves were a result of the direct reconstruction, and indicates that the algorithm is able to identify a large number of individual leaves. The majority of the grid voxels had less than 50% of the leaf area directly reconstructed, and very few required thinning (Fig. 12b). Leaf size prediction seemed to be reasonable (Fig. 12a) and resulted in a visually consistent tree reconstruction.

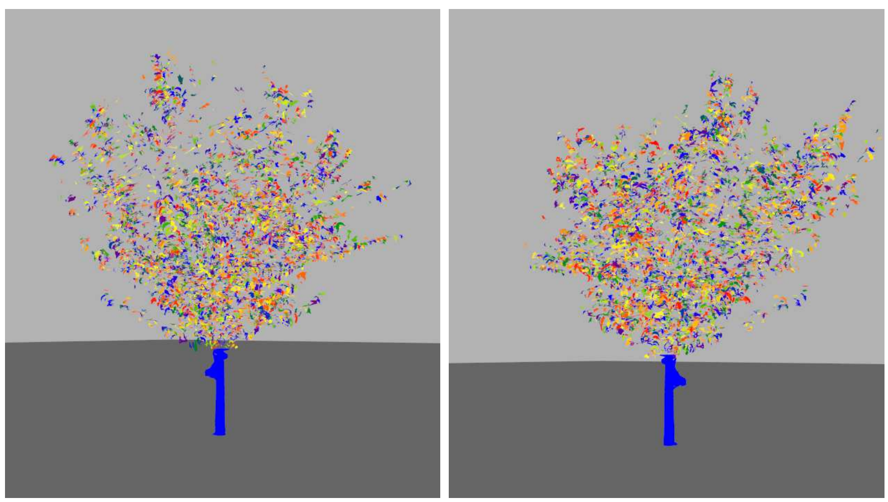

Figure 10: Visualization of the triangulated leaf groups used to determine the locations of directly reconstructed leaves in the reconstruction shown in Fig. 11b,d (actual elm tree) for two opposing viewpoints. Each independent leaf fill group is given a unique color.

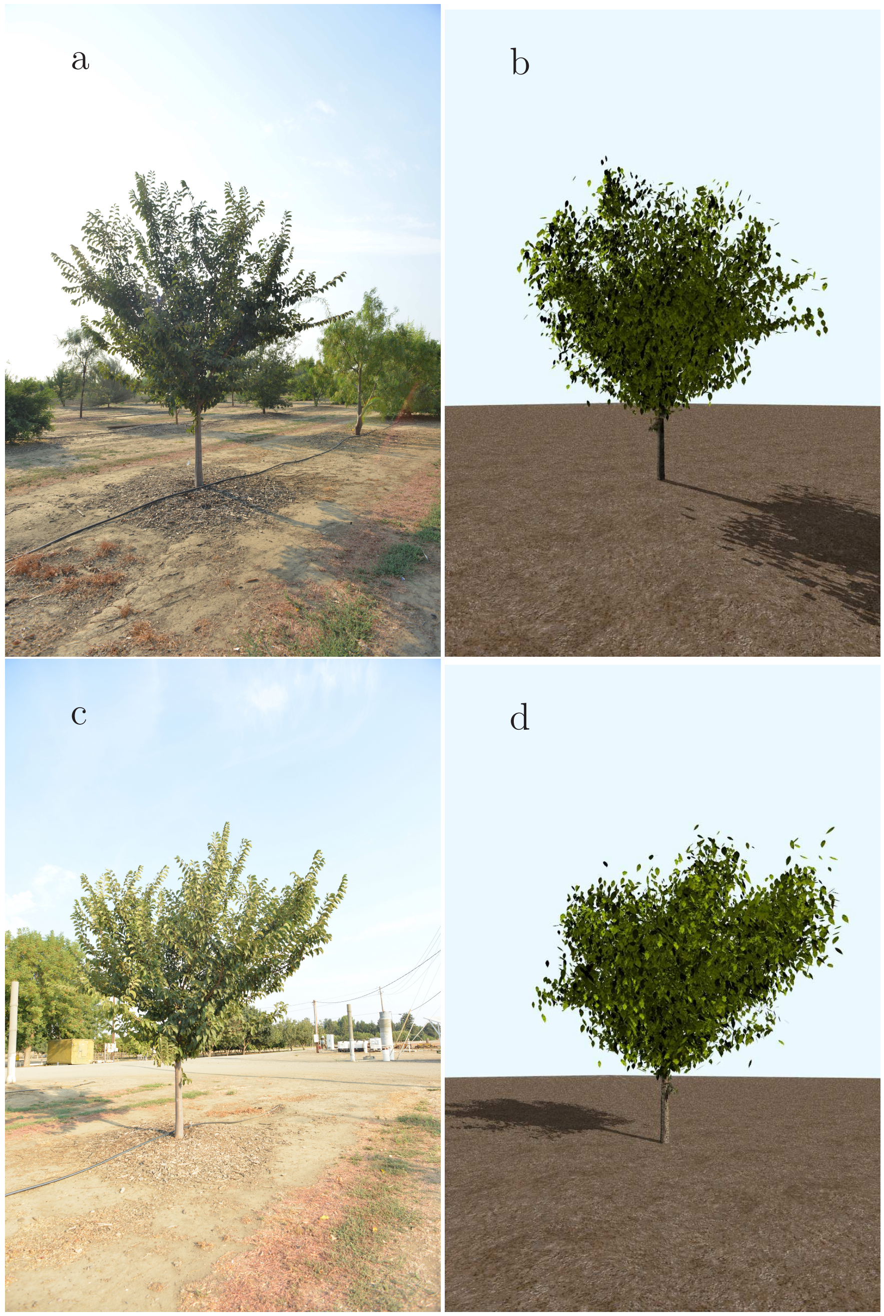

Figure 11: Visual comparison of actual elm tree photograph (a,c), and reconstructed elm tree (b,d) for two opposing viewpoints.

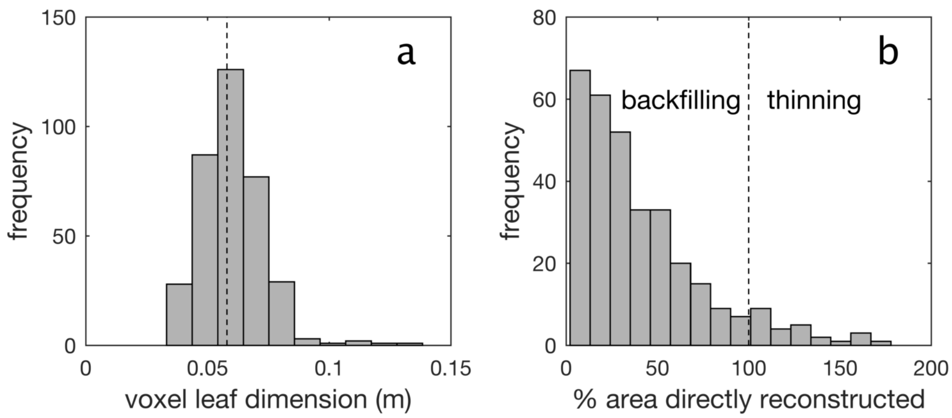

Figure 12: Histogram of characteristic leaf dimension in each grid voxel for the reconstructed tree in Fig. 11b,d (a), and histogram of the fraction of directly reconstructed leaf area within each grid voxel for the reconstructed tree in Fig. 11b,d (b).

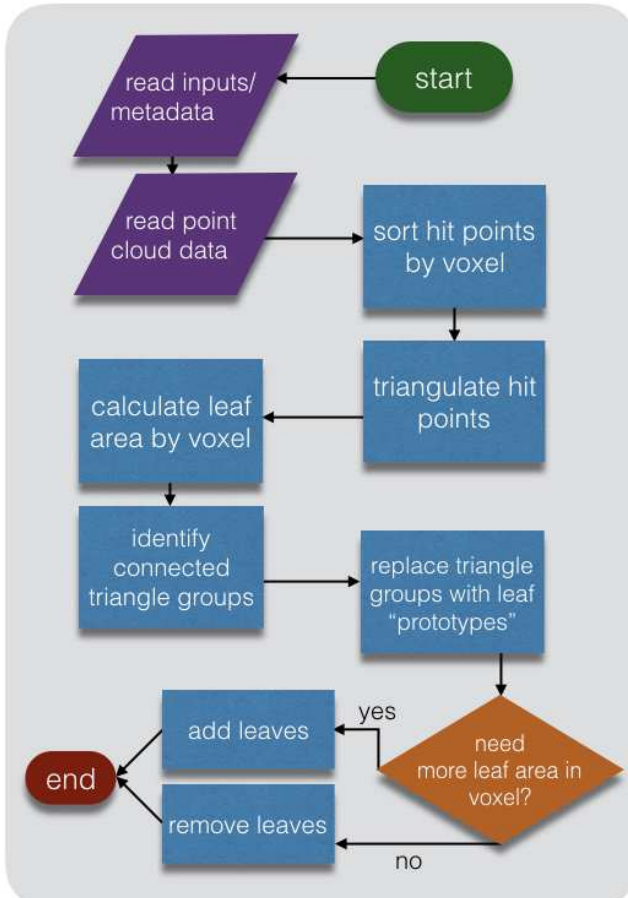

Figure 13: Flow chart illustrating plant reconstruction methodology.

6. Discussion and conclusions

A semi-direct method was developed and tested that uses terrestrial LiDAR scanning data to reconstruct the architecture of large plants such as trees. A summary of the overall reconstruction algorithm is presented in Fig. 13. The method is termed semi-direct because it seeks to directly reconstruct as many leaves as possible that are in view of the scanner. The resulting direct reconstruction typically represents only a fraction of the total leaf area of the plant. To reconstruct hidden or occluded leaf area, a statistical backfilling procedure was employed in which leaves were added (or removed) such that the overall leaf area and leaf orientation distributions matched that of the actual plant. This was accomplished by using the methods developed by Bailey and Mahaffee (2017a) and Bailey and Mahaffee (2017b) to measure the leaf angle and leaf area distributions within a user-defined grid of voxels, then adding leaves such that they are consistent with these measured distributions. Thus, the resulting reconstruction is not an exact replica of the plant, rather it is a statistical reconstruction that is consistent with the actual tree at the scale of the voxel grid at that particular instant in time.

In contrast with other methods that rely on the tree branch structure as a skeleton for reconstruction (e.g., Xu et al., 2007; Côté et al., 2009), the present method does not utilize branch structure in the reconstruction of leaves. As a result, the method is applicable to dense plants where little to no wood area is visible from the scanner. The leaf density does, however, affect the quality of the reconstruction. For relatively sparse plants, a larger fraction of leaves are visible to the scanner, and thus the direct portion of the reconstruction represents a larger fraction of the total reconstructed area, which preserves more of the vegetation structure. For dense plants, much of the leaf area is occluded from view of the scanner, and therefore less leaf area is directly reconstructed. Regardless, the reconstructed leaf area and orientation is still consistent with the actual plant at the voxel scale to within the accuracy that the instrument can measure leaf area and orientation for each voxel. A drawback of the present method is that it is dispersive, meaning that it tends to diminish plant structure by spreading out leaves.

Dense vegetation or large voxel sizes have the effect of diminishing the accuracy of the measurement of leaf area. This work suggested that voxels with denser leaves tended to have higher errors in predicted leaf area (Sect. 4.2).

Although not explored in detail, it appeared that for the case examined in this work, errors started to become significant when the voxels contained greater than about 1 m^2 of leaves (note that these values may be case-specific). Future work is needed to more thoroughly examine how various factors affect errors in the leaf area measurement method, as such an exercise was beyond the scope of this work which focused primarily on the reconstruction technique. Small voxels have an additional advantage that they reduce the tendency of the method to disperse or spread out leaves. However, using too small of voxels could become problematic if there are not enough ray samples per voxel.

Aside from the voxel size, there are relatively few tunable parameters in the reconstruction methodology itself. To utilize the triangulation algorithm, the user must specify the maximum allowable triangle dimension. This value is typically easy to specify, because results have shown little sensitivity over a wide range, as long as this dimension is much larger than the distance between adjacent hit points and much smaller than the typical distance between adjacent leaves (Bailey and Mahaffee, 2017b). The reconstruction algorithm requires the specification of threshold values for the minimum and maximum allowable surface area of a triangulated leaf "group". Regardless of how these threshold values are specified, the reconstructed tree will still be consistent with the actual tree at the voxel scale in terms of the leaf area and orientation distributions.

The results of this work have important implications in terms of the ability to provide accurate inputs to detailed biophysical models and analyses. Models are now able to represent plant-related processes at the leaf scale (e.g., Vos et al., 2010; Sarlikioti et al., 2011; Bailey, 2018), and combining such models with consistent, leaf-level plant reconstructions provides a means by which these processes can be scaled from leaf-to-tree-to-canopy without the need for often questionable assumptions of homogeneity. In addition to modeling-related efforts, reconstruction data can aid in studies seeking to understand relations between plant structure and function (Meinzer et al., 2011). In order to perform terrestrial scans of entire canopies, scanning throughput needs to be increased. Scanners can be placed on easily movable or autonomous platforms to increase throughput (e.g., Kukko et al., 2012). However, it is important to note that the data processing methods utilized in this work require a stationary sensing platform for the duration of the scan. This also makes utilization of aerial platforms a challenge. At the scan resolution used in this work, scans take only a couple of minutes each (if color photographs are not also collected) and can potentially scan several surrounding trees simultaneously. Canopy-scale reconstruction of very large trees (>10 m) is likely to introduce additional challenges such as requiring higher scan resolution and high occlusion toward the top of the canopy.

Acknowledgements

Financial support of this work by the American Vineyard Foundation grants 2015-1825/2016-1825/2017-1825, U.S. National Science Foundation grant AGS PREEVENTS 1664175, and the USDA National Institute of Food and Agriculture, Hatch project number CA-D-PLS-2401-H.

References

Amthor, J.S., 1994. Scaling CO_2-photosynthesis relationships from the leaf to the canopy. Photosyn. Res. 39, 321–350.

Bailey, B.N., 2018. Efficient ray-tracing methods for modeling radiation transfer in leaf-resolving plant canopy simulations. Ecol. Model. 398, 233–245.

Bailey, B.N., Mahaffee, W.F., 2017a. Rapid, high-resolution measurement of leaf area and leaf orientation using terrestrial LiDAR scanning data. Meas. Sci. Technol. 28, 064006.

Bailey, B.N., Mahaffee, W.F., 2017b. Rapid measurement of the three-dimensional distribution of leaf orientation and the leaf angle probability density function using terrestrial lidar scanning. Remote Sens. Environ. 193, 63–76.

Bailey, B.N., Overby, M., Willemsen, P., Pardyjak, E.R., Mahaffee, W.F., Stoll, R., 2014. A scalable plant-resolving radiative transfer model based on optimized GPU ray tracing. Agric. For. Meteorol. 198-199, 192–208.

Bailey, B.N., Stoll, R., Pardyjak, E.R., Miller, N.E., 2016. A new three-dimensional energy balance model for complex plant canopy geometries: Model development and improved validation strategies. Agric. For. Meteorol. 218-219, 146–160.

Béland, M., Baldocchi, D.D., Widlowski, J.L., Fournier, R.A., 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.

Binney, J., Sukhatme, G.S., 2009. 3D tree reconstruction from laser range data, in: IEEE International Conference on Robotics and Automation, pp. 1321–1326.

Côté, J.F., Fournier, R.A., Egli, R., 2011. An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. Environ. Model. Softw. 26, 761–777.

Côté, J.F., Widlowski, J.L., Fournier, R.A., Verstraete, M.M., 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. Remote Sens. Environ. 113, 1067–1081.

Delagrange, S., Rochon, P., 2011. Reconstruction and analysis of a deciduous sapling using digital photographs or terrestrial-LiDAR technology. Ann. Bot. 108, 991–1000.

DePury, D.G.G., Farquhar, G.D., 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant Cell Environ. 20, 537–557.

Ehleringer, J.R., 2000. Temperature and energy budgets, in: Pearcy, R.W., Ehleringer, J.R., Mooney, H., Rundel, P.W. (Eds.), Plant Physiological Ecology: Field Methods and Instrumentation. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 117–135.

Hackenberg, J., Spiecker, H., Calders, K., Disney, M., Raumonen, P., 2015. SimpleTree - an efficient open source tool to build tree models from TLS clouds. Forests 6, 4245–4294.

Henning, J.G., Radtke, P.J., 2006. Ground-based laser imaging for assessing three-dimensional forest canopy structure. Photogramm. Eng. Remote Sens. 72, 1349–1358.

Kukko, A., Kaartinen, H., Hyyppä, J., Chen, Y., 2012. Multiplatform mobile laser scanning: Usability and performance. Sensors 12, 11712–11733.

Lee, E., 1987. Region filling using two dimensional grammars, in: Proceedings. 1987 IEEE International Conference on Robotics and Automation, pp. 1475–1478.

Li, F., Chattopadhyay, S., Akbar, S.A., Elfiky, N.M., Kak, A., 2016. A novel visualization tool for evaluating the accuracy of 3D sensing and reconstruction algorithms for automatic dormant pruning applications, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 338–346.

Li, Y., Fan, X., Mitra, N.J., Chamovitz, D., Cohen-Or, D., Chen, B., 2013. Analyzing growing plants from 4d point cloud data. ACM Trans. on Gr. 32, 157.

Meinzer, F.C., Lachenbruch, B., Dawson, T.E. (Eds.), 2011. Size- and Age-Related Changes in Tree Structure and Function. Springer. 514 pp.

Mèndez, V., Catalàn, H., Rosell-Polo, J.R., Arnó, J., Sanz, R., 2013. LiDAR simulation in modelled orchards to optimise the use of terrestrial laser scanners and derived vegetative measures. Biosys. Eng. 115, 7–19.

Mèndez, V., Rosell-Polo, J.R., Pascual, M., Escolà, A., 2016. Multi-tree woody structure reconstruction from mobile terrestrial laser scanner point clouds based on a dual neighbourhood connectivity graph algorithm. Biosys. Eng. 148, 34–47.

Moore, G.E., 1965. Cramming more components onto integrated circuits. Electronics , 114–117.

Morsdorf, F., Meier, E., Kötz, B., Itten, K.I., Dobbertin, M., Allgöwer, B., 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. Remote Sens. Environ. 92, 353–363.

Pound, M.P., French, A.P., Murchie, E.H., Pridmore, T.P., 2014. Automated recovery of three-dimensional models of plant shoots from multiple color images. Plant Physiol. 166, 1688–1698.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge, U.K. 1256 pp.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., 2013. Fast automatic precision tree models from terrestrial laser scanning data. Remote Sens. 5, 491–520.

Rosell, J.R., Llorens, J., Sanz, R., Arnó, J., R.-D., M., Masip, J., Escolà, A., Camp, F., Solanelles, F., Gràcia, F., Gil, E., Val, L., Planas, S., Palacín, J., 2009. Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning. Agric. For. Meteorol. 149, 1505–1515.

Sarlikioti, V., de Visser, P.H.B., Marcelis, L.F.M., 2011. Exploring the spatial distribution of light interception and photosynthesis of canopies by means of a functional-structural plant model. Ann. Bot. 107, 875–883.

Shirley, P., Morley, R.K., 2003. Realistic Ray-Tracing. Second ed., A. K. Peters, Natick, MA. 225 pp.

Shlyakhter, I., Rozenoer, M., Dorsey, J., Teller, S., 2001. Reconstructing 3D tree models from instrumented photographs. IEEE Comp. Graph. Appl. , 53–61.

Sinclair, T.R., Murphy, C.E., Knoerr, K.R., 1976. Development and evaluation of simplified models for simulating canopy photosynthesis and transpiration. Brit. Ecol. Soc. 13, 813–829.

Sinoquet, H., Thanisawanyangkura, S., Mabrouk, H., Kasemsap, P., 1998. Characterization of the light environment in canopies using 3D digitising and image processing. Ann. Bot. 82, 203–212.

Suffern, K.G., 2007. Ray Tracing from the Ground Up. A K Peters/CRC Press, Boca Raton, FL. 784 pp.

Vos, J., Evers, J.B., Buck-Sorlin, G.H., Andrieu, B., Chelle, M., de Visser, P.H.B., 2010. Functional-structural plant modelling: a new versatile tool in crop science. J. Exp. Bot. 61, 2101–2115.

Weber, J., Penn, J., 1995. Creation and rendering of realistic trees, in: SIGGRAPH '95 Proceedings of the 22nd annual conference on computer graphics and interactive techniques, ACM. pp. 119–128.

Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184–194.

Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bull. Amer. Meteor. Soc. 63, 1309–1313.

Xu, H., Gossett, N., Chen, B., 2007. Knowledge and heuristic based modeling of laser-scanned trees. ACM Trans. on Gr. 26, 19.

Yang, X., Strahler, A.H., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., Dubayah, R.O., Woodcock, C.E., Ni-Meister, W., 2013. Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). Remote Sens. Environ. 135, 36–51.

Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. Int. J. Com. Vis. 13, 119–152.

List of Figures

Fig. 1: Schematic depiction of terrestrial LIDAR scanning. (a) scanning pattern in spherical coordinates, illustrating the range of scan zenithal angles (θ_{min} through θ_{max}) and azimuthal angles (φ_{min} through φ_{max}). (b) Cartesian coordinate $(x,\, y,\, z)_{hit}$ of hit point, and corresponding spherical coordinate $(\theta,\, \varphi)$.

Fig. 2: "Flood-fill" grouping of triangles. A two-dimensional grid of scan points in $\theta-\varphi$ space is shown, with "misses" denoted by open circles and "hits" denoted by filled circles. Connected triangle groups are identified and assigned a group identifier. In the example shown, four continuous triangle groups are formed, which are given identifiers of 0, 1, 2, and 3.

Fig. 3: Example of leaf prototype image. The solid portion of the image is colored, while the checkered portion of the image is considered transparent. The area of the solid portion is a, and the area of the total image is $A = wl$, where w and l are respectively the width and length of the prototype. The fraction of the total image that is solid is $s = a/A$.

Fig. 4: Visualization of (a,c) computer-generated or "reference" tree, and (b,d) reconstruction of the reference tree based on simulated LiDAR scanning data for two opposing viewing angles.

Fig. 5: Visualization of the triangulated leaf groups used to determine the locations of directly reconstructed leaves in the reconstruction shown in Fig. 4b,d for two opposing viewing angles. Each independent fill group is given a unique color.

Fig. 6: Comparison of exact values of leaf area (a) and leaf dimension (b) with values obtained from the synthetic LiDAR reconstruction for each voxel. In (a), the diagonal line denotes perfect agreement, and overall agreement is quantified by the index of agreement d, the root-mean-squared

error (RMSE), and the mean bias. In (b), the dashed vertical line denotes the (constant) exact value, and bars give a histogram of predicted values over all voxels. Note that the characteristic leaf dimension L was defined as $\sqrt{a}$, where a is the leaf surface area.

Fig. 7: Histogram of the fraction of leaf area within each voxel that was directly reconstructed. Bars to the left of the vertical dotted line correspond to voxels that had less reconstructed leaf area than actual leaf area, and thus required backfilling. Bars to the right of the vertical dotted line correspond to voxels that had more reconstructed leaf area than actual leaf area, and thus required thinning.

Fig. 8: Probability density functions (PDFs) of leaf inclination angle (θ_L) with a discrete bin size of $10°$ for eight different leaf zones. The solid black lines correspond to the inclination angle of N total leaves from the tree reconstruction, and the dashed red lines correspond to the inclination angle of N_r total leaves from the reference tree (exact). The leaf zones were determined by downsampling the $10 \times 10 \times 10$ voxel grid to a grid of $2 \times 2 \times 2$ voxels. The top and bottom rows of plots correspond to the top and bottom half of the tree crown, respectively, and each column of plots corresponds to a different azimuthal zone of the tree.

Fig. 9: Probability density functions (PDFs) of leaf azimuthal angle (φ_L) with a discrete bin size of $40°$ for eight different leaf zones. The solid black lines correspond to the azimuthal angle of N total leaves from the tree reconstruction, and the dashed red lines correspond to the azimuthal angle of N_r total leaves from the reference tree (exact). The leaf zones were determined by downsampling the $10 \times 10 \times 10$ voxel grid to a grid of $2 \times 2 \times 2$ voxels. The top and bottom rows of plots correspond to the top and bottom half of the tree crown, respectively, and each column of

plots corresponds to a different azimuthal zone of the tree.

Fig. 10: Visualization of the triangulated leaf groups used to determine the locations of directly reconstructed leaves in the reconstruction shown in Fig. 11b,d (actual elm tree) for two opposing viewpoints. Each independent leaf fill group is given a unique color.

Fig. 11: Visual comparison of actual elm tree photograph (a,c), and reconstructed elm tree (b,d) for two opposing viewpoints.

Fig. 12: Histogram of characteristic leaf dimension in each grid voxel for the reconstructed tree in Fig. 11b,d (a), and histogram of the fraction of directly reconstructed leaf area within each grid voxel for the reconstructed tree in Fig. 11b,d (b).

Fig. 13: Flow chart illustrating plant reconstruction methodology.