



Article

Generating Multispectral Point Clouds for Digital Agriculture

Isabella Subtil Norberto , Antonio Maria Garcia Tommaselli *  and Milton Hirokazu Shimabukuro 

Graduate Program in Cartographic Sciences, Faculty of Science and Technology, São Paulo State University (UNESP), Presidente Prudente 19060-900, SP, Brazil; isabella.subtil@unesp.br (I.S.N.); milton.h.shimabukuro@unesp.br (M.H.S.)

* Correspondence: a.tommaselli@unesp.br

Abstract

Digital agriculture is increasingly important for plant-level analysis, enabling detailed assessments of growth, nutrition and overall condition. Multispectral point clouds are promising due to the integration of geometric and radiometric information. Although RGB point clouds can be generated with commercial terrestrial scanners, multi-band multispectral point clouds are rarely obtained directly. Most existing methods are limited to aerial platforms, restricting close-range monitoring and plant-level studies. Efficient workflows for generating multispectral point clouds from terrestrial sensors, while ensuring geometric accuracy and computational efficiency, are still lacking. Here, we propose a workflow combining photogrammetric and computer vision techniques to generate high-resolution multispectral point clouds by integrating terrestrial light detection and ranging (LiDAR) and multispectral imagery. Bundle adjustment estimates the camera's position and orientation relative to the LiDAR reference system. A frustum-based culling algorithm reduces the computational cost by selecting only relevant points, and an occlusion removal algorithm assigns spectral attributes only to visible points. The results showed that colourisation is effective when bundle adjustment uses an adequate number of well-distributed ground control points. The generated multispectral point clouds achieved high geometric consistency between overlapping views, with displacements varying from 0 to 9 mm, demonstrating stable alignment across perspectives. Despite some limitations due to wind during acquisition, the workflow enables the generation of high-resolution multispectral point clouds of vegetation.



Academic Editor: Ray E. Sheriff

Received: 30 September 2025

Revised: 17 November 2025

Accepted: 18 November 2025

Published: 2 December 2025

Citation: Norberto, I.S.; Tommaselli, A.M.G.; Shimabukuro, M.H. Generating Multispectral Point Clouds for Digital Agriculture.

AgriEngineering **2025**, *7*, 407.

<https://doi.org/10.3390/agriengineering7120407>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: digital agriculture; terrestrial LiDAR; multispectral point cloud; data fusion

1. Introduction

Multispectral point clouds, which combine three-dimensional information with spectral attributes, have become a valuable tool in various fields, including precision agriculture, forest monitoring and ecological analysis. Recent methodological advances have made it possible to generate high-quality multispectral 3D point clouds in contexts such as plant phenotyping in near-surface environments [1]. In forestry, multimodal approaches that combine light detection and ranging (LiDAR) with multispectral imagery show great potential for distinguishing tree species and accurately classifying vegetational attributes [2]. Similarly, digital twin-driven simulations using RGB (red, green, blue) cameras and LiDAR sensors have enabled the efficient development and evaluation of machine vision pipelines for fruit detection and sizing, reducing the need for extensive real-field measurements [3]. These capabilities suggest that multispectral point clouds could facilitate the integration of

high-geometric-resolution and high-spectral-resolution data, paving the way for robust applications in plant-level analysis and mapping.

Digital agriculture relies on advanced remote sensing techniques and the analysis of images captured by orbital, aerial and terrestrial sensors to measure morphological attributes from plant height to leaf area [4]. Additionally, tree and shrub modelling allows the detailed mapping and analysis of the plant structure and leaf density, while crop monitoring methods assess plant health and biomass in agricultural fields. However, there are significant challenges in assessing plant characteristics, particularly due to the complexity of plant structures and phenotypic variation throughout plant growth, which makes it difficult to obtain accurate data. The integrity of image-based measurements may be affected by the inclination and deformation of leaves, as well as by plant–light interactions [1]. On the other hand, LiDAR is emerging as an advanced technique that provides detailed spatial information and offers robust solutions to mitigate these challenges in digital agriculture.

LiDAR point clouds are known for having high geometric quality and reliability, but they lack spectral information to describe 3D surfaces and fully extracted plant phenotypes. Conversely, using only an optical camera for surface reconstruction produces a 3D point cloud with a high point density and a large amount of spectral information, but this point cloud is affected by shading and occlusions [5]. Laser pulses can pass through vegetation open spaces and provide effective visibility even in areas that are shaded, such as the bottom of the canopy. A common method used by commercial terrestrial scanners is the integration of a digital camera into the LiDAR device; however, this is a high-cost solution.

An RGB and/or multispectral LiDAR point cloud can be generated through the fusion of the three-dimensional point cloud collected by the LiDAR device with spectral information collected by an optical camera. In the computer vision community, this product is called a ‘photorealistic scene’ and is distinguished by the fact that it reconstructs a real scene in 3D, with sufficient colour and texture information in the projections to provide an accurate representation of the scene from a variety of target perspectives [6]. Challenges in performing data fusion from different sensors include differences in spatial coverage and the orientation of imagery, different scales of observation (pixel size versus LiDAR footprint), unsynchronised data collection and varying conditions (attitude and angle of view), leading to the misalignment of colours (optical) and structures (LiDAR) [7–13].

The fusion of data from different sensors, for example the combination of LiDAR data with lightweight multispectral sensors with a high spatial resolution, can be advantageous. Portable multispectral cameras, for example the Micasense and the Mapir Survey3W, are known for their ability to capture detailed images in multiple spectral bands, allowing a more comprehensive analysis of plant characteristics. Another recent example is Agrowing’s camera [14], which has six spectral lenses for image capture. Agrowing’s camera simultaneously acquires six sub-images from which fourteen image bands are generated [15], allowing the display of the digital numbers (DNs) for each band according to the reflectance properties of the targets. Multispectral LiDAR scanners are also currently available, allowing the detailed data collection of information from the Earth’s surface in different spectral bands [16]. However, the cost of these systems has generally been quite high.

Fusing data to create a multispectral LiDAR product using a LiDAR sensor in combination with a lightweight multispectral camera is a viable option. An important part of this process is ensuring an accurate orientation between the 3D model generated by LiDAR and the optical images, since the data from both sensors are not acquired synchronously. To cover the same field of view as the LiDAR sensor, the camera must be positioned at different positions and orientations. Bundle adjustment (BA) can be used to determine the precise position and orientation of the camera relative to the LiDAR point cloud. This process

enables spectral information from camera images to be mapped onto corresponding points in the LiDAR cloud, ensuring that the correct DN is assigned to each visible point [17]. This requires an occlusion detection process for the identification and labelling of occluded points in the LiDAR point cloud so that double mapping does not occur when the data are fused [5]. The problem of double mapping refers to the overlapping or duplication of information, which leads to ambiguity in the scene when the same target is mapped from different viewpoints [18]. A direct consequence of this issue, when the object is captured at multiple locations in the point cloud, is the ghosting effect. This double mapping causes the duplication of pixels and the creation of 'ghost-artefact' representations of the original object [19]. In a three-dimensional model, colourising the same point from multiple viewpoints can reduce the visual consistency of the reconstruction. To address this issue, there has been much work on detecting and removing or labelling occluded points using various approaches. Two significant categories must be acknowledged: culling algorithms and visibility maps [20]. Culling algorithms are designed to discard points outside the camera's visible volume, thereby optimising processing. Examples include viewing frustum culling (VFC) [21], which eliminates objects outside the field of view; backface culling [22], which discards surfaces facing away from the camera; and occlusion culling [23], which hides points blocked by closer ones. These culling algorithms accelerate subsequent processes of hidden point detection. Visibility mapping algorithms are used to eliminate surfaces or parts of surfaces that are hidden from a given viewpoint (perspective centre of an image). These algorithms are also known as hidden surface removal or labelling. Commonly used techniques include Z-buffering [24], binary space partitioning [25], ray casting [26], ray tracing [27] and hidden points removal (HPR) [28].

Given the challenges involved in fusing data from different sensors, such as detecting occluded points, the double mapping problem and the need for effective visibility maps, in this context, the aim of this research project was the investigation and development of a processing sequence for the fusion of the terrestrial LiDAR point cloud with multispectral imagery. The proposed approach comprises several stages: (1) the geometric alignment of LiDAR data with images from a lightweight multispectral camera to ensure proper registration between clouds and images; (2) the implementation of techniques to optimise the processing workflow and reduce the computational cost; (3) the detection and labelling of occluded points to mitigate the double mapping problem; and (4) the projection of DNs onto the visible points of the 3D cloud to associate spectral information with the points measured by the LiDAR. This processing sequence addresses the main challenges of sensor fusion in an integrated manner, enabling the generation of a reliable multispectral point cloud. The fusion of terrestrial LiDAR and multispectral imagery remains relatively unexplored, particularly in agriculture. The proposed method is compatible with different types of cameras and does not require calibration between sensors. It also reduces the computational cost by processing only the visible points from each perspective. Consequently, the fused data provide robust support for agricultural applications, including monitoring plant health, detecting diseases early, optimising nutrition and improving productivity, thereby contributing to more informed and precise management decisions.

2. Materials

The experiments were carried out using an Agrowing multispectral camera (Agrowing Ltd., Tel Aviv, Israel) based on the Sony Alpha 7 IVR (Sony Corporation, Tokyo, Japan) model. Agrowing modified this camera to create a multispectral sensor with customised filters and lenses, enabling images to be captured in fourteen spectral bands through six multispectral lenses [14]. To obtain the spectral bands, the Bayer filter is combined with the custom filters of the lenses [15]. The camera (Figure 1a) has a nominal focal

length of 21.6 mm, a field of view (FoV) of 25°, a bandwidth of 30 nm and a resolution of 2600×2600 pixels. The bands are arranged into six sub-images (S1 to S6) (Figure 1b).

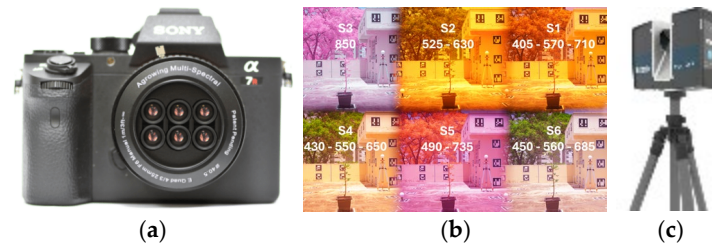


Figure 1. (a) Agrowing multispectral camera. (b) Camera sub-images with the range for bands (nm). (c) Faro TLS.

LiDAR point clouds were collected with the FARO Focus Premium Terrestrial Laser Scanner (TLS) (FARO Technologies Inc., Lake Mary, FL, USA) (Figure 1c), which has a $360^\circ \times 300^\circ$ FoV and a beam divergence of 0.3 mrad; it scans objects up to 350 m away. The data include the intensity of the return pulse processed in the Faro Scene 2022.2 software. Processing was executed on a computer with an Intel® Core™ i9-11900KF CPU (3.5 GHz, 11th Gen), 64 GB RAM and a 64-bit architecture, running Windows 11 Pro and using a set of software including Agisoft Metashape 2.2.2 (Agisoft LLC, St. Petersburg, Russia), MATLAB 2023b (MathWorks, Natick, MA, USA) and Visual Studio 2019 (Microsoft Corporation, Redmond, WA, USA) with Point Cloud Library (PCL) [29]. The LiDAR and camera were operated independently, with an asynchronous acquisition strategy.

Field Data Collection

The initial dataset was collected at a coffee plantation of the Federal University of Uberlândia, MG. Special targets and spheres were strategically positioned on the coffee trees and on the ground (Figure 2a). Due to space restrictions, a single scan was obtained with the TLS in front of the tree. Images were collected at different heights, positions and orientations. A second dataset was collected at São Paulo State University using a single potted apple tree and markers (Figure 2b). The Faro TLS and the camera were moved in a circle around the apple tree, at a distance of approximately 2 m. ARUCO targets were affixed to the walls and floor of the selected area, and they were utilised as ground control points (GCPs). Additionally, spheres were positioned to appear between scans of the point clouds.



Figure 2. (a) Coffee plantation with targets and spheres. (b) Experimental setup with ARUCO markers on the floor and walls, and a potted apple tree.

3. Methods

Figure 3 presents a flowchart illustrating the overall sequence of procedures for assigning spectral values to the 3D point cloud. The process begins with the acquisition of field data, followed by the generation of a high-resolution LiDAR cloud and the extraction of spectral information from the captured images. The initial processing focuses on im-

proving the quality of the LiDAR data by reducing noise and ensuring that the point cloud accurately represents the scanned surfaces.

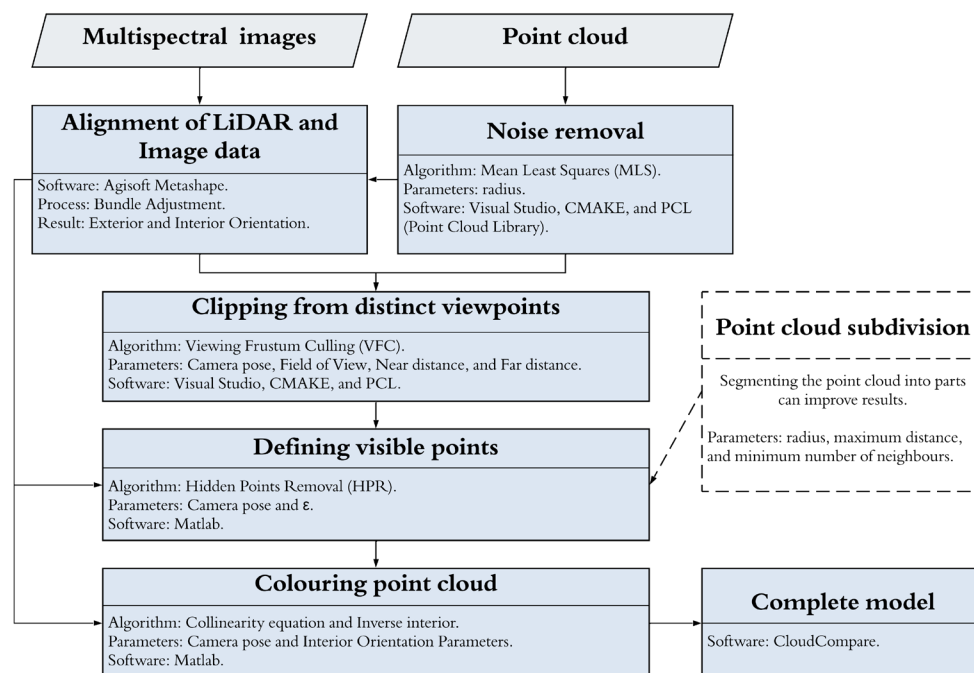


Figure 3. Flowchart of the processing sequence. Solid borders indicate the main steps of the workflow, while dashed borders represent optional processes.

Subsequently, the LiDAR point cloud and spectral images are aligned within a common reference frame, enabling each point in the 3D cloud to be associated with its corresponding spectral values. The methodology includes the identification of visible points and the detection of occlusions to prevent double mapping in overlapping areas. These steps allow the generation of a multispectral LiDAR point cloud in which each point contains both reliable geometric coordinates and the associated spectral data, making it suitable for the detailed analysis of plant traits and other applications that require combined 3D and spectral information.

3.1. Data Acquisition

Field data acquisition followed four main steps: (1) the planning and placement of coded targets to serve as GCPs, (2) camera configuration, (3) image acquisition at different stations and (4) the LiDAR scanning of the plant. The procedure differed between the two experiments (coffee and apple trees) according to the scene configuration and acquisition purpose. Table 1 summarises the main acquisition parameters for both datasets.

For the coffee tree dataset, spherical targets were initially considered for alignment purposes; however, they were not used because it was difficult to accurately identify their centres in all images, especially due to variations in the viewing angle and illumination. Therefore, planar coded targets were adopted as GCPs, as they allowed more reliable centre detection and measurement for bundle adjustment. A single LiDAR scan was performed at the front of the plant. Six central bands (490, 525, 550, 560, 570 and 850 nm) were processed using Agisoft Metashape 2.2.2 after extraction with AwBasic 1.23.5.16 (Tel Aviv, Israel) software. The image brightness was corrected in Irfanview through automatic colour adjustment and manual tuning (contrast 61, brightness 10, gamma 2.15).

For the apple tree dataset, four LiDAR scans were acquired from different positions to capture the entire plant structure. Point clouds were generated at four spatial resolutions

(102.1, 203.9, 407.8 and 814 pt/m²), but only the highest resolution was used for subsequent processing, as it preserved smaller branches. Corrections for blue and green bands were applied in Irfanview with parameters similar to those used for the coffee dataset. In both datasets, statistical filtering was applied to remove noise caused by wind effects.

Table 1. Data acquisition parameters.

Parameter	Coffee Tree Dataset	Apple Tree Dataset
Number of images	50	35
Camera-to-object distance	1.8 m	2 m
Camera orientation	Horizontal and frontal	Horizontal and convergent
Exposure settings	1/400 s (850 nm); 1/200 s (other bands)	1/500 s
LiDAR resolution (pt/m ²)	814 pt/m ²	102.1 pt/m ² , 203.9 pt/m ² , 407.8 pt/m ² and 814 pt/m ²
Weather conditions	Sunny (wind: 4.1 km/h)	Clear sky (wind: 11.6 km/h) *
Number of plants	1 coffee tree	1 apple tree
Sampling intervals (LiDAR)	One single station in front of the plant	Four stations surrounding the plant
Sampling intervals (camera)	5 height levels at 10 lateral positions with 0.35 m base	Convergent acquisition with a base length of approximately 1.8 m

* Area was surrounded by walls and less affected by wind; also, the meteorological stations were far from the acquisition area.

3.2. Filtering Noise from the LiDAR Point Cloud

Field surveys employing TLS equipment are often affected by several sources of noise, such as dust, moving humans, insects and wind, which can deform the representation of the object [30]. Moreover, with higher LiDAR resolutions in outdoor environments, the movement of non-static features caused by wind becomes more evident, often resulting in outliers due to the accumulation of multiple measurements at a single point. The need to filter out both noise and outliers to improve data quality is therefore evident. To address these challenges, different filtering and smoothing strategies have been explored, among which the moving least squares (MLS) technique has gained prominence for point cloud processing in vegetated environments. MLS is known for its ability to smooth data locally while preserving irregular characteristics [31–33]. It uses a projection operator that can smooth and estimate the surface from points in a 3D point cloud. The fitting process removes outliers and performs the smooth local segmentation of the surface.

In this work, MLS was adopted as the main statistical method for noise removal due to its implementation in the PCL [29]. MLS is described as a smoothing process over a point cloud, which involves a local polynomial fit to individual points and a simultaneous computation of the normal for each smoothed point [34]. The parameters used are consistent with the use of the 2nd-order polynomial method for smoothing the data and the use of the K-Dimensional tree for search optimisation. An empirically defined radius of 1.8 mm was used to smooth the surface; this radius is close to the laser footprint (1.2 mm) for the average range (2 m). Once the noise has been removed, the LiDAR point cloud is used to measure the GCPs, which are essential for the bundle adjustment to determine the exterior orientation parameters for each image.

3.3. Fusion of Data from Sensors

Fusing data from different sensors (camera and LiDAR) requires the data to be aligned within a common reference frame. Bundle adjustment is a method that enables this con-

catenation of data, ensuring the same reference system for both sensors. The coordinates of identified control points are then used to orient each multispectral band, with the exterior orientation parameters (EOPs) estimated through bundle adjustment and the interior orientation parameters (IOPs) refined via self-calibration within the adjustment process.

Figure 4 shows the following: (PC) represents the camera's perspective centre; (pp) represents the principal point; (o) represents the origin of the point cloud coordinate system; and (p) represents the point of interest in the image projected onto the 3D cloud (P). There are two coordinate systems: (X, Y, Z) for the point cloud, and (x, y, z) for the image.

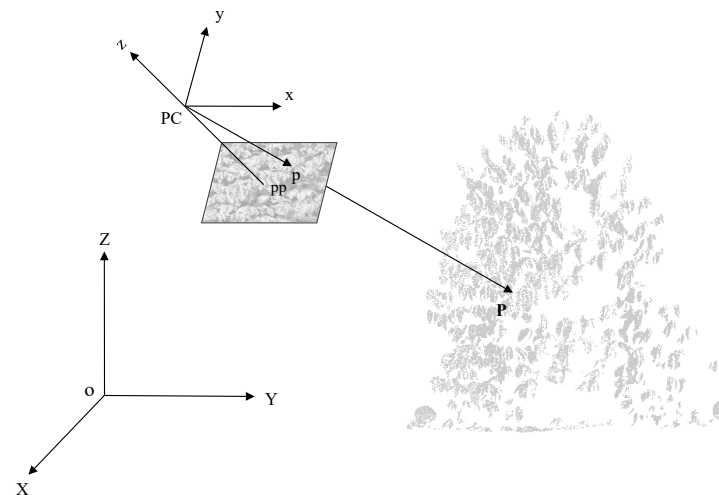


Figure 4. Visual representation of a tree point cloud. On the right, the 3D visualisation is shown; on the left, the relationship between the point in the cloud (P) and the corresponding point in the image (p) is shown, with each in its own reference frame (2D and 3D).

Before assigning spectral values to the LiDAR point cloud, double mapping must be avoided. Therefore, a preliminary step is required to label occluded points in the LiDAR point cloud with respect to each specific image. This topic is explored further in the following sections, which present methods for detecting occluded points based on point cloud visibility analysis.

3.3.1. Visibility Maps

The assignment of spectral values from an image onto a point cloud requires the accurate position and orientation of the observer (defined by the EOPs of the images) relative to the 3D point cloud. For each image viewpoint, the point cloud is evaluated to determine which points are visible. Points within the camera's field of view are considered visible, avoiding overlapping information from multiple perspectives and ensuring consistency in the 3D representation. A point is considered visible if it is not occluded by other points or objects. To support this process, two methodologies were applied. VFC acts as a clipping of viewpoints in the point cloud, reducing the computational cost by filtering points outside the camera's visible volume. HPR labels occluded points based on the camera's perspective, preventing double mapping and ensuring the proper assignment of spectral values to the points.

Clipping Viewpoints in the Point Cloud

The viewing frustum is a geometric representation of the camera's field of view in 3D space. It is defined as a volume bounded by six planes: four planes corresponding to the sides of the frustum (left, right, top, bottom), which are determined by the camera's horizontal and vertical FoV, and two planes parallel to the image plane (near and far

clipping planes), perpendicular to the optical axis. The frustum can take a perspective shape (truncated pyramid) or an orthogonal shape (box), representing the camera's viewing volume in 3D space, as illustrated in Figure 5 for a perspective shape.

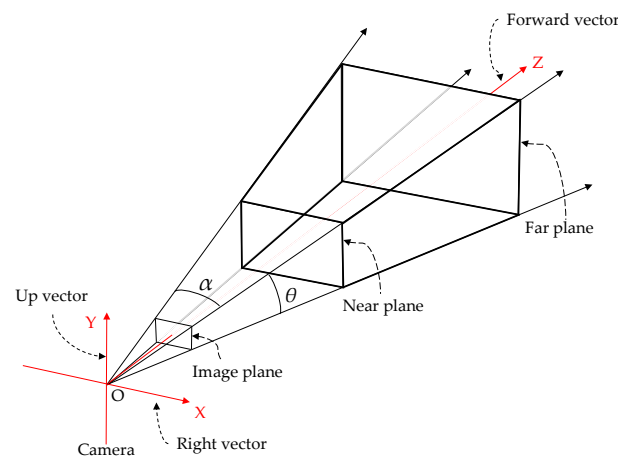


Figure 5. Schematic of the VFC shape: axes (X, Y, Z), camera perspective centre (O), image plane, near plane, far plane and horizontal (α) and vertical (θ) FoV. Solid lines represent the main schematic of the VFC, while dashed lines highlight auxiliary elements of the VFC for illustration purposes. Black labels refer to elements of the VFC, whereas red labels indicate the coordinate axes.

VFC algorithms use this frustum to discard points outside the camera's visible volume [21]. Conceptually, this process does not permanently remove points from the point cloud; rather, it labels points as occluded for a given camera viewpoint. By restricting further processing to only those points within the frustum, VFC reduces the computational complexity while ensuring that only potentially visible points are considered for subsequent steps, such as spectral value assignment.

According to [21], the VFC algorithm eliminates nodes from the scene graph, which is a hierarchical tree-like structure used to organise data (objects) in 3D computer graphics models. The tree is a directed acyclic graph composed of nodes and edges, where the root node (parent) serves as the basis of the structure by grouping all objects in the environment, and the child nodes organise the geometry of objects. With the addition of more child nodes to the scene graph, a more detailed representation of the scene is achieved. Each child node represents an object containing a bounding volume (BV). The BV of the node is tested against the frustum: if it is entirely inside, it is rendered; if it is entirely outside, it is discarded.

The main VFC algorithms described in the literature are those proposed by [35,36]. The bounding volume culling technique is addressed in two ways to solve the problem: first, by applying the perspective transformation (view and projection matrices) to the BVs, defined by either axis-aligned bounding boxes (AABBs) or oriented bounding boxes (OBBs) [21], and performing intersection tests between the transformed BVs and the viewing frustum; and second, by directly testing each BV against the frustum's six planes through acceptance/rejection intersection checks.

The usual intersection test to check visibility, based on the first approach and widely used in the PCL [37], involves the vectors represented by the axis directions: forward (Z), right (X) and up (Y) (Figure 5). First, the forward vector is tested against the near and far limits. Then, the right vector is checked to determine visibility to the left and right of the camera. Finally, the up vector is tested to verify visibility in the upper and lower parts. Points outside the field of view are discarded.

The implementation of the VFC algorithm follows the approach of [38], adapted for the step of defining the camera orientation, considering a rotation around the y-axis of 90° (φ angle). The other parameters required for the VFC operator are summarised in Table 2.

Table 2. VFC parameters.

Parameter	Description	Value	Source
Camera pose	Perspective centre	$(X_0, Y_0 \text{ rotated by } -90^\circ, Z_0)$	Exterior orientation parameters
FoV	Horizontal and vertical	H: 25° V: 25°	Camera specification
Near plane distance	Minimum distance from the camera to the frustum	1 m (coffee) 1 m (apple)	>0 m to define frustum shape [37]
Far plane distance	Maximum distance from the camera to the frustum	4 m (coffee) 3 m (apple)	Empirically defined based on scene depth

Labelling Occluded Points

The HPR operator stands out among the visibility map algorithms due to the fact that it does not rely on additional information from the dense point cloud. The HPR operator can efficiently determine visible points by extracting points located on the convex hull [28]. The theoretical basis considers a set of points P , representing a sampling of the continuous surface (S) viewed from the viewpoint (C). Thus, the operator determines, in two stages, whether a point (p_i), $\forall p_i \in P$, is visible from C (observer/camera position). The first stage is to apply a spherical flipping function in which all points undergo an inverted projection relative to the sphere with a centre C along the radius R , with the new position being defined proportionally to the distance from p_i to C . This operation is described by Equation (1), and the new coordinates are referred to as transformed coordinates:

$$f(\hat{p}_i) = p_i + 2(R - \|p_i\|) \frac{p_i}{\|p_i\|} \quad (1)$$

where p_i represents a point in the point cloud; R is the radius of the sphere; and $\|p_i\|$ is the norm of the position vector of the point in the cloud. The spherical flipping function calculates the norm ($\|p_i\|$) of the position vector for all points in the 3D point cloud (the distance from the point to the centre (C) of the projection). The radius (R) is determined based on the largest norm value of the points, which represents the point farthest from the centre. In addition, R is extended by a parameter (ϵ) ranging from zero to infinity, as used in Equation (2). The parameter ϵ in the spherical flipping algorithm controls the effective radius used to project points onto the convex hull. Using the default value ($\epsilon = 0$), as suggested by [28], is generally suitable for sparse point clouds, where points are already well separated. For denser point clouds, however, a larger ϵ is preferable, as increasing the radius projects the 3D points farther onto the sphere, ensuring that they lie on the convex surface. Furthermore, the ϵ parameter should be adjusted according to the scale of the scene and the average distance between the camera and the point cloud to ensure that visible points are correctly identified, according to [39]. The inverse transformation is computed considering that the points are in a valid visibility order for a given viewpoint of the point cloud:

$$R = \max \|p_i\| * 10^\epsilon \quad (2)$$

The geometric transformation represented in Equation (1) first computes the distance from the point p_i to the sphere surface ($R - \|p_i\|$). Multiplying this difference by 2 amplifies the displacement, ensuring that the transformed point is shifted beyond the sphere surface by the same amount. The ratio $\left(\frac{p_i}{\|p_i\|}\right)$ normalises the point vector, defining the outward direction of the displacement. Finally, the calculated vector is added to the original point

(p_i) to obtain the transformed point ($\hat{p}_i$). In effect, the transformation mirrors the point across the sphere, positioning it at an equal distance on the opposite side of the surface.

The implementation of these algorithms follows the approach described by [28]. Given the assumption that all points in the convex hull are visible, HPR can label the occluded points of the LiDAR point cloud. The computation of the convex hull can be performed through the utilisation of various functions, including `convhull`, `convexHull` and the `alphashape` version [40] implemented in MATLAB. As observed in the initial set of data, the `convhull` method was found to be more efficient when applied to scenarios involving data collection from a single frontal scan. However, in scenarios involving multiple scans or irregular geometry, such as those observed in the second dataset, complementary segmentation techniques become necessary. These techniques subdivide the point cloud into smaller subparts, enhancing the efficiency of the HPR operator.

3.3.2. Assigning Spectral Values to the Point Cloud

This process involves obtaining the image orientation data, including the interior orientation, exterior orientation, frame size (height and width) and points visible in the object space from the viewpoint (coordinates of the camera's perspective centre), followed by the assignment of spectral values to the points. Given a point with X, Y, Z coordinates, the corresponding photogrammetric image coordinates (x, y) can be calculated using the collinearity equations (Equations (3) and (4)):

$$x = -f \frac{m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)}, \quad (3)$$

$$y = -f \frac{m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)}. \quad (4)$$

The variables x and y represent the coordinates of the points in the image in the photogrammetric reference frame; f is the focal length of the camera; X, Y and Z are the coordinates of the point in the object space; m_{ij} represents the elements of the rotation matrix; and X_0, Y_0 and Z_0 are the coordinates of the camera perspective centre (PC) in the object space reference frame.

Next, the image coordinates (columns— c ; and rows— r) are obtained from x and y by the inverse process of interior orientation calculation, with the addition of lens distortions. Finally, to associate the DN value with the point coordinates, a bilinear interpolation is performed based on the pixel coordinates (c, r). For each point of the LiDAR point cloud, the inverse interior orientations are calculated for all points of the LiDAR point cloud.

Inverse interior orientation iterates when calculating lens distortions, allowing the refinement of radial symmetric and decentring distortions before the calculation of image coordinates, since the input argument is the radius calculated with the distorted coordinates. If the estimates of the IOPs are inaccurate, it is possible to reduce the size of the LiDAR cloud clipping based on smaller subpart segmentations (Section Subdivision of the Point Cloud). Thus, the cutout is limited to the central area of the image where the effect of radial symmetric distortion is reduced. Finally, bilinear interpolation is performed to ensure that DN values are interpolated for all coordinates present at image boundaries (width and height) based on the pixel neighbourhood.

Subdivision of the Point Cloud

To segment the dataset into smaller subparts, an iterative approach based on the spatial proximity between points was used. First, the 3D point cloud is loaded, and the search radius is defined, along with the maximum allowed distance between points, which sets a criterion for the minimum number of neighbours to determine the neighbourhoods.

These parameters are adjusted empirically according to the characteristics of the point cloud being analysed.

In each iteration, a reference point is selected, and the Euclidean distances between this point and others in the cloud are calculated. Additionally, points within the radius and maximum distance are considered as potential neighbours. As the process evolves, the point cloud undergoes a progressive reduction through the removal of distant neighbours at each iteration. This iterative process is designed to ensure the efficient segmentation of data, thereby isolating groups of points based on their proximity. This approach promotes local analysis and avoids overlap between regions that have been processed. The following criteria were implemented: a search radius of 1 m, a maximum distance between points of 0.3 m and a minimum number of neighbours of 50.

3.3.3. Radiometric Discrepancies

Variations in light incidence caused by the different capture angles of the frames obtained with the optical camera result in discrepancies in DN values within overlapping areas. These discrepancies make the direct integration of DN values into a complete point cloud model difficult, making radiometric adjustment an essential step to ensure the spectral consistency required for constructing the final multispectral point cloud [1]. These discrepancies render the direct integration of DN values into a comprehensive model of the point cloud challenging. Consequently, radiometric adjustment assumes a fundamental role in ensuring the spectral consistency for constructing the final multispectral point cloud.

The generation of the multispectral point cloud is initiated by the process of cropping the point cloud using the VFC method. Subsequently, the labelling of occluded points is executed through the HPR operator. Then, each point of the 3D cloud is assigned to a DN based on the image corresponding to the viewpoint. In instances where there was overlap between different viewpoints, the CloudCompare 2.14 alpha (Paris, France) software was used to remove these overlapping areas, thereby avoiding the need to adjust the DN values of the different viewpoints.

Nevertheless, in the second set of data (potted apple tree), which involves a 360° data collection process around a tree, the overlap is significantly greater and more complex to identify. In this scenario, simple cropping is inadequate to ensure the proper integration of the viewpoints. Consequently, radiometric adjustment is essential to adjust the DN values from the different viewpoints, thereby generating a full multispectral point cloud model. This step was not implemented and is left as a suggestion for future work.

4. Results

4.1. Bundle Adjustment

Bundle adjustment for each reference band was computed to make data from different sensors compatible, resulting in an estimate of the camera position and orientation for each lens using the reference image for each sub-image for the first and second datasets. The pixel size in object space units (PSOSU) of the first dataset was 0.3 mm, obtained with the camera positioned at a distance of approximately 1.80 m from a coffee tree. Eleven control points (Figure 6a) were distributed between targets on the tree and on the ground, with additional measurements made on branches and leaves.

The total root mean square error (RMSE) of the GCPs for the central bands of each sub-image generally exceeds the PSOSU value by a factor of 14, indicating that the estimated orientation of the images is affected by the low precision of the points, depending on the planimetric and altimetric PSOSU. However, sub-image S4 exhibited good results, with individual RMSEs of 1.59 mm, 2.06 mm and 1.13 mm in X, Y and Z, respectively, resulting

in a total RMSE of 2.84 mm. The other sub-images showed higher RMSE values, with averages exceeding 3.9 mm.

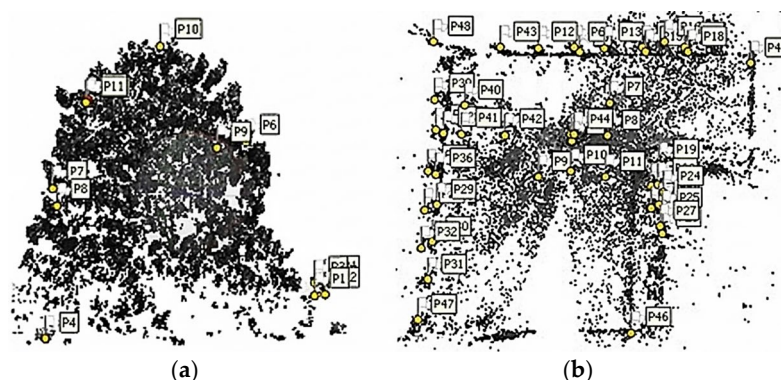


Figure 6. Distribution of control points for the (a) coffee tree and (b) apple tree datasets.

Due to the lack of targets and static features in the environment, the result was not sufficient to generate a multispectral point cloud with all bands, as revealed by the visual analysis of the coloured point cloud. However, sub-image S4, with wavelengths of 430 nm, 550 nm and 650 nm (Figure 1b), could be used because the interior and exterior orientation values enabled coherent colourisation. Improving the results mainly depends on increasing the measurement of GCPs. In the second dataset, more points were measured on targets in the scene, leading to a better overall accuracy (Figure 6b).

In the second dataset (apple tree), images taken around the plant at approximately 2 m of camera-object distance led to a PSOSU of 0.35 mm. Some of the GCPs (44 points) (Figure 6b) and four checkpoints were located along the walls and the floor. The RMSE at the control points (Table 3) after bundle adjustment for the central bands of each sub-image ranges from 1.8 to 2.4 times the PSOSU for the planimetric coordinates, while for the altimetric component, it ranges from 1.3 to 2 times the PSOSU. Furthermore, the total RMSE does not exceed four times the PSOSU. In comparison to the first dataset (coffee tree), the self-calibrating bundle adjustment of the second set achieved higher accuracy for the analysed points, with the planimetric and altimetric RMSEs expressed as multiples of the PSOSU.

Table 3. Total RMSE of the GCPs in the second dataset.

Sub-Images (Bands)	X (mm)	Y (mm)	Z (mm)
S1 (570 nm)	0.70	0.84	0.53
S2 (525 nm)	0.65	0.84	0.67
S3 (850 nm)	0.80	0.75	0.58
S4 (550 nm)	0.68	0.85	0.67
S5 (490 nm)	0.84	0.63	0.70
S6 (560 nm)	0.82	0.87	0.45

4.2. Point Cloud Registration

Both datasets were acquired at a LiDAR spatial resolution of 814 pt/m², which supports the definition of high resolution adopted in this study. This configuration enables the representation of small branches and fine structural details, providing a level of geometric detail suitable for close-range vegetation analysis.

For the apple tree dataset, the influence of the point density on the alignment accuracy was assessed by analysing registration errors, in terms of the mean absolute error (MAE),

for point clouds acquired at four different resolutions (Table 4). The coffee tree dataset was obtained from a single LiDAR scan and, therefore, did not require registration between multiple scans. The results demonstrate that increasing the point density directly improves the registration quality, as higher densities provide more stable points on rigid surfaces (e.g., walls and floor), reducing both horizontal and vertical errors.

Table 4. MAE registration errors for point clouds acquired at different resolutions (apple tree dataset).

Resolution (pt/m ²)	MAE XY (mm)	MAE Z (mm)
102.1	4.8	1.3
203.9	3.5	1.1
407.8	4.6	0.6
814	2.0	0.4

The configuration with 814 pt/m² not only produced the lowest registration errors but also enhanced the visualisation of fine-scale features such as small leaves and thin branches, which is crucial for accurate multispectral colourisation and detailed plant-level analysis.

4.3. Point Cloud Clipping

In the first dataset, the optical images of the coffee tree covered a narrow volume of the tree. On the other hand, the images in the second dataset covered the entire apple tree from top to bottom, enabling the generation of approximately 700,000 points. The VFC algorithm was used to crop the coffee tree point cloud, reducing it from 5 million to about 800,000 points per image, reducing the computational cost. Only eight images were needed to fully colourise the model, even though 35 images provided an ideal overlap; all 35 images were used for self-calibrating bundle adjustment.

The EOPs of 550 nm-band images (S4) confirmed that the position and attitude of the acquired images were well aligned with the viewer's position (Figure 7). The algorithm generated consistent results for the other camera's sub-images, with only small variations among the sections, which may have been caused by the parallax among the sub-images. These results indicate that even with a reduced number of GCPs, the EOPs of the sub-images were well defined. However, the other sub-images have been discarded due to the unreliability of their EOPs for assigning spectral values to the 3D points.

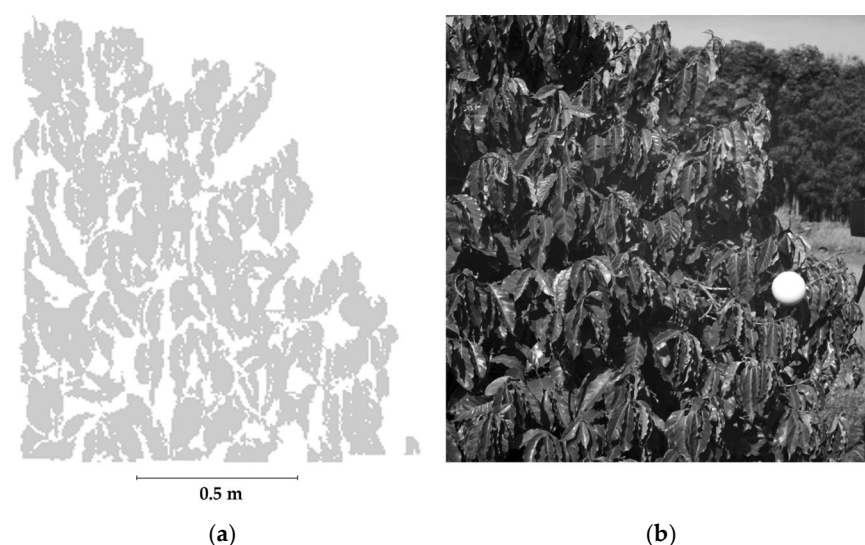


Figure 7. Cropping point cloud with VFC: (a) cropped point cloud and (b) corresponding image frame.

Figure 7a illustrates a cropped point cloud generated by the VFC algorithm from the original dataset, while Figure 7b shows the limits of the image whose parameters define the viewing frustum. The next step involves creating a visibility map for the point cloud. Thanks to the reduced number of points, the computational time decreased from approximately 5 h to only 20–23 min, representing a 14-fold reduction on average, depending on the remaining points after clipping.

4.4. Labelling Occluded Points

After applying VFC, some LiDAR points remained occluded, as they were not visible from the optical sensor's perspective. The HPR operator was then used to identify and label these occluded points. In this scenario, involving coffee tree canopies with closely spaced leaves and branches, the HPR parameter ϵ (Equation (2)) was set to 4 (Figure 8), according to the criteria described in Section 3. Since the average distance between the viewpoint and the point clouds was similar in both datasets, the ϵ value remained appropriate, with no significant variation. The only parameter that needs to be set by the user for HPR is ϵ , while other parameters are calculated. In Figure 8b, the leaf comprised 2750 points, which were reduced to 2190 points in Figure 8c.

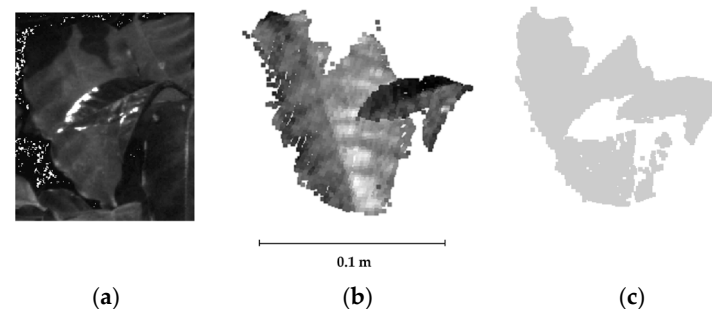


Figure 8. Example results from the first dataset: leaves shown in (a) the image, (b) the raw point cloud and (c) the post-processed point cloud.

The processing pipeline used on the first dataset was repeated for the second one. The orientation of the image sequences varied during data collection. In contrast to the uniformity observed in the first dataset, the plant analysed has an irregular behaviour. Varying the viewpoint revealed the different levels of features (branches, leaves and pot) in the point cloud, and it was decided to subdivide them into smaller parts. Nevertheless, if the images were taken with a shorter distance between the camera and the object, the use of the VFC technique could replace the subdivision technique and overcome the problem of irregular data. The HPR results for the data are shown in Figure 9.

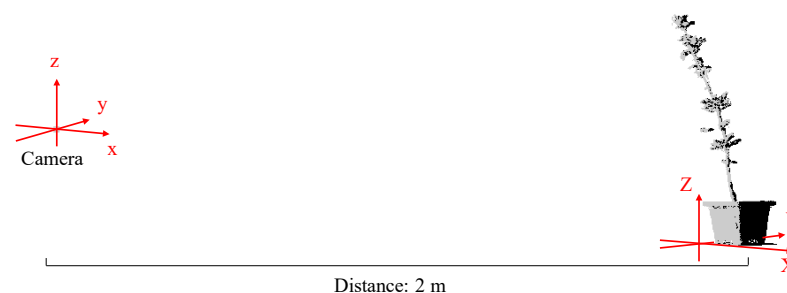


Figure 9. VFC clipping with $\epsilon = 4$ of the second dataset, showing occluded (black) and visible (grey) points.

In Figure 9, the coordinate system of the camera refers to the EOP data based on Agisoft Metashape 2.2.2 software, while the point cloud is expressed in the scanner's local

(arbitrary) coordinate system. The ideal parameter, ε , did not change in the second dataset, indicating that it was selected based on the volume of data rather than the specific structure of the elements studied. Despite this, some visible points of vegetation were occluded, as was the case in the first dataset. The occurrence of visible points considered occluded was lower in the pot bounded by regular faces.

4.5. Colouring

Assigning spectral values to LiDAR cloud points necessitates projecting the 3D points onto the images and interpolating the DN values. Eight images from the S4 sub-image were used to colourise the point cloud (Figure 10) generated from the coffee tree data. The point clouds clipped by the VFC algorithm showed a high level of overlap.



Figure 10. LiDAR RGB (430 nm, 550 nm, 650 nm) cloud for the coffee tree.

Some targets shifted between the LiDAR and camera acquisitions, and this positional difference is clearly visible in the coloured 3D point clouds (Figure 11). These targets were excluded from the bundle adjustment process. The small displacements observed in the leaves can be attributed to wind-induced movement, given the non-synchronous data collection of the camera and LiDAR.

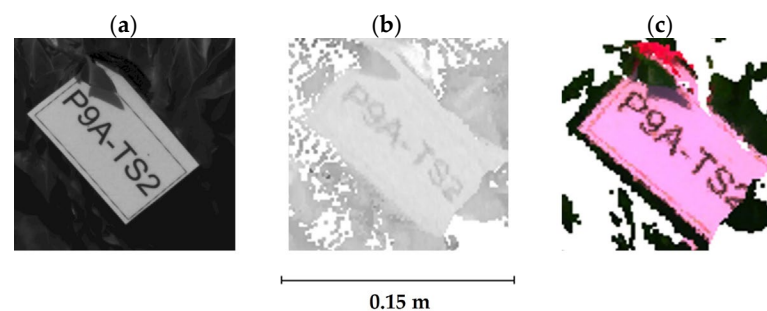


Figure 11. Example of a target that changed position between the LiDAR and camera surveys: (a) the captured image, (b) the point cloud and (c) the colourisation.

Comparing Figure 11a with Figure 11b, it can be seen that the target is slightly rotated downward, resulting in the displacement shown in Figure 11c. The observed discrepancy in colourisation can be investigated in future work to identify correction techniques. Clipping reduction was tested to improve colourisation, as described in Section Subdivision of the Point Cloud for the apple tree. Figure 12 shows an image (Figure 12a) used for colourising a VFC clipping (Figure 12b), along with a version with smaller clippings (Figure 12c).

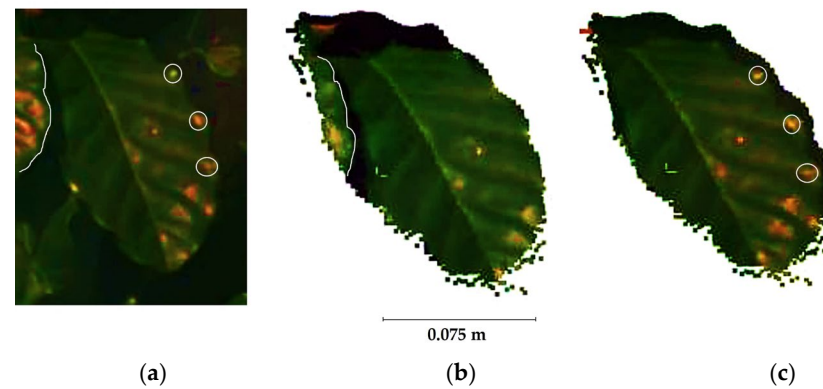


Figure 12. (a) Original image with the area of interest highlighted; white circles indicate spots on the leaf visible in the original image, and white lines show the leftmost leaf. (b) Colourised 3D cloud from the VFC cutout, showing a rightward shift in the colourisation that causes the leftmost leaf to invade the adjacent leaf, as indicated by the white lines. The spots visible in (a) are not seen here due to this displacement. (c) Colourised 3D cloud with smaller cutouts, where the spots on the leaf reappear, illustrating the effect of reducing the cutout size on correcting the colourisation displacement.

The results revealed some inconsistencies, as shown in Figure 12a, where the outlined top leaf indicates a potential overlap with the underlying leaf, likely caused by errors in the image orientation. However, the area of the bottom leaf, where the overlap occurs, was not fully reconstructed by the TLS scan due to occlusion. Nevertheless, this area is visible in the coloured point cloud of Figure 12b. By segmenting the point cloud into smaller pieces, the result of the colouring revealed some clear patches (highlighted in Figure 12a,c), which are also visible in the original image (circled in white). In Figure 12c, the location of the spot should be closer to the edge of the leaf; however, it has been shifted, as discussed in Section 5.

These problems may have been occurred due to the in situ determination of the interior orientation, which has been partially mitigated by the subdivision of the point cloud into smaller sections. However, possibly due to biases in the determination of the exterior orientation with bundle adjustment and wind effects, there are still some displacements at the boundary of the leaf. In addition, radiometric adjustment was not possible due to inconsistencies in the overlap of assigned DN values, which affected the efficient concatenation of the data. Figure 13 shows leaves from images used in the colourisation.

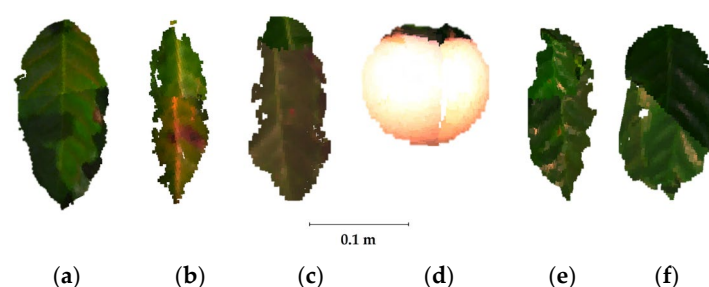


Figure 13. Colourised leaves in overlapping areas between different images: (a–f) examples of overlapping regions showing how colourisation varies depending on the contributing images.

Overlaying the images from different points of view resulted in significant geometric shifts and DN discrepancies, except for the example shown in Figure 13a, which is the least affected. In the future, it is recommended to develop techniques to reduce the geometric shifts and the radiometric discrepancies. Processing the second set of data resulted in a better quality for the colourisation (Figure 14), due to the greater number of control points

measured, which makes it possible to obtain better-quality IOPs and EOPs with lower RMSEs at the control points.

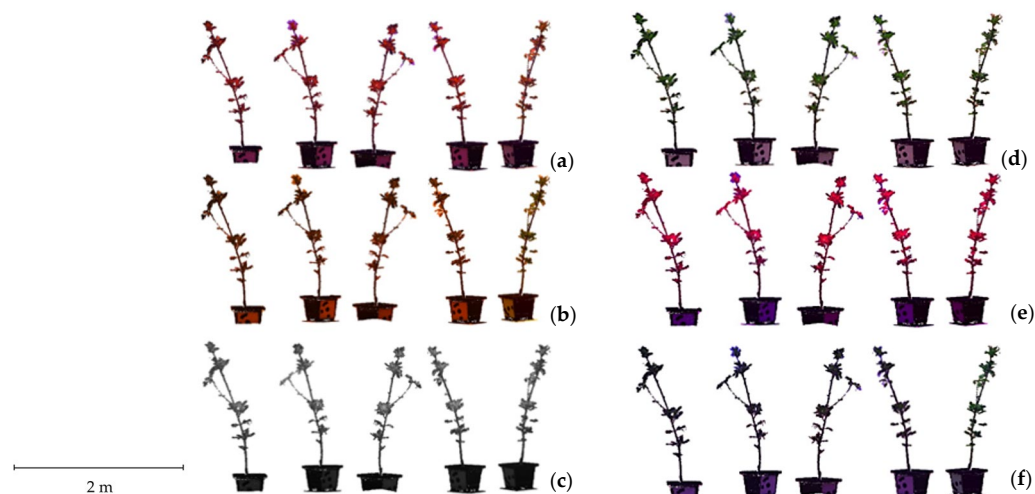


Figure 14. Multispectral 3D clouds for sub-images: (a) S1, (b) S2, (c) S3, (d) S4, (e) S5 and (f) S6.

The visual analysis of the results confirmed correct point colourisation in all cases. To identify potential problems, the discrepancies in both geometry and radiometry were assessed in the overlapping areas of the images (Figure 15).

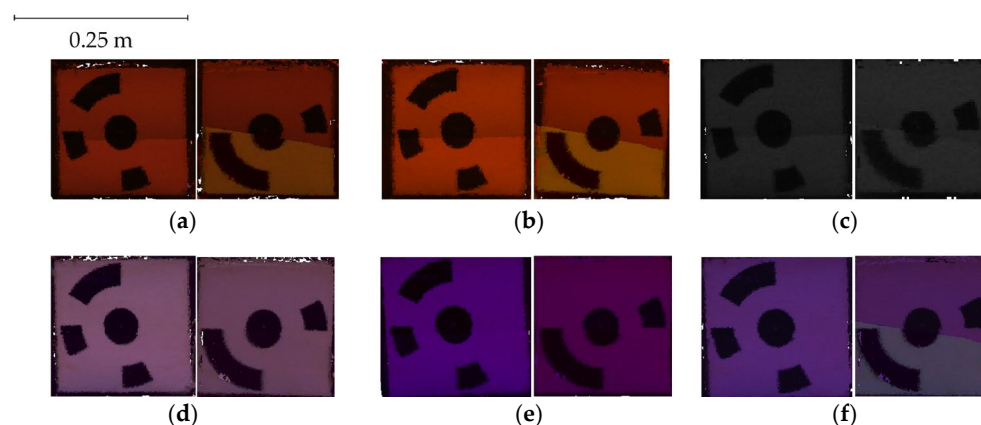


Figure 15. Point clouds of targets coloured with DN values from overlapping multispectral point clouds: each subfigure (a–f) shows targets coloured using pairs of sub-images from different view-points (S1–S6), illustrating how the combination of overlapping multispectral point clouds affects the result.

Figure 15a–f illustrate the geometric discrepancies in the point clouds generated from different viewpoints. Colourisation with sub-images S2 (Figure 15b), S4 (Figure 15d) and S5 (Figure 15e) resulted in negligible displacements (0–1 mm), indicating high consistency between the overlapping point clouds. In contrast, when using sub-images S1 (Figure 15a), S3 (Figure 15c) and S6 (Figure 15f), larger displacements were noted between the left and right viewpoints, reaching approximately 5.8 mm, 9.6 mm and 7.1 mm, respectively. These variations reflect the sensitivity of the alignment to small orientation differences between sub-images. Among all the sub-images, S4 (550 nm) stands out as the most stable configuration for multispectral colourisation.

The quantitative measurements of these types of discrepancies are discussed in Section 5. Furthermore, due to the environmental wind, the positions of the leaves at the top of the tree were the most affected in the colourisation step (Figure 16a–f). In S1 (Figure 16a)

and S5 (Figure 16e), colour saturation is strongest at 710 nm and 735 nm, respectively. Possible leaf movement can be seen, as this phenomenon does not occur in other views, as shown in Figure 16g–l.

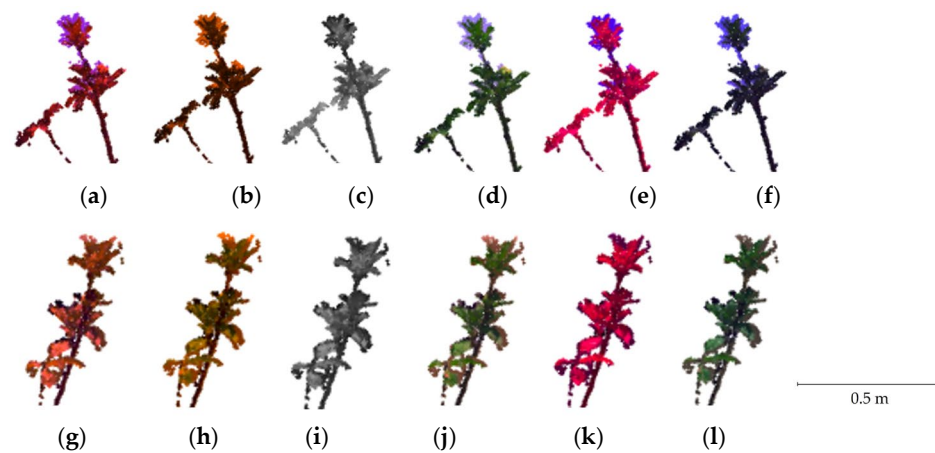


Figure 16. Leaves at the top of the tree coloured using sub-images from different viewpoints: (a–f) S1 to S6, showing the effect of wind-induced movement; (g–l) S1 to S6, showing the same leaves from a different viewpoint without wind effects.

Moreover, due to the influence of wind and the presence of small leaves on the plant, the three-dimensional reconstruction using LiDAR data showed inferior results compared to the first dataset. Fine details of the surface, such as subtle variations in height and roughness, were not captured due to the reduced density of points on the leaves, which affected the reconstruction of the plant's texture. It is recommended to avoid windy periods to improve the quality of the reconstruction.

4.6. Complete Model

Building the full model ideally should include the radiometric adjustment of the data coming from different sub-images. However, as shown in Figure 13, there are shifts in the DN's coming from overlapped images when values are assigned to the 3D points. To mitigate this problem, it is possible to crop the overlapping areas (Figure 17a) while partially preserving the individual contributions of the images (Figure 17b), which are represented by different colours.

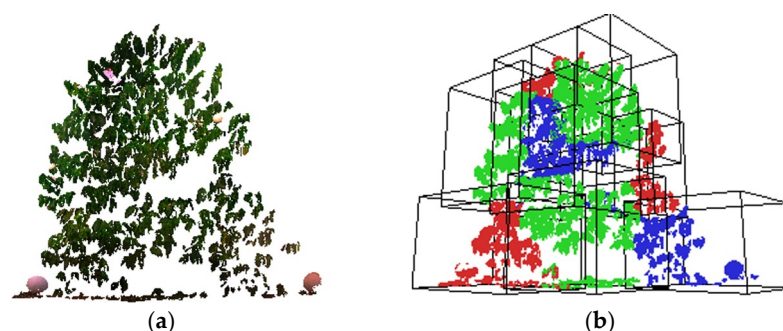


Figure 17. Complete model of the (a) coloured point cloud (S4) and (b) colours representing the individual contributions of the images, where the red, blue, and green colours are used to distinguish the non-overlapping image regions.

Due to the data collection method, this approach is only feasible for the first dataset. Coffee plants do not allow 360-degree variations in the camera position, which is common with individual plants. On the other hand, for the second dataset, it is recommended to

apply a radiometric correction to reduce discrepancies caused by different viewing angles for each image.

5. Discussion

For the first dataset, the LiDAR point cloud was coloured with RGB bands (S4: 430 nm, 550 nm, 650 nm) due to the insufficient number of sub-images available with orientation parameters compatible with the process. A primary challenge was identifying the causes of the lack of accuracy in the EOPs of the images, despite this orientation being adequate for generating the sections (VFC). The prevailing hypothesis is that unsuitable self-calibration is due to the reduced number of GCPs. To investigate this hypothesis, a comparison was performed between dense point clouds generated by Agisoft Metashape and the LiDAR point cloud. In Metashape, optical point clouds are reconstructed through a photogrammetric workflow, which includes image alignment, camera calibration and dense stereo matching to generate 3D coordinates from overlapping images. This comparison was intended to determine whether the shift observed in the colourisation has also occurred between the dense point clouds. This was performed to ensure an adequate distribution in the displacement analysis (Table 5).

Table 5. RMSE of discrepancies between LiDAR coordinates and dense band cloud in coffee tree.

Sub-Images (Bands)	RMSE X (mm)	RMSE Y (mm)	RMSE Z (mm)
S1 (570 nm)	9.7	9.9	4.8
S2 (525 nm)	4.5	25.6	5.9
S3 (850 nm)	28.6	17.3	5.6
S4 (550 nm)	5.1	5.8	2.1
S5 (490 nm)	5.5	8.3	4.7
S6 (560 nm)	9.2	18.3	4.8

Table 5 summarises the alignment errors between the LiDAR data and the optical dense multispectral point cloud for the coffee tree. The point cloud generated with the sub-images of the 550 nm band (S4) exhibited the smallest discrepancies in the planimetric (X and Y) and altimetric (Z) components, while the other bands demonstrated substantial shifts in one or both components. The EOPs of S4 enabled the suitable colourisation of the point cloud. These findings confirm the hypothesis that interior orientation has a substantial influence on the suitable projection of points to the image for colourisation and the generation of a comprehensive multispectral point cloud for the coffee tree. The results obtained from colourising other sub-images are presented in Figure 18.

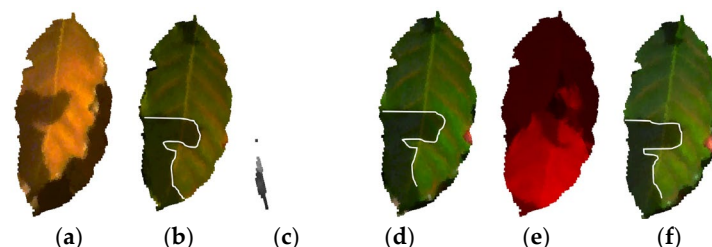


Figure 18. Coloured point cloud of a leaf for the sub-images (a) S1, (b) S2, (c) S3, (d) S4, (e) S5 and (f) S6.

In Figure 18c, the leaf was out of the field of view due to errors in the EOPs of S3 (850 nm). Similar results were obtained with the S2 (525 nm), S4 (550 nm) and S6 (560 nm) sub-images (Figure 18b,d,f). While S2 and S4 were found to be compatible (Figure 18b,d),

S6 exhibited distortions in the projected shadow (white line). However, only S4 permitted the colourisation of the point cloud based on the perspective of eight images, thereby resulting in a complete model. In Figure 19, sequential cropping by VFC for sub-images S2 (Figure 19a) and S6 (Figure 19b) is demonstrated. It is important to note that omitting an image from the sequence can prevent the formation of a complete model, since the reduced overlap between images limits coverage, resulting in missing parts of the 3D cloud.

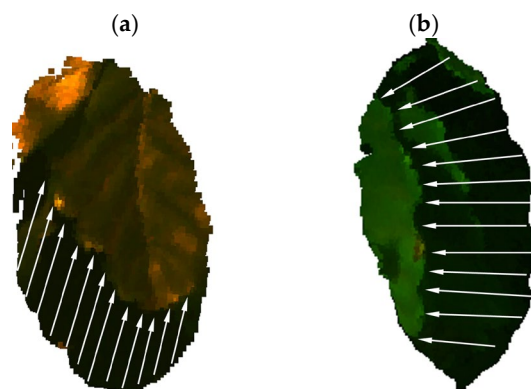


Figure 19. Example leaves for sub-images (a) S2 and (b) S6, where the arrows indicate the displacement occurred.

The colourisation of six stations was possible utilising S2 images, while four stations required the employment of S6 images. The errors in the exterior orientation of the images affected the remaining stations. The displacements of the DN values are indicated in Figure 19a (S2 displacement of approximately 14 mm) and Figure 19b (S6 displacement of approximately 22 mm). The space with points related to the leaf is coloured with DNs of the bottom of the canopy. This issue can be attributed to both the interior orientation, determined with in situ calibration, and the exterior orientation, which was limited by the reduced number of control points available for bundle adjustment, as discussed in Section 4.1. To expand upon these findings, the evaluations were expanded to encompass the second dataset. Initially, the dense point clouds generated by Agisoft Metashape were analysed (Table 6).

Table 6. RMSE of discrepancies between LiDAR coordinates and dense band cloud in an apple tree.

Sub-Images (Bands)	RMSE X (mm)	RMSE Y (mm)	RMSE Z (mm)
S1 (570 nm)	4.1	0.6	1.7
S2 (525 nm)	4.0	2.7	2.5
S3 (850 nm)	6.3	6.5	5.0
S4 (550 nm)	3.5	2.7	1.9
S5 (490 nm)	5.8	2.9	1.4
S6 (560 nm)	4.3	2.9	2.8

The analysis of Table 6 revealed that millimetric shifts were found in the point clouds (LiDAR and optical) of the apple tree. The results showed that the combination of the bands leads to a less accurate assignment of the DNs to the points, and thus, the visualisation of the multispectral point cloud in the CloudCompare 2.14 alpha software should be limited to each quadrant individually. Furthermore, one of the selected images for colourisation, depicting the posterior region of the tree, exhibited significant shifts when overlapping with another point of view of a coloured cloud (Figure 20).

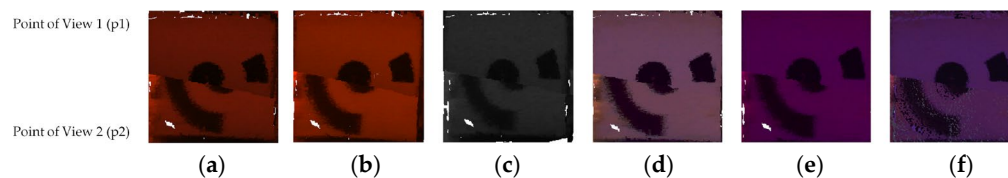


Figure 20. A marked shift between two points of view (p1 and p2) in the point clouds for the bands (a) S1, (b) S2, (c) S3, (d) S4, (e) S5 and (f) S6.

The effect, as observed from viewpoint p2, can be explained by alignment issues between the images (image orientation), leading to inaccurate overlap between point clouds captured from different viewpoints. These issues can be explained by errors in image orientation due to inadequate overlap between images, a limited number of control points and challenges in estimating the interior orientation. The images employed to densify the dense point cloud in Agisoft, for the rear part, had previously indicated significant coordinate discrepancies in relation to the reference (LiDAR). Consequently, the utilisation of the coloured point cloud derived from this image was unsuitable. Thus, the image captured in the same position during the second 360° data collection process, exhibiting a slight variation in inclination, was selected for utilisation. It was observed that some images exhibited errors during the alignment process. However, the colourisation at that position was recovered using another image, with a smaller base, to replace the affected location.

Although alignment issues and orientation errors limited the use of some coloured point clouds, these challenges did not prevent the extraction of valuable information. By adopting alternative images and adjusting the processing workflow, it was possible to recover spectral attributes and maintain the integrity of the dataset. In fact, our recent research [41] has shown that, once properly processed, multispectral point clouds can provide highly relevant results for plant-level assessment in digital agriculture. Building on this type of workflow, multispectral data integration enables the precise discrimination of leaves, wood and fruits, illustrating how the combination of geometric and spectral information supports high-resolution plant analysis and critical tasks such as growth monitoring, leaf area estimation, the detection of reproductive structures and the assessment of plant health. These capabilities are particularly valuable in digital agriculture, where accurate plant-level information can optimise resource use and improve overall crop management strategies.

Previous studies [42] have demonstrated the benefits of combining LiDAR with satellite-based multispectral imagery, which offers wide spatial coverage and accessibility. Nevertheless, these medium-scale approaches often struggle to capture detailed variations at the individual plant level. By contrast, terrestrial multispectral point clouds provide the necessary resolution to capture fine geometric and spectral characteristics, enabling more precise analyses at the plant level.

6. Conclusions

In this study, a methodology for colourising terrestrial LiDAR point clouds was developed and evaluated. The results confirm the objectives of this study, demonstrating that the proposed workflow effectively integrates LiDAR and multispectral data, reduces the computational cost and mitigates occlusion issues, enabling the generation of consistent and high-resolution multispectral point clouds. The proposed methodology integrates three-dimensional data with spectral information to create a multispectral model, thereby addressing the challenges associated with fusing data from different sensors. The proposed methodology involves several steps, including the projection of the DNs from the images to the LiDAR points, the utilisation of the HPR operator to remove occluded points and the implementation of the VFC method to optimise computational performance by reducing

the number of points to be processed. The efficacy of this approach is evidenced by a substantial enhancement in processing efficiency; it achieved a 14-fold improvement compared to conventional processing methods that do not employ minor segmentations. The conclusions derived from the experiments with two datasets indicate that the results are susceptible to the influence of wind conditions, which were found to have a substantial impact on the outcomes. Nevertheless, the results obtained were suitable for the generation of multispectral point clouds.

The initial data collection process faced substantial challenges, primarily due to a reduced number of control points, which affected the precision of image orientation. It is noteworthy that the S4 sub-image was not affected by these issues. Small displacements of the targets utilised as control points, resulting from human interference or natural factors, also contributed to the reduction in accuracy. Avoiding the use of spheres as GCPs proved to be advantageous, mitigating additional problems. Despite the substantial overlap between images, it was found that this was inadequate to replace images with incorrect orientations or distortions. Regarding the LiDAR data collection, even in the presence of wind interference, the capture of leaves and branches was effective, thereby ensuring a comprehensive geometric description. Considering the challenges identified during the initial survey, a second survey was conducted to address the limitations and identify more effective solutions.

Although the second dataset reduced some of the limitations of the first one, some challenges persist, particularly in the collection of LiDAR data. The strong influence of the wind and the small leaf size resulted in a less detailed LiDAR point cloud, affecting the reconstruction of subtle variations in the surface, such as small elevations and roughness, and affecting the accuracy of the texture representation. The application of the HPR operator required the division of the dataset into smaller components due to the irregularity of the leaves and branches, in contrast to the uniformity of the geometry observed in the first dataset. The implementation of the VFC method, which involves the prior segmentation of the data according to the size of the image, could be further improved by reducing the distance between the camera and the object. These issues demonstrate the need for improving the data collection configuration to optimise the quality of the models generated. Notably, the data collection strategy was identified as the most effective approach for generating an accurate multispectral point cloud.

The absence of prior camera calibration emerged as a key issue, making it necessary to define a strategy in scenarios with a reduced number of GCPs. It is therefore recommended that camera calibration be implemented prior to data collection, considering the inherent challenges of obtaining sufficient static points in vegetated environments. Additionally, the observed discrepancies in the colourisations of some point clouds, potentially arising from errors in the bundle adjustment, showed the need for a more detailed analysis of 360° data collection for future research and the refinement of the technique. Collecting with greater overlap between images, or repeating the process in 360° around the plant, can offer suitable solutions to the problems of discrepancies in the images, as shown for both sets of data analysed. For future experiments, we recommend the following improvements: a priori calibration of the camera, improvement of the geometric distribution of the control points, increasing the overlap between images or repeating the data collection with minor variations in inclination, utilisation of a higher-resolution configuration of the LiDAR acquisition device, even with the high number of spurious points, and employment of VFC as a tool to reduce computational costs for large datasets.

This study presented a processing flow for fusing data from different sensors using a multispectral camera and integrating three-dimensional data to generate multispectral point clouds. A future challenge relates to the radiometric adjustment necessary to achieve

a comprehensive model, which requires small shifts between images. This was successfully demonstrated in the second dataset and presented as an ideal acquisition model. Furthermore, the exploration of more advanced radiometric adjustment techniques, such as the use of the bidirectional reflectance distribution function (BRDF), can facilitate the collection of the complete model and the generation of a final product. An additional step of applying the empirical line will be necessary to transform the DN values to ensure a closer representation of the spectral response of the targets. This research has the potential to significantly impact the field of phenotyping, resulting in a final product with a high spatial resolution.

Author Contributions: Conceptualisation, I.S.N. and A.M.G.T.; Methodology, I.S.N., A.M.G.T. and M.H.S.; Validation, I.S.N.; Formal Analysis, I.S.N., A.M.G.T. and M.H.S.; Investigation, I.S.N. and A.M.G.T.; Data Curation, I.S.N. and A.M.G.T.; Software, I.S.N.; Writing—Original Draft Preparation, I.S.N.; Writing—Review & Editing, I.S.N., A.M.G.T. and M.H.S.; Supervision, A.M.G.T. and M.H.S.; Project Administration, A.M.G.T.; Resources, I.S.N., A.M.G.T. and M.H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the National Council for Scientific and Technological Development, CNPq (Grants: 130414/2022-0, 303670/2018-5), and by the São Paulo Research Foundation, FAPESP (Grant number: 2021/06029-7).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AABB	Axis-Aligned Bounding Boxes
BV	Bounding Volume
DEM	Digital Elevation Model
DN	Digital Number
EOP	Exterior Orientation Parameter
FoV	Field of View
GCP	Ground Control Point
HPR	Hidden Points Removal
IOP	Interior Orientation Parameter
LAI	Leaf Area Index
LiDAR	Light Detection and Ranging
MLS	Moving Least Squares
OBB	Oriented Bounding Box
PCL	Point Cloud Library
PSOSU	Pixel Size in Object Space Units
RGB	Red, Green, Blue
TLS	Terrestrial Laser Scanner
VFC	Viewing Frustum Culling

References

- Xie, P.; Du, R.; Ma, Z.; Cen, H. Generating 3D Multispectral Point Clouds of Plants with Fusion of Snapshot Spectral and RGB-D Images. *Plant Phenomics* **2023**, *5*, 40. [CrossRef]
- Ruoppa, L.; Oinonen, O.; Taher, J.; Lehtomäki, M.; Takhtkeshha, N.; Kukko, A.; Kaartinen, H.; Hyypä, J. Unsupervised Deep Learning for Semantic Segmentation of Multispectral LiDAR Forest Point Clouds. *ISPRS J. Photogramm. Remote Sens.* **2025**, *228*, 694–722. [CrossRef]
- Mirbod, O.; Choi, D.; Schueller, J.K. From Simulation to Field Validation: A Digital Twin-Driven Sim2real Transfer Approach for Strawberry Fruit Detection and Sizing. *AgriEngineering* **2025**, *7*, 81. [CrossRef]
- Rivera, G.; Porras, R.; Florencia, R.; Sánchez-Solís, J.P. LiDAR Applications in Precision Agriculture for Cultivating Crops: A Review of Recent Advances. *Comput. Electron. Agric.* **2023**, *207*, 107737. [CrossRef]
- Gharibi, H.; Habib, A. True Orthophoto Generation from Aerial Frame Images and LiDAR Data: An Update. *Remote Sens.* **2018**, *10*, 581. [CrossRef]
- Seitz, S.M.; Dyer, C.R. Photorealistic Scene Reconstruction by Voxel Coloring. *Int. J. Comput. Vis.* **1999**, *35*, 151–173. [CrossRef]
- Hudak, A.T.; Lefsky, M.A.; Cohen, W.B.; Berterretche, M. Integration of Lidar and Landsat ETM+ Data for Estimating and Mapping Forest Canopy Height. *Remote Sens. Environ.* **2002**, *82*, 397–416. [CrossRef]
- Popescu, S.; Wynne, R. Seeing the Trees in the Forest: Using Lidar and Multispectral Data Fusion with Local Filtering and Variable Window Size for Estimating Tree Height. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 589–604. [CrossRef]
- Mundt, J.T.; Streutker, D.R.; Glenn, N.F. Mapping Sagebrush Distribution Using Fusion of Hyperspectral and Lidar Classifications. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 47–54. [CrossRef]
- Geerling, G.W.; Labrador-Garcia, M.; Clevers, J.; Ragas, A.M.J.; Smits, A.J.M. Classification of Floodplain Vegetation by Data Fusion of Spectral (CASI) and LiDAR Data. *Int. J. Remote Sens.* **2007**, *28*, 4263–4284. [CrossRef]
- Anderson, J.E.; Plourde, L.C.; Martin, M.E.; Braswell, B.H.; Smith, M.-L.; Dubayah, R.O.; Hofton, M.A.; Blair, J.B. Integrating Waveform Lidar with Hyperspectral Imagery for Inventory of a Northern Temperate Forest. *Remote Sens. Environ.* **2008**, *112*, 1856–1870. [CrossRef]
- Packalén, P.; Suvanto, A.; Maltamo, M. A Two Stage Method to Estimate Species-Specific Growing Stock. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1451–1460. [CrossRef]
- Kampe, T.U.; Johnson, B.R.; Kuester, M.A.; Keller, M. NEON: The First Continental-Scale Ecological Observatory with Airborne Remote Sensing of Vegetation Canopy Biochemistry and Structure. *J. Appl. Remote Sens.* **2010**, *4*, 043510. [CrossRef]
- Agrowing Development Team. Available online: <https://agrowing.com/> (accessed on 3 September 2025).
- Tommaselli, A.M.G.; Berveglieri, A.; Imai, N.N.; Santos, G.H.; Moriya, E.A.S.; Watanabe, F.S.Y.; Neto, L.S. Geometric Performance of a Camera with Single Sensor and Multiple Heads. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 389–396. [CrossRef]
- Okhrimenko, M.; Coburn, C.; Hopkinson, C. Investigating Multispectral Lidar Radiometry: An Overview of the Experimental Framework. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8745–8748. [CrossRef]
- Crombez, N.; Caron, G.; Mouaddib, E. 3D Point Cloud Model Colorization by Dense Registration of Digital Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 123–130. [CrossRef]
- Nielsen, M.Ø. True Orthophoto Generation. Master's Thesis, Technical University of Denmark, Kgs. Lyngby, Denmark, 2004.
- Eisemann, M.; Magnor, M. *Filtered Blending and Floating Textures: Ghosting-Free Projective Texturing with Multiple Images*; TU Braunschweig: Braunschweig, Germany, 2007.
- Zhang, H.; Manocha, D.; Hudson, T.; Hoff, K.E. Visibility Culling Using Hierarchical Occlusion Maps. In Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 3–8 August 1997; ACM Press/Addison-Wesley Publishing Co.: New York, NY, USA, 1997; pp. 77–88.
- Assarsson, U.; Moller, T. Optimized View Frustum Culling Algorithms for Bounding Boxes. *J. Graph. Tools* **2000**, *5*, 9–22. [CrossRef]
- Blinn, J.F. Backface Culling Snags (Rendering Algorithm). *IEEE Comput. Graph. Appl.* **1993**, *13*, 94–97. [CrossRef]
- Cohen-Or, D.; Chrysanthou, Y.L.; Silva, C.T.; Durand, F. A Survey of Visibility for Walkthrough Applications. *IEEE Trans. Vis. Comput. Graph.* **2003**, *9*, 412–431. [CrossRef]
- Catmull, E.E. A Subdivision Algorithm for Computer Display of Curved Surfaces. Ph.D. Thesis, The University of Utah, Salt Lake City, UT, USA, 1794.
- Fuchs, H.; Kedem, Z.M.; Naylor, B.F. On Visible Surface Generation by a Priori Tree Structures. In Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques, Seattle, WA, USA, 14–18 July 1980; pp. 124–133.
- Appel, A. Some Techniques for Shading Machine Renderings of Solids. In Proceedings of the AFIPS '68 (Spring): Spring Joint Computer Conference, Atlantic City, NJ, USA, 30 April–2 May 1968; pp. 37–45.
- Whitted, T. An Improved Illumination Model for Shaded Display. In Proceedings of the 6th Annual Conference on Computer Graphics and Interactive Techniques, Chicago, IL, USA, 8–10 August 1979; p. 14. [CrossRef]

28. Katz, S.; Tal, A.; Basri, R. Direct Visibility of Point Sets. In Proceedings of the SIGGRAPH '07: ACM SIGGRAPH 2007 Papers, San Diego, CA, USA, 5–9 August 2007; Volume 26, p. 11. [\[CrossRef\]](#)
29. Rusu, R.B.; Cousins, S. 3D Is Here: Point Cloud Library (PCL). In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4. [\[CrossRef\]](#)
30. Hu, F.; Lin, C.; Peng, J.; Wang, J.; Zhai, R. Rapeseed Leaf Estimation Methods at Field Scale by Using Terrestrial LiDAR Point Cloud. *Agronomy* **2022**, *12*, 2409. [\[CrossRef\]](#)
31. Alexa, M.; Behr, J.; Cohen-Or, D.; Fleishman, S.; Levin, D.; Silva, C.T. Computing and Rendering Point Set Surfaces. *IEEE Trans. Vis. Comput. Graph.* **2003**, *9*, 3–15. [\[CrossRef\]](#)
32. Shen, C.; O'Brien, J.F.; Shewchuk, J.R. Interpolating and Approximating Implicit Surfaces from Polygon Soup. In *Computer Graphics Proceedings, Annual Conference Series*; Association for Computing Machinery: New York, NY, USA, 2004; pp. 896–904.
33. Fleishman, S.; Cohen-Or, D.; Silva, C.T. Robust Moving Least-Squares Fitting with Sharp Features. *ACM Trans. Graph.* **2005**, *24*, 544–552. [\[CrossRef\]](#)
34. Jenke, P.; Wand, M.; Bokeloh, M.; Schilling, A.; Straßer, W. Bayesian Point Cloud Reconstruction. *Comput. Graph. Forum* **2006**, *25*, 379–388. [\[CrossRef\]](#)
35. Bishop, L.; Eberly, D.; Whitted, T.; Finch, M.; Shantz, M. Designing a PC Game Engine. *IEEE Comput. Graph. Appl.* **1998**, *18*, 46–53. [\[CrossRef\]](#)
36. DirectModel. *DirectModel: 1.0 Specification*, Hewlett Packard Company: Corvallis, OR, USA, 1990.
37. Krishnan, A. Point Cloud Library (PCL). Available online: https://pointclouds.org/documentation/frustum_culling_8h_source.html (accessed on 20 January 2024).
38. Argueta, C. Filtering a Point Cloud to Match the Field of View of the Camera. Available online: <https://medium.com/@kidargueta/filtering-a-point-cloud-to-match-the-field-of-view-of-the-camera-ccab0d189b58> (accessed on 20 January 2024).
39. Katz, S.; Tal, A. On the Visibility of Point Clouds. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1350–1358.
40. MathWorks. Computing the Convex Hull—MATLAB & Simulink. Available online: <https://www.mathworks.com/help/matlab/math/computing-the-convex-hull.html> (accessed on 21 January 2024).
41. Norberto, I.S.; Faria Junior, C.D.S.; Tommaselli, A.M.G.; Galo, M.; Silva, R.M. MLP-Based Classification of Multispectral Point Clouds for Digital Agriculture. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Dubai, United Arab Emirates, 6–11 April 2025.
42. Reji, J.; Nidamanuri, R.R. Deep Learning Based Fusion of LiDAR Point Cloud and Multispectral Imagery for Crop Classification Sensitive to Nitrogen Level. In Proceedings of the 2023 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS), Hyderabad, India, 27–29 January 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–4.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.