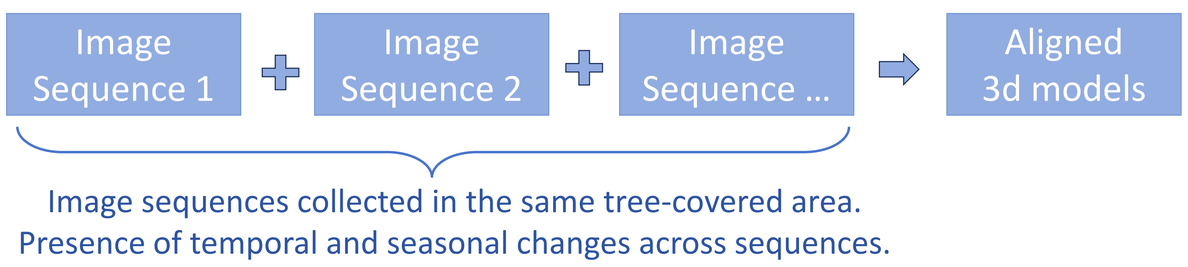
Temporal- and Viewpoint-Invariant Registration for Under-Canopy Footage using Deep-Learning-based Bird's Eye View Prediction

Jiawei Zhou[1], Ruben Mascaro[1], Cesar Cadena[2], Margarita Chli[1], Lucas Teixeira[1]
[1]V4RL - University of Cyprus and ETH Zürich [2]RSL - ETH Zürich

1. Motivation

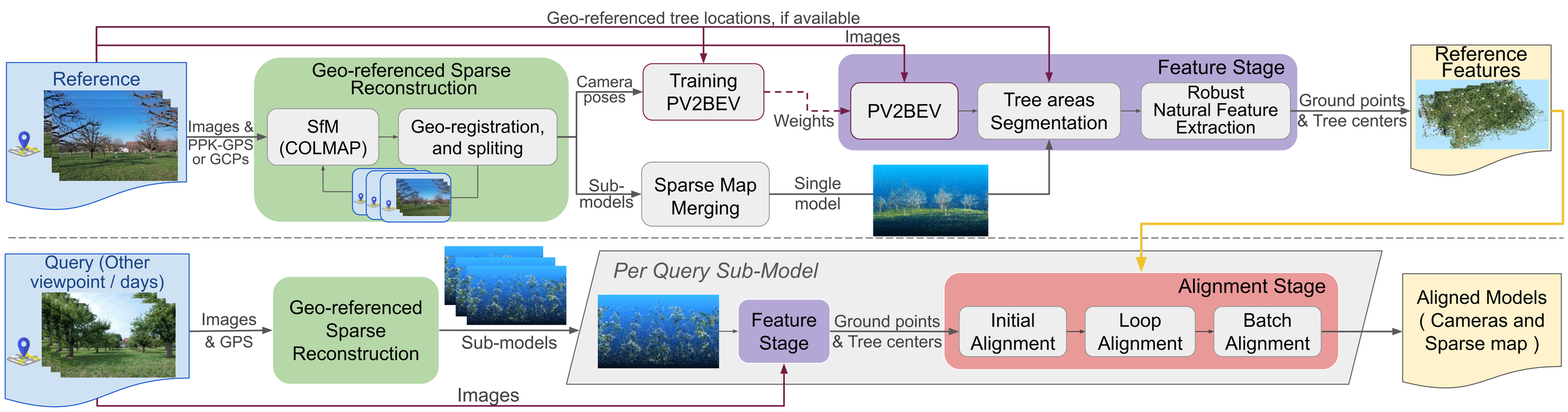
Achieving precise registration between successive data capture sessions is essential in smart farming. This work focuses on aligning noisy, sparse point clouds reconstructed from multiple geo-referenced RGB image sequences collected in areas covered by trees.



2. Contribution

- Deployment of learning-based perspective to bird's-eye view image conversion for estimating the position of trees.
- Feature extraction strategy that targets the clean segmentation of trunk and ground points.
- Iterative alignment process that alternates focus between the segmented ground and the estimated tree centers for enhanced precision.
- Strategic division of the 3D models into batches, followed by a local refinement phase to fine-tune the alignment results for each batch.

3. Method



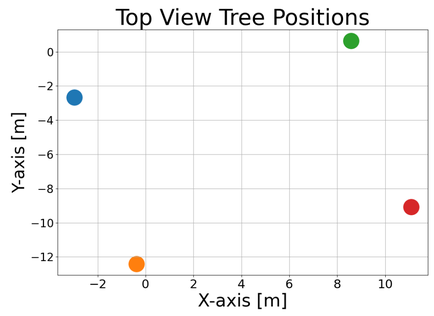
I Sparse Reconstruction

Reconstruction using **SfM**

Initial alignment with **GPS**

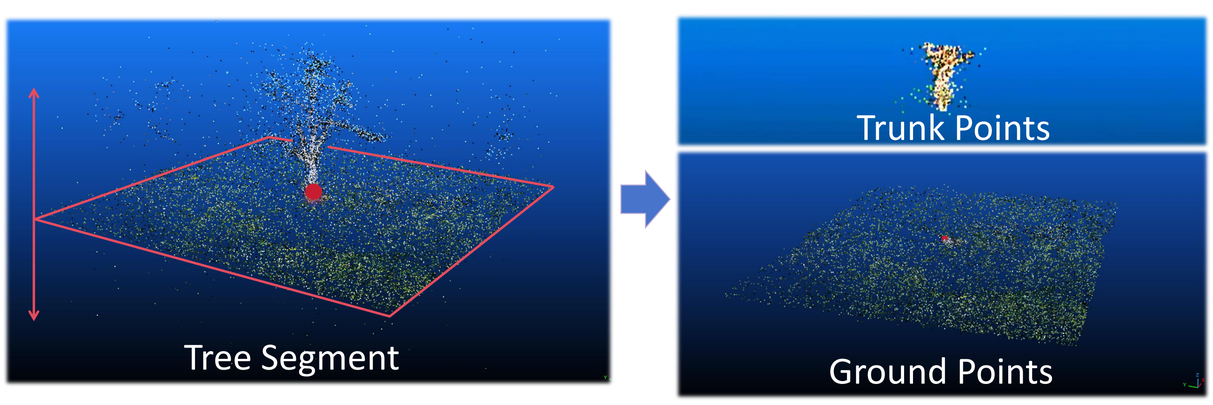
II Perspective View to Bird's Eye View

PV2BEV



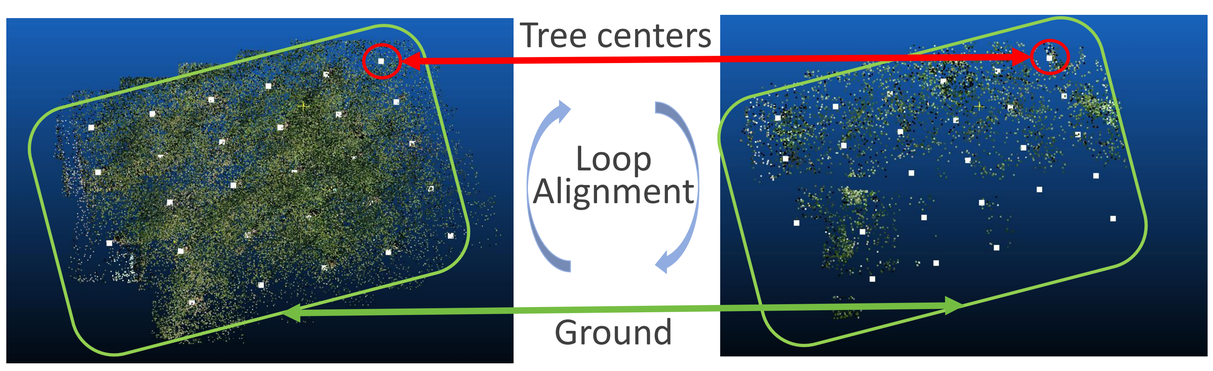
III Feature Stage

1. **Tree Area Segmentation:** Area extraction around each tree center.
2. **Robust Natural Feature Extraction:** Slicing and Gaussian fitting for trunk segmentation, and DBSCAN for outlier removal.



IV Alignment Stage

1. **Initial Alignment**: ICP alignment of extracted ground points.
2. **Loop Alignment:** Alternation between tree centers and ground points.



3. **Batch Alignment**: Clustering of tree centers into well-defined groups, which are refined locally to address accumulated drift.

4. Experiments

Evaluation Metric	a) Different Seasons Case						b) Perpendicular Viewpoints Case						c) Opposite Viewpoints Case					
	Ours	ICP[1]	FGR[2]	FICP[3]	RICP[3]	TEASER[4]	Ours	ICP[1]	FGR[2]	FICP[3]	RICP[3]	TEASER[4]	Ours	ICP[1]	FGR[2]	FICP[3]	RICP[3]	TEASER[4]
Camera Translation Error [m]	**0.44**	1.77	4.66	5.88	5.65	22.58	**0.26**	0.88	6.49	4.62	0.83	46.28	**0.40**	15.83	118.37	23.75	14.80	200.77
Camera Rotation Error [°]	**0.74**	2.48	0.77	6.82	7.50	19.62	**0.18**	0.39	1.33	10.60	0.28	165.44	**0.54**	3.64	1.29	3.92	2.35	158.65
Tree Centers 3D Error [m]	**0.28**	1.88	4.77	5.39	4.95	21.77	**0.20**	0.84	6.53	4.94	0.73	36.00	**0.32**	14.93	107.92	22.91	14.06	203.17

References:

1. P.J. Besl and N.D. McKay, "A method for registration of 3-D shapes", TPAMI 1992
2. Q.Y. Zhou et al., "Fast Global Registration", ECCV 2016
3. J. Zhang et al., "Fast and Robust Iterative Closest Point", TPAMI 2021
4. H. Yang et al., "TEASER: Fast and Certifiable Point Cloud Registration, T-RO 2020

Code and datasets available!