Markerless Aerial-Terrestrial Co-Registration of Forest Point Clouds using a Deformable Pose Graph

Benoît Casseau[1] Nived Chebrolu[1] Matias Mattamala[1] Leonard Freissmuth[1,2] Maurice Fallon[1]

***Abstract*— For biodiversity and forestry applications, end-users desire maps of forests that are fully detailed—from the forest floor to the canopy. Terrestrial laser scanning and aerial laser scanning are accurate and increasingly mature methods for scanning the forest. However, individually they are not able to estimate attributes such as tree height, trunk diameter and canopy density due to the inherent differences in their field-of-view and mapping processes. In this work, we present a pipeline that can automatically generate a single joint terrestrial and aerial forest reconstruction. The novelty of the approach is a marker-free registration pipeline, which estimates a set of relative transformation constraints between the aerial cloud and terrestrial sub-clouds without requiring any co-registration reflective markers to be physically placed in the scene. Our method then uses these constraints in a pose graph formulation, which enables us to finely align the respective clouds while respecting spatial constraints introduced by the terrestrial SLAM scanning process. We demonstrate that our approach can produce a fine-grained and complete reconstruction of large-scale natural environments, enabling multi-platform data capture for forestry applications without requiring external infrastructure.**

I. Introduction

Terrestrial laser scanning (TLS, tripod-based) and mobile laser scanning (MLS, human or robot carried) play an emerging role in biodiversity monitoring and forestry by capturing the detailed geometric data required for the study of individual trees. These data can be used to recover highly detailed reconstructions of the under-canopy, which can be used to extract traits such as the diameter-at-breast-height (DBH), tree species and general forest density [29],[21]. Such traits can be used for decision-making in forestry as well as robotic forest harvesting [18]. However, TLS/MLS data alone are not able to capture other traits such as the tree height or crown volume, which cannot be accurately observable from the ground due to limitations in the sensor range as well as blocking foliage and other occlusions [28].

Aerial vehicles, such as MAVs and fixed-wing aircraft, fly above the canopy to carry out aerial laser scanning (ALS). These data captures structural details that cannot be obtained from terrestrial scanning. ALS survey can also reconstruct much larger areas in a single flight (tens of hectares). Nevertheless, they intrinsically produce lower density reconstructions of the understory than TLS or MLS, making it difficult to determine stem properties in dense forests [28]. Given the complementary nature of the sensing modalities, combining the terrestrial and aerial reconstructions would allow comprehensive characterization of tree and forest structures in dense and diverse ecosystems. A further advantage of the combined map is that it serves as an effective reference map, improving robot localization capabilities.

Fig. 1: A forest map built using a terrestrial map from our mobile scanning system (orange) and an aerial map from a drone (blue). Together, the scans produce a map reconstructing the canopy and understory in sufficient detail for individual-tree forest inventory.

The standard way of achieving joint ALS-MLS reconstruction is to manually place reflective markers in the forest which can be detected by both ground and aerial systems. This facilitates the alignment of the clouds, either through manual intervention or automated processes [14]. This method involves a time-consuming setup of markers, and ensuring marker visibility for both ALS and MLS systems is challenging in dense forests. Hence, there is an interest in developing marker-free registration methods, with recent advancements showing promise [26], [5]. These methods perform a rigid alignment of the aerial and terrestrial clouds, assuming that both are independently accurate. However, while uniform accuracy can be guaranteed for ALS thanks to the fusion of inertial measurement units (IMU) and global navigation satellite system (GNSS); a mobile scanner cannot rely on GNSS under canopy due to occlusion and signal loss. Therefore, MLS typically estimates their pose via a SLAM system using onboard LiDAR and IMU measurements with a crude $5\,\mathrm{m}$ to $10\,\mathrm{m}$ accurate GNSS position estimate available. While MLS (from onboard LiDAR

[1]The authors are with the University of Oxford, UK. {benoit, nived, matias, mfallon}@robots.ox.ac.uk .

[2]The author is also with the Technical University of Munich, Germany. l.freissmuth@tum.de

odometry) can produce accurate local reconstruction, it can still suffer from drift in long sequences (Fig. 2).

In this paper, we present an approach to combine aerial and terrestrial data to achieve high-fidelity georeferenced point cloud reconstructions by automatically finding matches between individual, local terrestrial sub-maps and a single global, georeferenced ALS map without requiring specialized markers. The approach combines global aerial maps with the locally accurate MLS sub-maps, by jointly optimizing them in a factor graph-based formulation that combines local and global constraints. We demonstrate that our method enables denser, complete reconstructions of the full trees, which enable further forestry analyses, as shown in Fig. 1.

The specific contributions of our work are:

- A method to compute local aerial-terrestrial registration that does not require custom targets for alignment, only requiring a rough GNSS prior position estimate.
- A pose graph-based formulation that performs fine-grained alignment of terrestrial clouds to a global aerial reference, applicable to both sequences captured with MLS as well as a set of individual TLS scans.
- Experimental validation using aerial and terrestrial datasets obtained in a Swiss mixed forest and a Finnish conifer forest.

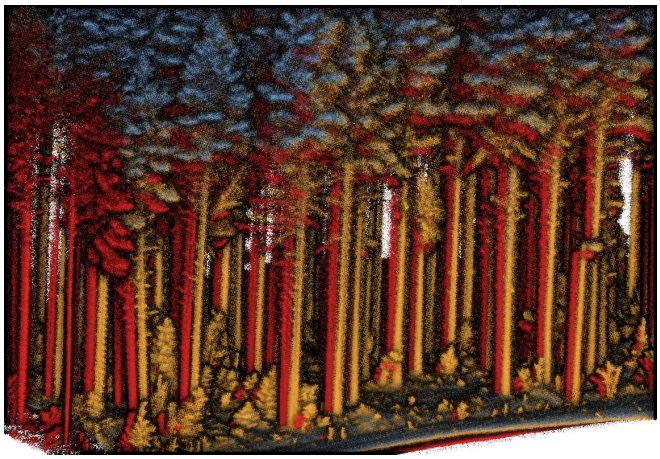


Fig. 2: An unoptimized MLS point cloud (in red) is poorly aligned with the ALS map (blue). On the other hand, the optimized cloud (in orange) is directly aligned with the ALS map. The shearing in the red MLS point cloud is due to the small drift in the SLAM system (a few meters across a kilometer scale map).

II. Related Work

The problem of aerial and terrestrial cloud registration has been motivated by the different scanning perspectives provided by each technique [6]. Early approaches achieved this by manual alignment of the different different sources, which was generally assisted by some automatic components. Lovell et al. [17] presented one of the earliest examples of measuring canopy in an Australian forest by combining both data sources. Omasa et al. [22] also combined aerial and terrestrial scans to model an urban park and trees using rendered bird-eye views. More recent approaches have used automatic tree detection instead but still required manual registration [19].

To achieve semi-automatic matching, specialized markers (e.g. reflective targets) have been used for the co-registration of independent clouds [14], [13]. Chasmer et al. [6] used targets to study the distribution of scans captured by ground and aerial sensing. Similarly, Holopainen et al. [15] used reflective spheres to align clouds obtained in urban forests. Nevertheless, these solutions require careful positioning of the markers to maximize the co-visibility in aerial and terrestrial scans.

Instead, we are interested in automatic solutions, which do not require any specialized markers to perform the co-registration. This is generally achieved by having an aerial-terrestrial matching step, and then a registration procedure that provides a relative transformation between aerial and terrestrial scans. For matching, previous approaches have explored strategies such as GNSS priors [20], as well as identifying matching features between the scans, including tree density analysis [8], tree crown analysis [23] or stem position matching [4]. In this work we leverage some of the previous techniques, particularly detecting tree stems, to deal with partial scans due to the incremental nature of MLS scanning.

Different registration approaches have been explored depending on the application domain. For forestry applications where the objective is to provide denser reconstructions, previous works have explored techniques such as graph matching to provide a registration solution [24], which can be further refined with standard point cloud alignment algorithms such as iterative closest point (ICP) [3]. Alternatively, non-linear optimization refinement has been proposed to achieve fine-grained alignment [8], [5]. Other solutions based on particle filtering have been proposed for related problems such as robot localization using an aerial map [9], which are beyond the scope of this work.

Our system builds upon the previous ideas to combine some of the co-registration techniques with an optimization-based refinement step, similarly as done by Castorena et al. [5]. However, we explicitly aim to introduce spatial consistency in the refinement using a factor graph formulation, which is *deformed* to better align the aerial scan.

III. Method

In this section, we present our aerial-terrestrial co-registration pipeline, shown in Fig. 3. The inputs are an aerial point cloud in UTM frame (*ALS cloud*), and terrestrial scans from MLS (*MLS clouds*). The MLS clouds can be in either a *pose graph format* or a *tile format*.

In *pose graph format* the MLS clouds are represented by a SLAM pose graph in UTM frame with local terrestrial point clouds attached to each pose. The clouds—which we name *data payloads*—are obtained by temporally aggregating LiDAR scans at sensor rate (10 Hz) within a fixed distance; an example is shown in Fig. 4 (a). The data payload is converted from a local coordinate frame to the UTM frame using the GNSS measurements acquired by the on-board

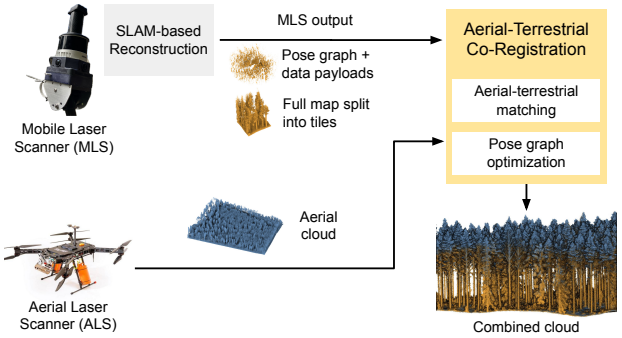
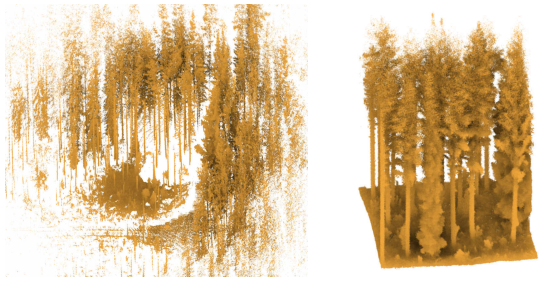


Fig. 3: General system overview. Our aerial-terrestrial co-registration system combines the aerial clouds from ALS, with a terrestrial reconstruction obtained by means of a LiDAR-Inertial odometry and pose-graph SLAM system.



(a) Data payload (b) Tile cloud

Fig. 4: We consider two formats for the MLS data. The *pose graph format* is defined by a pose graph and a collection of data payloads (a), obtained by temporal aggregation of consecutive LiDAR scans; payloads cover a larger area but are sparse and lack canopy points. The *tile format* represents the MLS mission as a collection of tiles (b), obtained by aggregating all the mission scans into a single, dense cloud and partitioning them into a fixed-size grid.

receiver. This pose graph format enables fine-grained aerial-terrestrial alignment by exploiting the temporal information about how the data was acquired.

In *tile format* we consider the full, single cloud obtained by joining all the data payloads acquired over a mission and then partitioning the data into a fixed-size, gravity-aligned 2D grid to produce a collection of *tiles*. We use a tile of size $20\,\mathrm{m} \times 20\,\mathrm{m}$ in this work, as shown in Fig. 4 (b). This representation is independent of the MLS device, and does not require access to the SLAM information—only the relative transformation between the tiles specified by the grid.

A. Aerial-terrestrial Matching

The first step aims to find a relative transformation between each MLS cloud and the aerial cloud (Fig. 5). The procedure is independent of the MLS cloud format.

We denote $\mathbf{T}_{\mathtt{AM}}$ the transformation that maps an individual MLS cloud (payload/tile) in $\mathtt{M}$ frame into the ALS frame $\mathtt{A}$. The set of relative transformations obtained for each of the MLS clouds will provide global constraints for the pose graph optimization procedures presented later in Sec. III-B.

1) Pre-processing: Since the ALS cloud can be very large ($\sim$600 m$\times$600 m in our datasets), we reduce the search space by cropping the ALS cloud around a neighborhood of the

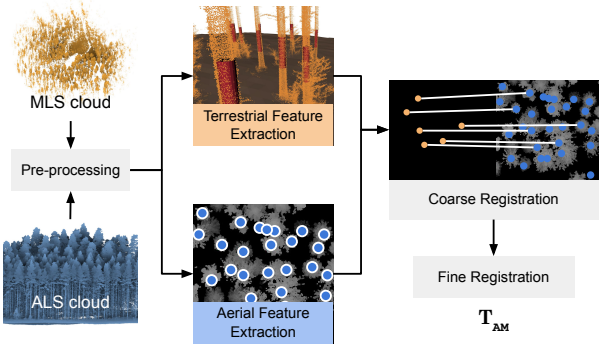


Fig. 5: The aerial-terrestrial registration step aims to find the relative transformation between a large aerial and a much smaller local terrestrial cloud, $\mathbf{T}_{\mathtt{AM}}$.

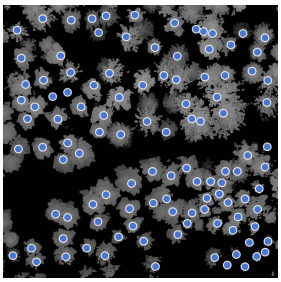
center of the MLS cloud. This is achieved by using the approximate GNSS position of the MLS cloud (accurate to about $5\,\mathrm{m}$ to $10\,\mathrm{m}$) and a padding factor depending on the MLS cloud's size. On average, this produced crops of $30\,\mathrm{m}\times30\,\mathrm{m}$ for both payloads and tiles in our experiments.

While the MLS clouds are gravity-aligned and approximately aligned with North (as with the ALS cloud), we observe a height offset compared to the ALS cloud. This may be attributed to an imprecise GNSS fix obtained by the MLS under the forest canopy. We correct this misalignment of the clouds along the height axis by matching the ground planes from both the MLS and ALS clouds. To rectify this misalignment along the height axis, we estimate the correction by matching the ground planes derived from both the MLS and ALS clouds. This process involves identifying ground points based on their normal information and employing a RANSAC procedure to fit a plane to each MLS cloud as well as the ALS cloud. Using the fitted planes from ALS and MLS clouds, we determine a rotation correction by aligning the normal vectors to the plane as well as the vertical offset between the ALS and terrestrial MLS clouds, similar to the method by Shao et al. [26].

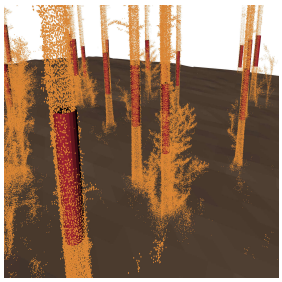
2) Aerial Feature Extraction: The next step of the method is to extract features from the cropped ALS cloud — specifically the peaks of each tree present. We rasterize a top-down view of the cloud by selecting the maximum height of the peak points along the vertical direction, as also done by De Lima et al. [9]. This effectively represents a *canopy height map* (CHM) of the aerial cropped cloud (see Fig. 6 (a))

We apply non-maxima suppression on the CHM to extract local peaks, which we assume corresponds to the tree stem locations, as also considered previously by other authors [16]. The output tree location corresponds to the features of the aerial map that are used for matching, as shown in Fig. 6. A limitation of this approach is that for dense forest environments it is likely to be more difficult to distinguish individual treetops due to canopy overlap.

3) Terrestrial Feature Extraction: The MLS clouds are acquired by an MLS with a horizontal LiDAR and narrow field-of-view ($\sim$30°), inclined at a pitch angle of 20 °. Hence,



(a) Peak extraction in the rasterized ALS cloud.



(b) Tree segmentation in the MLS clouds.

Fig. 6: Feature extraction approaches to determine the tree positions in (a) ALS and (b) MLS clouds.

the tree canopy remains obstructed for trees around the sensing device, making it unfeasible to generate CHM which is complete enough to extract tree locations. To address this limitation, we use an alternative method for terrestrial tree stem estimation, similar to Proudman et al. [25] and Freißmuth et al. [12]. We extract the points up to $5\,\mathrm{m}$ above the ground plane and apply a density-based spatial clustering algorithm (DBSCAN [10]). We then fit a cylinder within a RANSAC [11] loop for each cluster, thereby obtaining the principal axis and the center position of each tree. These tree center positions serve as the features of the MLS cloud, as shown in Fig. 6 (b).

4) Coarse Registration using Maximum Clique: After extracting the positions of tree trunks from both the terrestrial and aerial clouds, we match individual trees to estimate the relative 2D transformation between both clouds.

To find correspondences, we use a maximum clique algorithm inspired by Bailey et al. [1]. We first build an aerial graph $\mathcal{G}_\mathrm{a}$ by using the tree locations from the CHM as the vertices, resulting in a *completely connected* graph. Similarly, we create a complete terrestrial graph $\mathcal{G}_\mathrm{t}$ using the estimated tree positions estimated in Sec. III-A.3.

Both graphs are used to construct a *correspondence graph* $\mathcal{C}$ between the aerial and local terrestrial graph [7]. Technically, this correspondence graph is formed by taking the Cartesian product of the vertices from $\mathcal{G}_\mathrm{a}$ and $\mathcal{G}_\mathrm{t}$. For example, given $u, u' \in \mathcal{G}_\mathrm{a}$ and $v, v' \in \mathcal{G}_\mathrm{t}$, $\mathcal{C}$ has as vertices (u, v), (u, v'), (u', v) and (u', v'). The edges of $\mathcal{C}$ encode pairwise consistency: an edge between vertices (u, v) and (u', v') exists if and only if $|d(u, u') - d(v, v')| < \tau$, where $d(\cdot)$ is a distance metric between vertices of the graphs, and τ is a tolerance threshold.

The graph correspondence problem between $\mathcal{G}_\mathrm{a}$ and $\mathcal{G}_\mathrm{t}$ then reduces to finding the maximum clique of the correspondence graph $\mathcal{C}$ [2]. This allows us to obtain matches between the 2D tree positions of the aerial and terrestrial clouds, used to solve for the relative planar pose using Umeyama's method [27]. As previously mentioned, by cropping the input clouds, we manage to limit the number of trees, thereby controlling the size of the correspondence graph $\mathcal{C}$—ensuring that it remains manageable for efficient maximum clique detection. This operation typically takes only a few seconds using the Python library *NetworkX*.

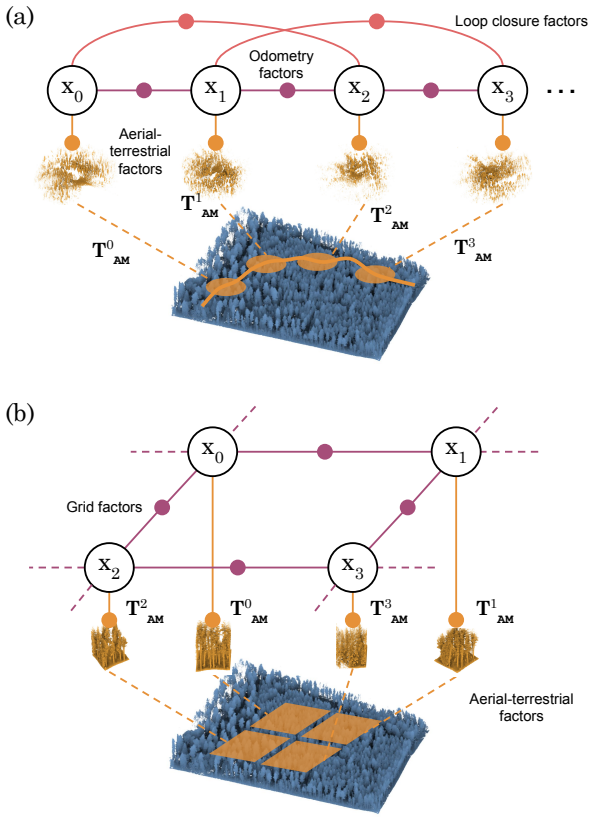


Fig. 7: The factor graph formulation for aerial-terrestrial registration for our two data formats. Top: For pose graph/payload optimization we have standard odometry (purple) and loop closure factors (red) to which we add individual aerial-terrestrial prior factors (in orange). Bottom: For the tile format, the grid representation forms local constraints (purple) and the individual tiles are registered to add aerial-terrestrial prior factors (in orange).

5) Fine Registration using ICP: The planar pose estimate from the maximum clique method, combined with the z-offset and pitch and roll corrections from the pre-processing step provide a coarse estimate of the relative 6-DoF transformation between the local MLS cloud and the aerial/global frame, denoted by $\mathbf{T}_{\mathrm{AM}}$.

In practice, we observed offsets of about $20\,\mathrm{cm}$ in all directions for the registered clouds at this step—particularly noticeable on the tree trunks. Therefore, to achieve a finer registration, we conduct a final Iterative Closest Point (ICP) refinement [3]. Finally, we filter out those matches where the number of ICP inliers falls below a specified threshold, as we do not require all tiles/payloads to be precisely registered to the aerial cloud.

B. Aerial-terrestrial Pose Graph Optimization

To achieve global consistency between the ALS and MLS clouds, previous approaches have proposed to solve an optimization problem that minimizes the overall registration cost between aerial and terrestrial scans, using additional constraints provided by cycle consistency costs [5]. Here we explore a unified method, based on factor graph optimization,

which enables us to co-register aerial and terrestrial clouds for both the SLAM and tile formats.

Using the approach described in Sec. III-A, we obtain a set of 6-DoF transformations $\mathbf{T}_{\text{AM}}^{l}$ between the MLS clouds and the ALS, where l is the index of all the MLS clouds successfully registered after the ICP registration step. Each transformation provides a constraint $\mathbf{r}_{\text{AM}}^{l}$ in the factor graph, which we refer to as *Aerial-terrestrial factors*. We discard the index l for simplicity.

We formulate an optimization that leverages the structure of the MLS data format, as well as the aerial-terrestrial matches:

$$\{\mathbf{T}_i\} = \text{argmax} \underbrace{\sum \|\mathbf{r}_{\text{structure}}\|^2_{\mathbf{\Sigma}}}_{\text{MLS structure factors}} + \underbrace{\sum \|\mathbf{r}_{\text{AM}}\|^2_{\mathbf{\Sigma}}}_{\text{Aerial-terrestrial factors}} , \quad (1)$$

where $\{\mathbf{T}_i\}$ is the optimized set of poses that define the factor graph. The technical details about the residuals are as follows.

1) Structure factors (pose graph case): First, we consider the typical SLAM case in which the MLS mission is described by a pose graph and a set of payload clouds. The structure factors are given by

odometry factors between the poses:

$$\mathbf{r}_{\text{odometry}} = \text{Log}\left(\Delta\mathbf{T}^{-1}\mathbf{T}_i^{-1}\mathbf{T}_{i+1}\right), \quad (2)$$

where $\mathbf{T}_i^{-1}$,$\mathbf{T}_{i+1}$ are two consecutive poses and $\Delta\mathbf{T}$ is the relative odometry change. Similarly, we also consider *loop closure factors* between poses i and j proposed by a place recognition system:

$$\mathbf{r}_{\text{loop}} = \text{Log}\left(\Delta\mathbf{T}_{i,j}^{-1}\mathbf{T}_i^{-1}\mathbf{T}_j\right). \quad (3)$$

Fig. 7(a) illustrates these factors.

2) Structure factors (tiles case): When we use the MLS clouds in tiles format, the structure is given by binary factors that connect a node i to its neighbors $n \in \mathcal{N}$:

$$\mathbf{r}_{\text{grid}} = \text{Log}\left(\Delta\mathbf{T}_{i,n}^{-1}\mathbf{T}_i^{-1}\mathbf{T}_n\right), \quad (4)$$

Similarly, Fig. 7(b) shows how these factors are defined graphically.

3) Aerial-Terrestrial factor: Lastly, the aerial-terrestrial constraints impose the 6-DoF relative transformations between the MLS clouds (payloads or tiles) with the ALS cloud:

$$\mathbf{r}_{\text{AM}} = \text{Log}\left(\mathbf{T}_{\text{AM}}^{i}\mathbf{T}_i\right), \quad (5)$$

This is illustrated for both cases in Fig. 7.

IV. Experiments

A. Datasets

We evaluated our approach using data collected from two field campaigns conducted in Finland and Switzerland. The first campaign took place in Evo, Finland, in a Southern boreal forest characterized by conifers consisting mostly of pines and spruces. ALS point cloud data was collected using an Avartek Bower drone covering up to $730\,\text{m} \times 540\,\text{m}$. MLS data was captured by a backpack-carried Hesai XT32 LiDAR, covering areas of more than a hectare.

The second campaign was in Stein am Rhein (SaR), Switzerland, in a mixed forest. The aerial data was collected by a DJI M600 drone carrying a Velodyne HDL-32E LiDAR and an RTK GNSS receiver in a single mission covering $1850\,\text{m} \times 500\,\text{m}$. For the MLS we used the same setup as in Evo, and we collected data in four different missions, covering about $1\,\text{ha}$ each.

B. Deformable Registration from SLAM Missions

In the first experiment, we demonstrate and evaluate our co-registration approach across different datasets. Fig. 8 shows a large-scale example of the co-registrations produced by our method across multiple missions in Evo, Finland.

We register four MLS missions to an aerial map covering an overall area of about $4\,\text{ha}$. Each MLS mission trajectory is represented by distinct colored lines overlaid on the ALS point cloud (in blue). The cross-sectional views at different points on the map highlight the precise alignment accomplished through our approach. The consistency in trunk reconstructions along the length of trees suggests a high level of accuracy in co-registration.

To quantitatively evaluate the advantages of the proposed pose graph refinement, we combined all the payloads in the MLS mission before and after our registration process. Then, we cropped the ALS cloud or the MLS cloud to obtain the maximum overlapping subset of points, and computed the *point-to-point* error obtained before and after the pose graph refinement.

Tab. I presents the results obtained for all missions. We report three different metrics: (1) point-to-point error against the ALS cloud before and after optimization, (2) average point-wise offset after to the optimization, and (3) average pose graph correction.

Regarding (1), in general, we observed a decrease in the error against the ALS cloud after the pose graph refinement — as expected. We then inspected how much the clouds and pose graph nodes move after optimization (2) and (3). We report a displacement of the clouds of the order of $10\,\text{cm}$ on average across all missions, while the nodes effectively moved $25\,\text{cm}$ on average. The discrepancies between these numbers can be explained by the metric used to evaluate the cloud offset (point-to-point distance). Considering the nearest neighbors to determine the error does not necessarily reflect how much a point moved after the optimization correction but rather suggests a measure of the overall alignment. However, both the metrics reflect a substantial improvement due to the optimization step, as qualitatively shown in Fig. 8.

C. Deformable Registration for Tile-based Data

In this experiment, we demonstrate pose graph-based optimization for tile-based MLS data. We partitioned one of the MLS missions from the Evo dataset into square tiles measuring $20\,\text{m} \times 20\,\text{m}$ each. We then executed our pipeline, including the tile-wise aerial-terrestrial matching step (see Sec. III-A), to assess the performance of registration. The

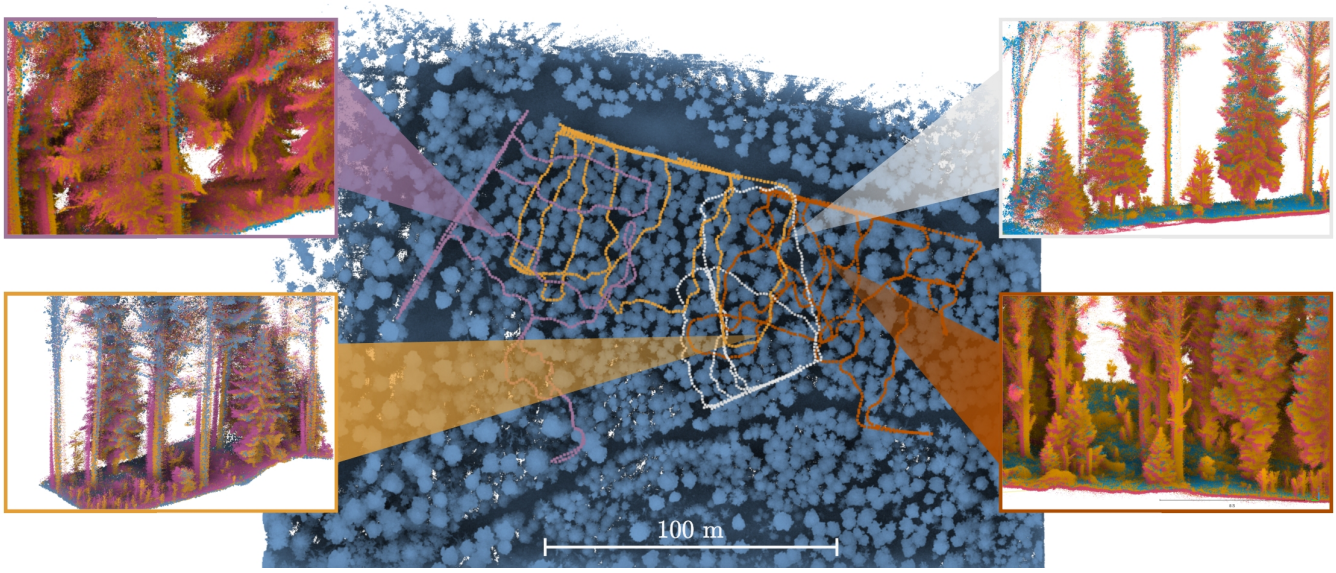


Fig. 8: Illustrative examples ALS and MLS point cloud co-registrations using our proposed pipeline for four SLAM missions in Evo, Finland. MLS clouds are shown before optimization (pink), after pose graph refinement (orange), and the aerial cloud (blue). Discrepancies observed before and after the alignment highlight the corrections introduced through the addition of aerial-terrestrial constraints to the pose graph structure.

Place	Mission	ALS-MLS error [m]		Cloud shift [m]	Pose graph shift [m]
		Pre	Post		
Evo	1	0.64	**0.61**	0.05	0.21
	2	0.28	**0.26**	0.06	0.23
	3	0.39	**0.31**	0.09	0.45
	4	**0.36**	0.37	0.07	0.28
SaR	1	0.78	**0.39**	0.38	0.20
	2	0.47	**0.41**	0.13	0.38

TABLE I: Co-registration metrics for SLAM-based MLS missions. We evaluated the point-to-point error before and after the pose graph optimization w.r.t the ALS cloud, as well the point cloud and pose graph correction magnitude.

Area	Mission	Manual		Ours	
		mean	std	mean	std
Evo	1	0.67	2.11	**0.41**	2.20
	2	0.31	0.31	**0.21**	0.30
	3	0.38	1.27	**0.35**	1.53
	4	0.40	0.86	**0.29**	0.94
SaR	1	0.31	0.48	**0.24**	0.37
	2	0.32	0.40	**0.26**	0.38

TABLE II: Accuracy of aerial-terrestrial registration for individual payloads. We evaluated the point-to-point error for our registration method (Sec. III-A) and manual alignment as a reference.

main distinction between the MLS payload data and the tiled data lies in the density of the respective point clouds. However, since the coarse registration step only uses sparse tree location positions, it does not impact computation time significantly. We present the results of the initial aerial-terrestrial registration in Fig. 9(a). Here, we colorized the tiles individually based on the ICP fitness score obtained after the matching step, where blue indicates a higher fitness score. Other colored tiles represent lower fitness scores. This is often due to an insufficient number of trees present in the corresponding tiles for unambiguous matching within the aerial cloud.

We address this limitation by leveraging the pose graph structure to propagate corrections to tiles that initially failed to register successfully with the aerial clouds. After optimization, guided by the prior imposed by the tile grid, alignment is attained across all tiles, as depicted in Fig. 9(b).

D. Accuracy Analysis of Aerial-Terrestrial Transformation

To assess the quality of the proposed aerial-terrestrial factors, we evaluated the accuracy of the individual payload-based matches. We compared the results obtained by our matching module against a subset of manually registered clouds. We chose a subset of MLS payload clouds and manually registered them using CloudCompare, selecting four corresponding points in the payload and the aerial map, followed by ICP refinement. Tab. II presents the root-mean-square error (RMSE) computed from point-to-point distances for both the manually aligned clouds and our proposed method (Sec. III-A). This evaluation was conducted on a subset of mission payloads for which successful alignment was achieved using our approach. On average, we observed a point-to-point RMSE of $0.44\,\mathrm{m}$ for the Evo dataset and $0.31\,\mathrm{m}$ for the SaR dataset. The results demonstrate that our registration method achieves accurate local alignment, with our approach showing better point-to-point error compared to the manual alignment.

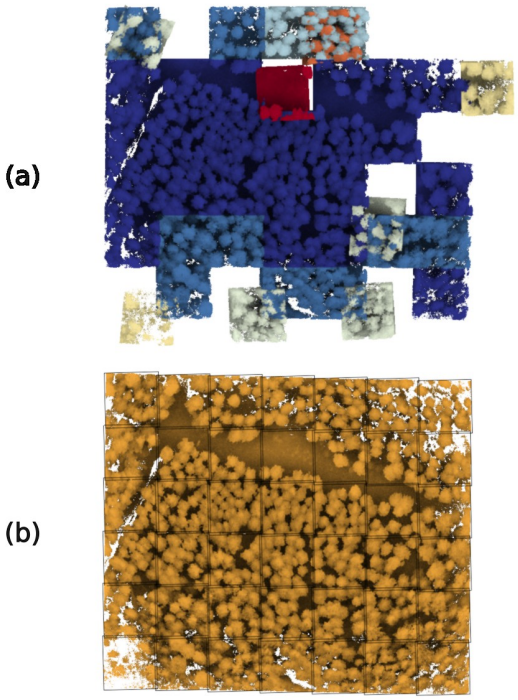


Fig. 9: Deformable registration for tile-based data. Each tile measure 20 m × 20 m. (a) Registration results after the initial aerial-terrestrial tile-wise matching step. Blue denotes higher ICP matching scores, while other colors indicate lower scores. (b) Results after pose graph optimization, achieving alignment across all tiles despite initial registration failures for some of the tiles.

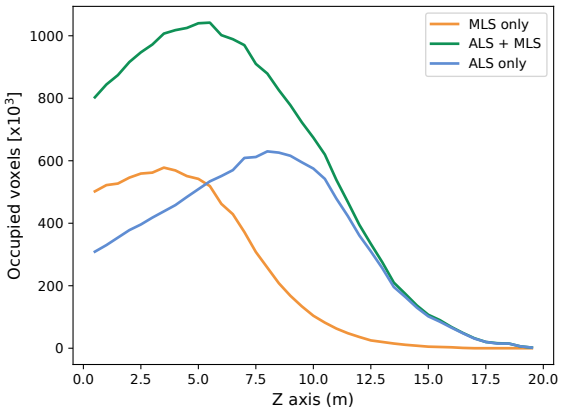


Fig. 10: Completeness analysis. Voxel occupancy (with 5 cm resolution) at different heights for the aerial, terrestrial, and co-registered point clouds. The co-registered point cloud shows a higher occupancy through the height of the tree.

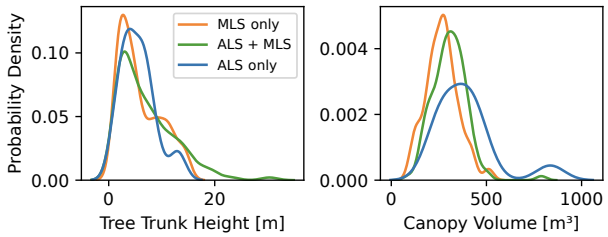


Fig. 11: Distributions of tree height and canopy volume. The combined point cloud (green) shows a shift to the right, indicating greater height and canopy volume captured compared to MLS (orange) or ALS (blue) only.

E. Accessing Completeness of Co-Registered Point Clouds

Lastly, we evaluated the benefits of co-registering terrestrial and aerial clouds in terms of *completeness*. We quantified cloud completeness through density measurements across various height intervals, utilizing a voxelized representation to ensure occupancy evaluation rather than relying solely on raw point density, which may vary depending on the LiDAR type. We achieved this by voxelizing the point cloud with a resolution of 5 cm for the ALS, MLS, and combined clouds. Fig. 10 shows the results obtained for Mission 1 of the Evo dataset. First, we observed that the distribution of occupied voxels for ALS and MLS confirms the differences expected for both modalities, with MLS concentrating the low-height points, and ALS the canopy and top of the trees. The combined ALS + MLS cloud obtained with our proposed method indeed provided more complete reconstructions, as qualitatively illustrated in the previous experiments.

To further demonstrate the advantages of combining aerial and terrestrial clouds, we ran a tree trait extraction pipeline to segment individual trees and determine the tree height and canopy volume from ALS, MLS, and co-registered clouds. For estimating the tree height, we used an algorithm that modeled the stem as a set of local cylinders along the trunk—the set of estimated cylinders was used to estimate the tree height. For the canopy volume, we used a convex hull-based algorithm that enclosed points associated to the tree that were not used to estimate the stem model, effectively capturing points associated to branches and leaves. We used the volume of the hull as a proxy for the canopy volume.

Fig. 11 shows the results obtained for this experiment. Regarding the tree height estimates, Fig. 11 (a) shows the normalized distribution of estimated heights. We observed that the combined cloud distribution effectively covers a larger range of tree heights, observed by the shift of the right-hand tail. This can be explained for the higher degree of completeness of the combined clouds compared to independent ALS and MLS clouds, which improved the robustness of the stem modeling approach for different tree heights. We reported a similar result for the canopy volume estimates, with the combined cloud displaying longer tails in the volume distribution. This demonstrates that by estimating tree traits from a combined aerial-terrestrial point cloud we can effectively produce better tree estimates for forest inventory and monitoring.

V. Conclusion

This work presented a novel approach for co-registering aerial and terrestrial point clouds. By leveraging a deformable pose graph formulation, we achieved precise align-

ment by exploiting the underlying MLS SLAM representation in a manner which the existing methods do not now. We validated our approach on diverse forest datasets in multiple countries and demonstrate the effectiveness in creating more accurate and complete forest inventories. This capability holds promise for enhancing forest management and conservation efforts.

ACKNOWLEDGMENTS

This work was supported by the European Union's Horizon Europe under grant agreement 101070405, a Royal Society University Research Fellowship and a UKRI/EPSRC Programme Grant [EP/V000748/1]. The authors acknowledge the support of PreFor Oy for providing the ALS data used in the experiments. We also thank Janine Schweier, Holger Griess and their team at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) for supporting the data collection in Stein am Rhein.

REFERENCES

[1] T. Bailey, E.M. Nebot, J. Rosenblatt, and H.F. Durrant-Whyte. Data association for mobile robot navigation: A graph theoretic approach. In *IEEE Int. Conf. Robot. Autom. (ICRA)*, pages 2512–2517. IEEE, 2000.

[2] H.G. Barrow and R.M. Burstall. Subgraph isomorphism, matching relational structures and maximal cliques. *Inf. Process. Lett.*, 4(4):83–84, 1976.

[3] P.J. Besl and N.D. McKay. A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(2):239–256, 1992.

[4] A. Bienert and H.G. Maas. Methods for the automatic geometric registration of terrestrial laser scanner point clouds in forest stands. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 93–98, 2009.

[5] J. Castorena, L.T. Dickman, A.J. Killebrew, J.R. Gattiker, R.R. Linn, and E.L. Loudermilk. Automated structural-level alignment of multi-view TLS and ALS point clouds in forestry. *ArXiv*, abs/2302.12989, 2023.

[6] L. Chasmer, C. Hopkinson, and P. Treitz. Assessing the three-dimensional frequency distribution of airborne and ground-based lidar data for red pine and mixed deciduous forest plots. *ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(8):W2, 2004.

[7] C. Chen and D.Y. Yun. Unifying graph-matching problem with a practical solution. In *Proceedings of International Conference on Systems, Signals, Control, Computers*, volume 55. Citeseer, 1998.

[8] W. Dai, B. Yang, X. Liang, Z. Dong, R. Huang, Y. Wang, and W. Li. Automated fusion of forest airborne and terrestrial point clouds through canopy density analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 156:94–107, 2019.

[9] L.C. de Lima, M. Ramezani, P.V.K. Borges, and M. Brünig. Air-ground collaborative localisation in forests using lidar canopy maps. *IEEE Robot. Autom. Lett. (RA-L)*, 8(3):1818–1825, 2023.

[10] M. Ester, H. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis, J. Han, and U.M. Fayyad, editors, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*, pages 226–231. AAAI Press, 1996.

[11] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981.

[12] L. Freißmuth, M. Mattamala, N. Chebrolu, S. Schaefer, S. Leutenegger, and M. Fallon. Online tree reconstruction and forest inventory on a mobile robotic system. *IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2024.

[13] X. Ge and Q. Zhu. Target-based automated matching of multiple terrestrial laser scans for complex forest scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179:1–13, 2021.

[14] T. Hilker, N.C. Coops, D.S. Culvenor, G. Newnham, M.A. Wulder, C.W. Bater, and A. Siggins. A simple technique for co-registration of terrestrial lidar observations for forestry applications. *Remote sensing letters*, 3(3):239–247, 2012.

[15] M. Holopainen, V. Kankare, M. Vastaranta, X. Liang, Y. Lin, M. Vaaja, X. Yu, J. Hyyppä, H. Hyyppä, H. Kaartinen, A. Kukko, T. Tanhuanpää, and P. Alho. Tree mapping using airborne, terrestrial and mobile laser scanning – a case study in a heterogeneous urban forest. *Urban Forestry & Urban Greening*, 12(4):546–553, 2013.

[16] M. Hussein, M. Renner, M. Watanabe, and K. Iagnemma. Matching of ground-based lidar and aerial image data for mobile robot localization in densely forested environments. In *IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, pages 1432–1437. IEEE, 2013.

[17] D.S.C. J L Lovell, D L.B. Jupp and N.C. Coops. Using airborne and ground-based ranging lidar to measure canopy structure in australian forests. *Canadian Journal of Remote Sensing*, 29(5):607–622, 2003.

[18] E. Jelavic, D. Jud, P. Egli, and M. Hutter. Towards autonomous robotic precision harvesting: Mapping, localization, planning and control for a legged tree harvester, 2021.

[19] V. Kankare, J. Vauhkonen, T. Tanhuanpää, M. Holopainen, M. Vastaranta, M. Joensuu, A. Krooks, J. Hyyppä, H. Hyyppä, P. Alho, and R. Viitala. Accuracy in estimation of timber assortments and stem distribution – a comparison of airborne and terrestrial laser scanning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97:89–97, 2014.

[20] E.N. Marius Hauglin, Vegard Lien and T. Gobakken. Geo-referencing forest field plots by co-registration of terrestrial and airborne laser scanning data. *International Journal of Remote Sensing*, 35(9):3135–3149, 2014.

[21] A. Murtiyoso, S. Holm, H. Riihimäki, A. Krucher, H. Griess, V.C. Griess, and J. Schweier. Virtual forests: a review on emerging questions in the use and application of 3d data in forestry. *International Journal of Forest Engineering*, 35(1):29–42, 2024.

[22] K. Omasa, F. Hosoi, T.M. Uenishi, Y. Shimizu, and Y. Akiyama. Three-Dimensional Modeling of an Urban Park and Trees by Combined Airborne and Portable On-Ground Scanning LIDAR Remote Sensing. *Environmental Modeling & Assessment*, 13(4):473–481, November 2008.

[23] C. Paris, D. Kelbe, J. van Aardt, and L. Bruzzone. A novel automatic method for the fusion of als and tls lidar data for robust assessment of tree crown structure. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):3679–3693, 2017.

[24] P. Polewski, W. Yao, L. Cao, and S. Gao. Marker-free coregistration of uav and backpack lidar point clouds in forested areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147:307–318, 2019.

[25] A. Proudman, M. Ramezani, S.T. Digumarti, N. Chebrolu, and M. Fallon. Towards real-time forest inventory using handheld lidar. *Robotics and Autonomous Systems*, 157:104240, 2022.

[26] J. Shao, W. Yao, P. Wan, L. Luo, J. Lyu, and W. Zhang. Efficient divide-and-conquer registration of UAV and ground lidar point clouds through canopy shape context. *ArXiv*, abs/2201.11296, 2022.

[27] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(4):376–380, 1991.

[28] Y. Wang, J. Pyörälä, X. Liang, M. Lehtomäki, A. Kukko, X. Yu, H. Kaartinen, and J. Hyyppä. In situ biomass estimation at tree and plot levels: What did data record and what did algorithms derive from terrestrial and aerial point clouds in boreal forest. *Remote Sensing of Environment*, 232:111309, 2019.

[29] J.C. White, N.C. Coops, M.A. Wulder, M. Vastaranta, T. Hilker, and P. Tompalski. Remote sensing technologies for enhancing forest inventories: A review. *Canadian Journal of Remote Sensing*, 42(5):619–641, 2016.