

# Efficient View Planning Guided by Previous-Session Reconstruction for Repeated Plant Monitoring

Sicong Pan, Luca Lobefaro, Moein Taherkhani, Xuying Huang, Rohit Menon, Cyrill Stachniss, Maren Bennewitz

**Abstract**—Repeated plant monitoring is essential for tracking crop growth, and 3D reconstruction enables consistent comparison across monitoring sessions. However, rebuilding a 3D model from scratch in every session is costly and overlooks informative geometry already observed previously. We propose efficient view planning guided by a previous-session reconstruction, which reuses a 3D model from the previous session to improve active perception in the current session. Based on this previous-session reconstruction, our method replaces iterative next-best-view planning with one-shot view planning that selects an informative set of views and computes the globally shortest execution path connecting them. Experiments on real multi-session datasets show that, in greenhouse crop-row monitoring, the proposed method achieves comparable or higher surface coverage with fewer executed views and shorter paths than baseline planners.

## I. INTRODUCTION

Monitoring crop growth is essential for understanding plant development and supporting timely agricultural decision-making, yet repeated 3D monitoring remains labor-intensive. To compare plant geometry consistently across monitoring sessions, robotic systems increasingly rely on 3D reconstruction. Recent work has explored active perception for plant reconstruction, often by solving next-best-view (NBV) planning problems to mitigate self-occlusions in complex plant structures [1]. In greenhouse production, however, reconstruction is performed repeatedly over time, and many existing systems still treat each session independently, restarting the pipeline from scratch and discarding informative geometry already observed in previous sessions.

We ask whether a 3D reconstruction from a previous monitoring session can be exploited to make view planning for the current session more efficient. Building on our earlier workshop study on individual-plant monitoring [3], we adopt an efficient one-shot view planning strategy, motivated by the fact that globally selected view sets [4] can reduce redundant observations and acquisition cost compared with iterative NBV planning when path efficiency needs to be considered jointly.

In this work, we focus on extending this idea to greenhouse crop rows. Compared with individual plants, greenhouse-row monitoring takes place at meter scale, involves restricted fields of view, and places stronger emphasis on path-efficient execution. To address this setting, we first obtain a coarse current-session reconstruction by passive mapping [2], then align the previous-session reconstruction to it via non-rigid registration [5], and conservatively inflate the transferred geometry to improve robustness to growth and residual alignment errors. The resulting approximation is used to jointly plan camera positions and viewing directions, guiding observations toward under-observed structures while reducing unnecessary motion. Fig. 1 illustrates our approach.

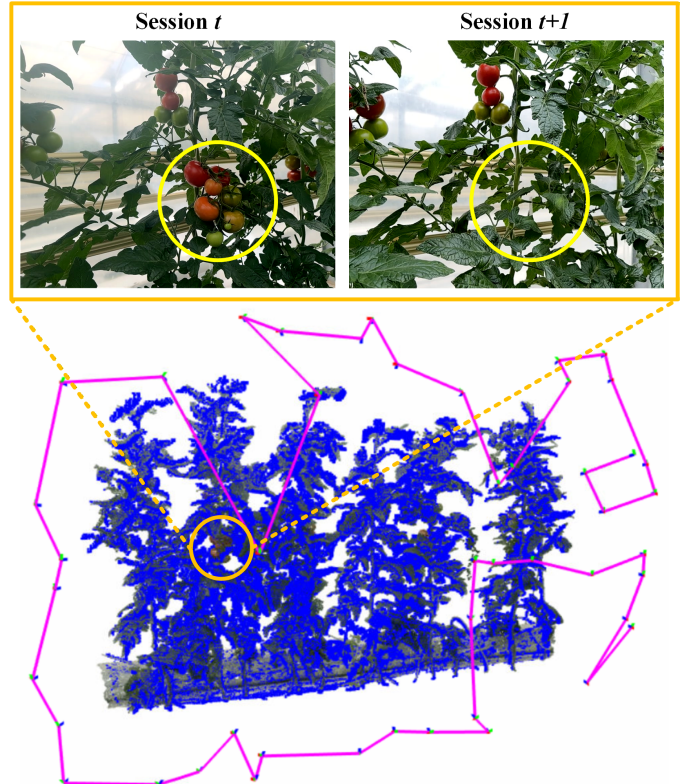


Fig 1: Example of view planning guided by a previous-session reconstruction in the greenhouse-row setting. Top: Representative RGB views from two monitoring sessions ( $t$  and  $t+1$ ) illustrate session-to-session changes in foliage and fruit configuration (yellow circles), e.g., due to plant growth and routine harvesting. Bottom: We first build a coarse reconstruction (blue) for the current session and then align the previous-session reconstruction to obtain an actionable approximation (colored) that highlights under-observed structures and guides planning toward poorly covered regions. Based on this approximation, the proposed method selects informative views and computes the globally shortest execution path connecting them (purple) for efficient data acquisition with an eye-in-hand camera on a robotic arm using a greenhouse mobile robot platform [2].

The main progress presented in this work is threefold. First, we extend one-shot view planning guided by previous-session reconstruction from individual plants to repeated monitoring of greenhouse crop rows. Second, we introduce a new multi-session greenhouse crop-row dataset<sup>1</sup>. Third, we show that the proposed row extension, which jointly plans camera positions and viewing directions, improves the coverage–effort trade-off over baseline planners. To facilitate reproducibility, we release our implementation on GitHub<sup>2</sup>.

<sup>1</sup>[www.kaggle.com/datasets/sicongpan/greenhouse-multi-session-row-dataset](http://www.kaggle.com/datasets/sicongpan/greenhouse-multi-session-row-dataset)

<sup>2</sup>[www.github.com/HumanoidsBonn/TPVP](https://www.github.com/HumanoidsBonn/TPVP)

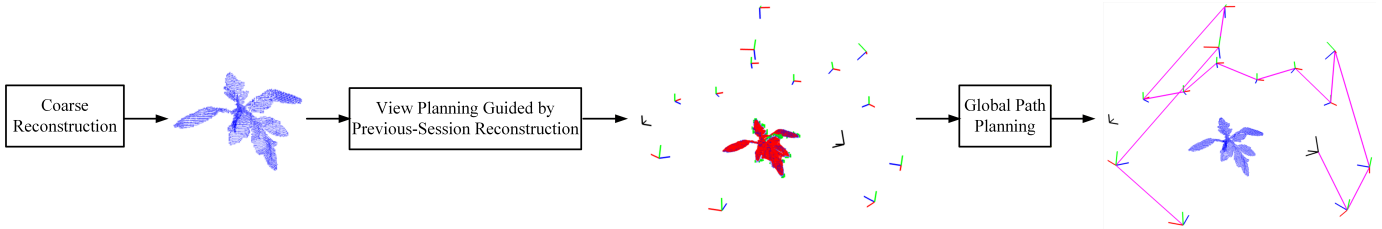


Fig 2: Overview of the proposed one-shot view planning pipeline. Shown on an individual-plant example, the approach starts from a coarse reconstruction (blue); in this example, it is obtained from an initial view fused with one warm-up NBV observation (black). We then transfer the previous-session reconstruction (red) to obtain actionable geometric guidance and select a minimal view set (colored axes) that covers the plant. Finally, the selected views are ordered by global path planning, yielding a global execution path (purple) that reduces travel during data collection.

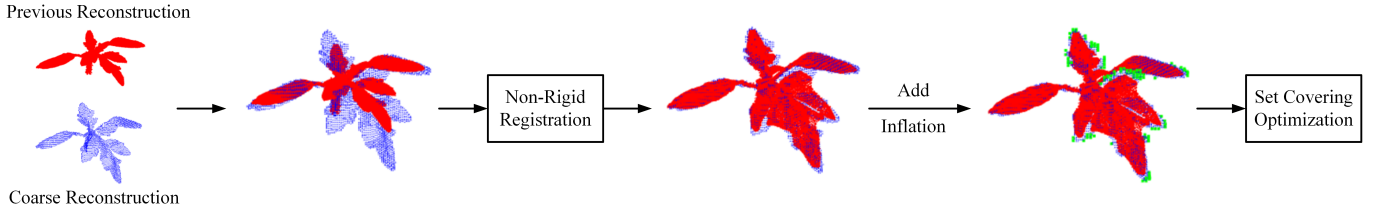


Fig 3: Overview of view planning guided by a previous-session reconstruction. Shown on an individual-plant example, the approach first aligns the previous-session plant reconstruction (red) with a coarse reconstruction of the current session (blue) via non-rigid registration, yielding a transferred geometric approximation. It then applies a conservative inflation (green) to account for inter-session changes and residual registration errors. Finally, a set-cover optimization is solved to select the minimum view set that achieves the desired surface coverage of the inflated planning geometry.

## II. OUR APPROACH

We study view planning for *repeated plant monitoring* guided by a previous-session reconstruction. Given a plant 3D model built from the previous monitoring session and a coarse 3D reconstruction in the current session, the goal is to plan a sequence of sensor views that efficiently reconstructs the plant. In greenhouse crop rows, each candidate view is defined by both a camera position and a 3D look-at target.

We use a one-shot view planning pipeline, illustrated in Fig. 2, which first selects an informative set of views and then computes a globally short execution path through them.

An overview of the proposed view planning guided by previous-session reconstruction is shown in Fig. 3 and consists of three core steps: (1) *non-rigid registration* to align the previous-session reconstruction with the current-session coarse reconstruction under plant deformation; (2) *conservative inflation* to obtain a robust planning geometry that accounts for inter-session changes and residual registration errors; and (3) *set-cover-based view selection* to choose a minimal set of views that covers the inflated geometry approximation.

## III. EXPERIMENTS

We evaluate the proposed greenhouse-row extension on a new multi-session crop-row dataset, focusing on coverage-effort trade-offs, the role of transferred geometry in look-at planning, and robustness to execution noise.

1) *Test Platform*: We use the view-planning simulator of [4], which renders visible surface points by OctoMap-based ray casting from candidate views defined by a 3D camera position and a 3D look-at target.

2) *Evaluation Metrics*: We report oracle-visible surface coverage, the number of executed views, and the travel cost

measured as the accumulated Euclidean path length connecting consecutive executed views. Following [4], coverage is normalized by the oracle visible set of surface points that are observable from the feasible view space, so methods are not penalized for inherently unreachable regions caused by self-occlusion or platform constraints.

3) *Data and Protocol*: We collected a greenhouse crop-row dataset containing two crop types (tomato and cherry tomato), captured over five monitoring sessions spanning two weeks (Sep. 15, 18, 22, 25, and 29, 2025). In each session, one crop row per crop type was scanned using Polycam<sup>3</sup> on a LiDAR-equipped iPhone/iPad Pro, providing RGB images, per-frame depth maps, and estimated camera poses. For each session, we reconstruct a reference 3D model with Nerfstudio<sup>4</sup> and export a point cloud representation. For reliable simulation-based evaluation, we lightly post-process the reconstructions by cropping the target row region, removing background structures and supports, filtering outliers, and aligning all sessions to the first capture day using landmark-guided ICP on static greenhouse structures. We evaluate four consecutive session pairs (Sep. 15→18, 18→22, 22→25, and 25→29), resulting in  $2 \times 4 = 8$  test cases.

4) *View Space Configuration*: We define the feasible view space in the greenhouse-row setting based on oriented bounding boxes (OBBs) of the reconstructed row. Specifically, we construct (i) a *plant OBB* that bounds the crop structure and (ii) a *view OBB* that approximates the accessible camera workspace on the aisle side, consistent with typical greenhouse monitoring platforms [2]. Within the view OBB, we discretize candidate camera positions by Poisson-disk sampling and use

<sup>3</sup><https://poly.cam/>

<sup>4</sup><https://docs.nerf.studio/>

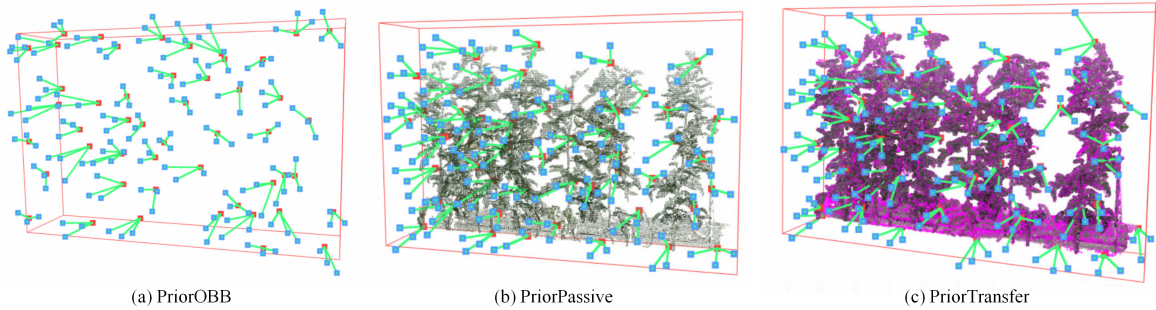


Fig 4: Candidate 6D view generation under the prior ladder for look-at sampling. (a) PriorOBB: look-at targets are sampled inside the plant OBB (red). (b) PriorPassive: look-at targets are sampled from surface frontiers of the passive coarse reconstruction (raw plant colors). (c) PriorTransfer: look-at targets are sampled on the transferred surface obtained from the previous-session reconstruction (purple), excluding regions already observed in the coarse reconstruction. In all cases, camera positions (blue) are drawn from the same rail-side view OBB and assigned to look-at targets using a  $k$ -NN rule with balanced allocation to form 6D candidates (green directions). For clarity, the visualization shows a subset of 50 targets and  $k = 3$  assignments; experiments use 100 targets with  $k = 8$  (800 candidates) per strategy.

Look-At Sampling Strategy	PriorOBB	PriorPassive	PriorTransfer
Max Reach of Oracle Visible Set (%)	94.48 $\pm$ 1.17	97.21 $\pm$ 0.90	<b>97.34 <math>\pm</math> 0.77</b>

TABLE I: Effect of geometric guidance on reachable coverage. We report the maximum reachable fraction of the oracle visible set under three look-at sampling strategies (OBB-only, passive, and previous-session transfer), using the same 1000 candidate positions (mean  $\pm$  std across test cases). Passive and transfer-based guidance improve the reachable upper bound over OBB-only sampling. The similar upper bounds of PriorPassive and PriorTransfer suggest that previous-session transfer mainly improves planning reliability rather than expanding the theoretically observable surface.

the same set of 1,000 view positions for all compared methods. Unlike the object-centric individual-plant setting, greenhouse-row monitoring requires explicit planning of viewing directions under restricted fields of view. We therefore parameterize each view by both a 3D camera position and a 3D *look-at target* within the plant OBB.

5) *Implementation Details of the Proposed Method*: In the greenhouse-row setting, we obtain the coarse current-session reconstruction from a passive warm start consisting of 15 fixed viewpoints within the view OBB, whose combined fields of view cover the plant OBB. The rendered point clouds from these views are fused and used for the previous reconstruction transfer. For view selection, we evaluate the proposed method at covering resolutions  $r_{\text{cover}} \in \{10, 5, 4, 3\}$  cm, which provide different operating points in the coverage–effort trade-off.

6) *Baselines*: We compare against two baseline families: a sequential NBV baseline and one-shot baselines based on a prior ladder for look-at generation. (i) Block-Wise NBV: We reimplement the sequential NBV pipeline of Zaenker *et al.* [1] with information gain over unknown voxels and online look-at sampling from surface frontiers. To reflect restricted greenhouse workspaces and reduce inefficient long-range back-and-forth motions, we use a block-wise execution scheme [6] that partitions the view OBB into local workspace regions. (ii) Prior ladder baselines: To isolate the effect of geometric guidance for look-at planning, we construct a three-level prior ladder, illustrated in Fig. 4, ranging from OBB-only sampling to passive-reconstruction-guided sampling and finally previous-session-transfer-guided sampling. For all three strategies, we sample 100 look-at targets and assign nearby camera positions using a balanced  $k$ -NN rule ( $k = 8$ ), yielding

Method	Number of Views	Surface Coverage (%)	Path Length (m)
Coarse Reconstruction	15	44.48 $\pm$ 1.09	<b>2.58 <math>\pm</math> 0.02</b>
Block-Wise NBV	100	73.73 $\pm$ 2.46	64.94 $\pm$ 4.14
	300	90.06 $\pm$ 2.12	174.20 $\pm$ 6.95
	500	94.70 $\pm$ 1.60	279.02 $\pm$ 7.68
PriorOBBRandom	100	61.44 $\pm$ 3.25	28.78 $\pm$ 0.85
	300	82.94 $\pm$ 2.11	60.25 $\pm$ 1.28
	500	89.69 $\pm$ 1.44	87.09 $\pm$ 1.22
PriorPassiveRandom	100	71.99 $\pm$ 2.69	28.93 $\pm$ 0.74
	300	88.15 $\pm$ 1.87	57.99 $\pm$ 0.77
	500	93.45 $\pm$ 1.17	83.04 $\pm$ 0.67
PriorTransferRandom	100	72.43 $\pm$ 2.20	28.91 $\pm$ 0.86
	300	88.37 $\pm$ 1.40	57.98 $\pm$ 0.59
	500	93.57 $\pm$ 1.07	83.07 $\pm$ 0.48
Ours	71.3 $\pm$ 7.5 (10 cm)	69.89 $\pm$ 2.58	23.73 $\pm$ 1.64
	255.9 $\pm$ 19.1 (5 cm)	90.55 $\pm$ 1.53	53.21 $\pm$ 2.28
	345.8 $\pm$ 18.9 (4 cm)	93.26 $\pm$ 1.30	64.54 $\pm$ 2.36
	476.0 $\pm$ 17.1 (3 cm)	<b>95.46 <math>\pm</math> 1.01</b>	80.28 $\pm$ 2.19

TABLE II: Coverage–effort trade-off in the greenhouse-row setting. We report oracle-visible surface coverage, executed views, and path length for a passive warm start (Coarse Reconstruction), a sequential baseline (Block-Wise NBV), one-shot baselines using different look-at sampling strategies with random selection (PriorOBB/Passive/Transfer), and our method. Block-Wise NBV is sequential, so results are shown at fixed view budgets (100/300/500), whereas our method solves a one-shot set-cover problem and outputs a variable number of views. All one-shot methods (including ours) use the same global path planner; only the view-selection step differs. The four Ours rows correspond to different covering resolutions (voxel sizes 10/5/4/3 cm), yielding multiple operating points on the coverage–effort frontier. For methods that use random selection, the experiments were repeated five times for evaluation. For methods that use the passive warm start, the reported views and path lengths include the 15 fixed passive viewpoints. Values are mean  $\pm$  std across test cases. Overall, our method achieves substantially higher coverage than random one-shot baselines at comparable path lengths, and it reaches NBV-level coverage with dramatically shorter path lengths.

800 candidate 6D views per strategy. Since the OBB-only and passive-only strategies do not provide a reliable full-surface prior for set-cover reasoning, we use random selection over their generated candidates as one-shot baselines, denoted *PriorOBBRandom*, *PriorPassiveRandom*, and *PriorTransferRandom*. Our method uses the transfer-prior strategy together with set-cover-based view selection.

7) *Analysis on Upper-Limit Surface Coverage*: Table I compares the maximum reachable fraction of the oracle visible set under the three look-at sampling strategies.

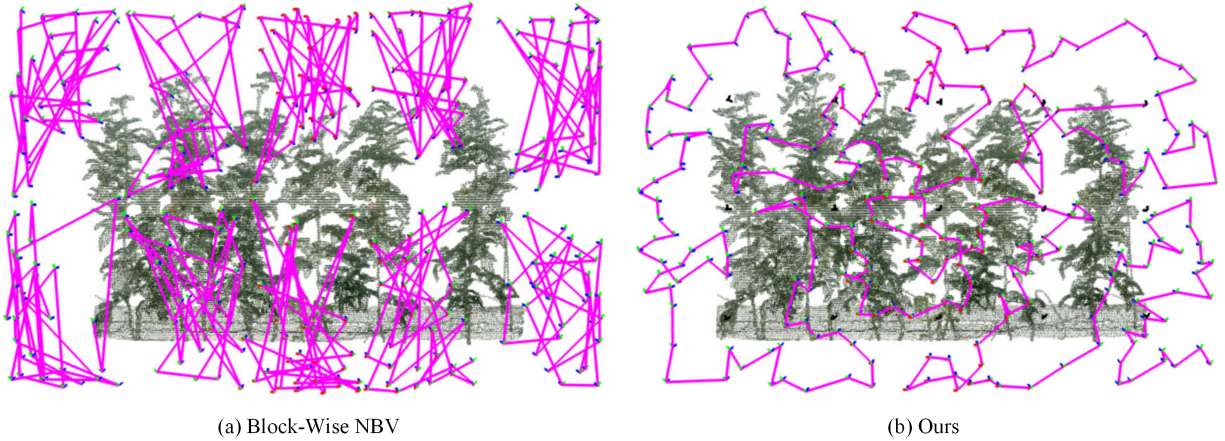


Fig 5: Qualitative comparison of execution paths in the greenhouse-row setting. Block-wise NBV (a) and Ours (b) are shown at a matched coverage level: NBV reaches 92.42% surface coverage with 300 views, while Ours (5 cm) reaches 92.68% with 280 views.

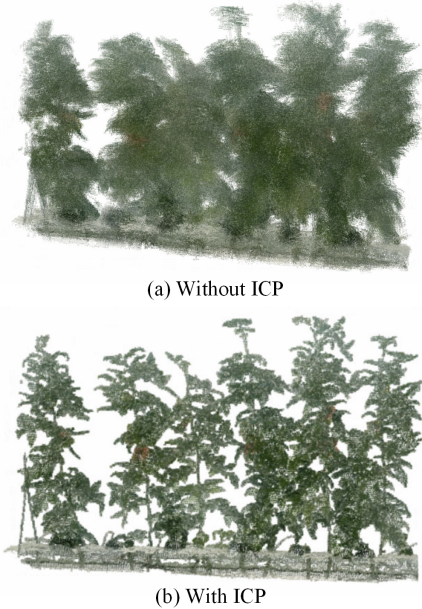


Fig 6: Pose-noise robustness example in the greenhouse-row setting (our method at 4 cm). We perturb the executed active-view poses and fuse the resulting observations. Without refinement (a), pose noise leads to drift and severe ghosting in the fused point cloud. With lightweight post-fusion ICP (b), the reconstruction is largely re-aligned, recovering sharp plant structures.

8) *Evaluation of View Planning Performance:* Table II compares the coverage–effort trade-off in greenhouse-row monitoring. Block-Wise NBV improves coverage with increasing view budget, but at a dramatic motion cost. In contrast, stronger look-at priors consistently improve one-shot performance at a fixed budget, with passive- and transfer-guided sampling clearly outperforming OBB-only random sampling. Our method further improves the trade-off by replacing random selection with set-cover-based view selection. At 5 cm resolution, it reaches 90.55% coverage with a 53.21 m path, compared with 90.06% and 174.20 m for Block-Wise NBV at 300 views; at 3 cm resolution, it achieves the highest coverage in the table (95.46%) while maintaining a much shorter path. Fig. 5 shows a matched-coverage example with a substantially

more compact execution path.

9) *Pre-Deployment Pose-Noise Robustness:* To assess robustness to execution errors before real-robot deployment, we perturb the executed active-view poses and evaluate the resulting reconstruction quality with and without lightweight ICP refinement, as shown in Fig. 6. ICP makes reconstruction quality much less sensitive to the pose noise.

10) *Discussion:* An important next step is end-to-end validation on a real greenhouse mobile manipulation platform under live sensing and actuation conditions. The main computational bottleneck is visibility evaluation by ray casting before set-cover selection, which can be naturally parallel and substantially accelerated in practice.

#### IV. CONCLUSION

We presented one-shot view planning guided by previous-session reconstruction to repeated monitoring of greenhouse crop rows. Experiments on real multi-session datasets show that the proposed method improves the coverage–effort trade-off over baseline planners by providing actionable geometric guidance for look-at planning and global view selection. These results support previous-session reconstruction as a useful geometric prior for efficient repeated plant monitoring.

#### REFERENCES

- [1] T. Zaenker, C. Smitt, C. McCool, and M. Bennewitz, “Viewpoint planning for fruit size and position estimation,” in *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2021.
- [2] C. Smitt, M. Halstead, T. Zaenker, M. Bennewitz, and C. McCool, “Pathobot: A robot for glasshouse crop phenotyping and intervention,” in *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, 2021.
- [3] S. Pan, X. Huang, and M. Bennewitz, “Temporal-prior-guided view planning for periodic 3d plant reconstruction,” 2025, Active Perception Workshop at ICRA. Available at: [https://activep-ws.github.io/accepted\\_papers/5\\_Temporal\\_Prior\\_Guided\\_View\\_P.pdf](https://activep-ws.github.io/accepted_papers/5_Temporal_Prior_Guided_View_P.pdf).
- [4] S. Pan, H. Hu, H. Wei, N. Dengler, T. Zaenker, and M. Bennewitz, “Integrating one-shot view planning with a single next-best view via long-tail multiview sampling,” *IEEE Transactions on Robotics (T-RO)*, 2024.
- [5] L. Lobefaro, M. V. Malladi, T. Guadagnino, and C. Stachniss, “Spatio-temporal consistent mapping of growing plants for agricultural robots in the wild,” in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 6375–6382.
- [6] A. I. Jose, S. Pan, T. Zaenker, R. Menon, S. Houben, and M. Bennewitz, “Go-vmp: Global optimization for view motion planning in fruit mapping,” *arXiv preprint arXiv:2503.03912*, 2025.