

CropCraft: Complete Structural Characterization of Crop Plants From Images

Supplementary Material

Abstract

In the following supplementary material, we provide additional experimental details (Sec. A), additional experimental results (Sec. B), and a discussion of limitations (Sec. C). We invite readers to watch the supplementary video (supp.mp4) for further visualization.

A. Additional Experimental Details

A.1. Additional Procedural Model Details

Soybean Model. Our procedural soybean model is based on the mCanopy model from Song et al. [4]. Each plant consists of a main stem with up to 14 trifoliate nodes, and up to 6 branch stems with up to 2 nodes each. Unifoliate nodes and cotyledons are ignored for simplicity. Each node consists of three coplanar leaves, with the middle leaf connected by a small petiole, and is connected to its stem by a larger petiole. The leaf length (and width), the petiole length, the angle with the stem, and the distance between nodes are variable, as they are affected by the multiplier parameters. On the main stem, when all the multipliers are set to 1, the variables follow a profile from top to bottom based on field data measured by Song et al. [4]. Branches start growing with every new main stem node starting from the 8th, with one branch per main stem node starting from the bottom and up to the 6th. The number of nodes per branch is selected randomly according to the distribution defined by Song et al. [4]. The variables for the branch nodes are the same across nodes (but are affected by the multipliers). When the number of nodes is not an integer, it is rounded up and the last node is scaled down in size linearly according to the remainder. Additionally, random noise is added to each petiole angle following a Gaussian distribution with mean 0 and standard deviation 5 degrees. Each node grows on alternating sides of their stem, with additional azimuthal rotation following a Gaussian distribution with mean 0 and standard deviation 60 degrees. The range of possible values for each model parameter are provided in Tab 1.

Maize Model. Our procedural maize model is based on the coupled maize model from Qian et al. [3]. Each plant consists of a main stem with up to 18 nodes with one leaf each. Each leaf consists of 10 trapezoidal segments that bend towards the ground following a polynomial function that depends on the leaf order [3]. To allow for different bending angles, we introduce a parameter that additively shifts all of the leaf orders by the same amount. The leaf

Parameter	Range
Leaf length multiplier	0.5 - 1.5
Petiole length multiplier	0.5 - 2.0
Petiole angle multiplier	0.5 - 4.0
Internode length multiplier	0.5 - 2.0
Number of nodes	1.0 - 14.0

Table 1. Soybean model parameter ranges.

lengths, widths, and internode lengths follow a profile from bottom to top based on field data measured in the U.S. Midwest. Each node grows on alternating sides of their stem, with additional azimuthal rotation following a Gaussian distribution with mean 0 and standard deviation 60 degrees. The range of possible values for each model parameter are provided in Tab 2.

Parameter	Range
Leaf length multiplier	0.8 - 1.2
Leaf order shift	-4.0 - 4.0
Internode length multiplier	0.8 - 1.2
Number of nodes	1.0 - 18.0

Table 2. Maize model parameter ranges.

A.2. Additional Row Fitting and Depth Rendering Details

Our row fitting procedure operates on a point cloud sampled from the NeRF, using the sampling implementation in Nerfstudio [5]. For the soybean data, we sample 100K points in a $2\text{ m} \times 2\text{ m} \times 3\text{ m}$ box centered at $(0, 0, -1.5)\text{m}$. Voxel-downsampling is performed at a resolution of 1 cm. Color-thresholding is performed in LAB color space. All points with $L^* < 0$ and $b^* < 1$ are removed. Then, points with $a^* < 2$ are classified as ground points, and the rest are classified as plant points. RANSAC is used to fit a plane to the ground points using an inlier threshold of 5 cm and 1000 max iterations. For row fitting, the points whose distance from the ground are below the 50th percentile are discarded. RANSAC is applied sequentially to fit lines with an inlier threshold of 20 cm until the number of points remaining is either less than 1000 or less than 20% of the starting number. We then take the row with the most inliers and fit a line through the inliers again using least-squares fitting. The camera pose for depth rendering is set to be directly above the mean of the inlier points at a height of 1 m from the ground, facing downwards and with the camera's

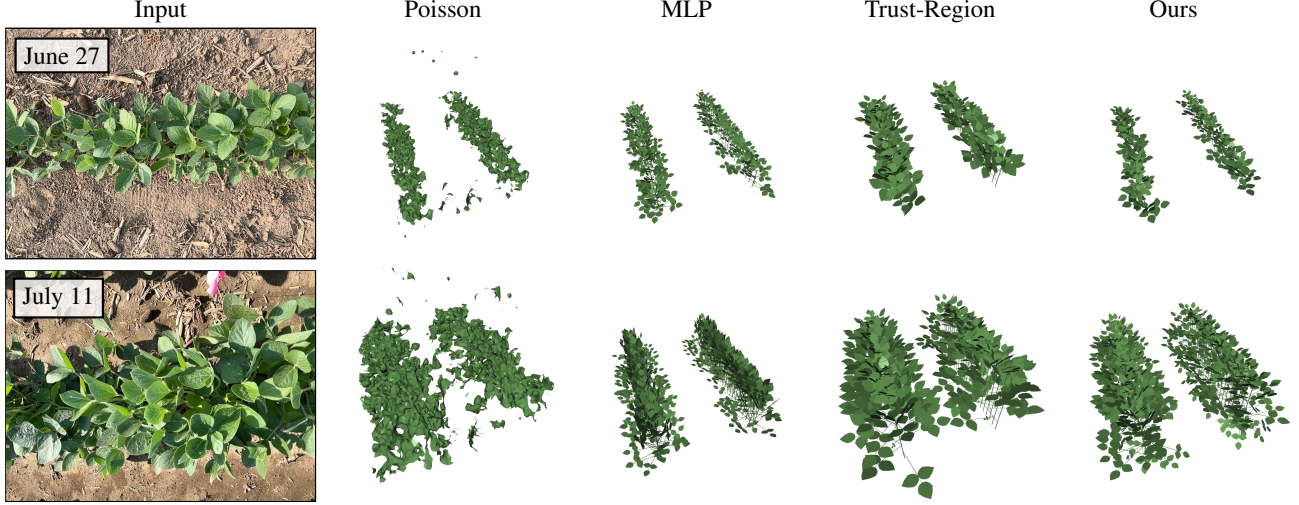


Figure 1. **Qualitative reconstruction comparison.** We compare example reconstructions from each of the baseline methods. The first column shows example images from the multi-view data collected at each time point, and the remaining columns show the corresponding mesh reconstructions.

x -axis aligned with the row line. If the number of initial plant points is greater than 75% of the total, we instead use the 70th percentile for distance from the ground, a row inlier threshold of 25 cm, and a render height of 1.25 m. The render resolution is 994×738 with a vertical FOV of 50 degrees.

For the maize data, we sample 5M points in a $2 \times 2 \times 2$ box around the origin in the scaled scene. The L^* , a^* , and b^* thresholds are set to 32, 0, and 0 respectively. The ground-inlier threshold is set to 10 cm and the render height is set to 5 m. The row fitting is done in a ROI with the shape of a cylinder with radius 2 m whose axis is perpendicular to the ground plane.

A.3. Mask and Histogram Details

The foreground mask M_{obs} is obtained by color-thresholding on the RGB render combined with thresholding on 3D coordinates. For soybean, points are marked as background if their a^* channel is less than -8 and their z -coordinate is greater than 0.25 m short of the render height, or their z -coordinate is less than 0.1 m, or their y -coordinate is at least 0.5 m away from zero. For maize, points are marked as background if a^* channel is less than -8 and their L^* channel is greater than 40 and their z -coordinate is greater than 2 m short of the render height, or their z -coordinate is less than 0.1 m, or their y -coordinate is at least 3 m away from zero. M_{pred} is obtained trivially by not rendering any background.

For soybean, the depth histogram has 20 equally spaced bins from 0.1 m to the render height. The lateral histogram has 10 equally spaced bins from zero to half the render height. The depth derivative histogram has 10 equally spaced bins from zero to 0.004. A Gaussian blur with ker-

nel size 25 is applied to the depth map before the calculation of Sobel derivatives, since the NeRF rendering is generally blurrier than the procedural mesh rendering. For maize, the depth histogram has 10 equally spaced bins from 2 m to the render height, and the blur kernel size is 55.

A.4. Baseline Implementation Details

We provide implementation details for the baseline reconstruction methods:

- **Poisson:** We use the Poisson surface reconstruction implementation by Open3D [6] integrated into Nerfstudio [5]. For soybean, we use 100K points with 50K faces in the same bounding box as in our method and filter out points whose z -coordinates are lower than 0.2 plus the 1st percentile. For maize, we use a $4 \text{ m} \times 4 \text{ m}$ bounding column and 0.1 m instead of 0.2.
- **MLP:** The MLP input consists of the same histograms used for optimization, concatenated with the render height and mask area. The training data is generated by uniformly randomly sampling parameter values and computing the corresponding MLP inputs. For soybean, we sample 10K pairs with render height 1.0 m and 10K pairs with render height 1.25 m. For maize, we sample 20K pairs with render height 5.0 m. The MLP has two hidden layers of size 512, and we train for 200 epochs using an Adam optimizer with $\alpha = 0.001$, $(\beta_1, \beta_2) = (0.9, 0.999)$, and weight decay of 10^{-5} .
- **Trust-Region:** We use the trust-region method implemented by the “trust-constr” option of SciPy’s `optimize.minimize` function [2], with the maximum number of iterations set to 500. For fair comparison with our Bayesian optimization method, we run it 10

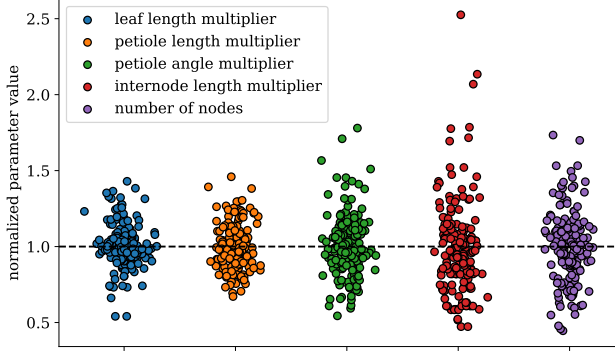


Figure 2. **Jitter plots of optimized parameter values.** The values are normalized by dividing by the per-scene average. The individual optimization runs generally do not stray far from the average, but taking the average ultimately gives a better result.

times with different random initializations. We take the solution with the lowest loss value instead of averaging, since we find it gives better performance.

B. Additional Results

B.1. Additional Qualitative Results

We provide qualitative comparisons of the different reconstruction baselines in Fig. 1. We observe that Poisson reconstruction gives incomplete and implausible canopy shape, motivating the use of inverse procedural modeling. Compared to the other procedural generation baselines, our method generally produces the parameters that best reflects the true real-world canopy. In the examples shown, the MLP produces leaves that are too small, and the trust-region optimization produces leaves that are too large.

B.2. Synthetic Scenes Evaluation

We perform additional experiments using synthetic scenes to evaluate additional aspects of canopy structure for which it is costly to obtain ground-truth in the real world. We generated 30 synthetic scenes for both soybean and maize by uniformly randomly sampling procedural model parameters and evaluated our method and the Trust-Region baseline on these scenes. We introduce two metrics to measure *stem structure* and *topology*: **Node Distance Percent Error (NDPE)** is the percent error on the average stem segment length between nodes (branching points), and **Node Count Percent Error (NCPE)** is the percent error on the total number of nodes. The results (Tab. 3) show that our method can capture these structural variables fairly accurately (but not perfectly). Note that the Poisson baseline cannot be applied here as it has no stem structure, and the MLP will simply overfit.

	Soybean			Maize		
	LAIE (↓)	NDPE (↓)	NCPE (↓)	LAIE (↓)	NDPE (↓)	NCPE (↓)
T-R	2.49	0.48	0.91	1.46	0.13	0.70
Ours	0.97	0.21	0.23	0.13	0.05	0.08

Table 3. **Synthetic data evaluation. Bold: best.**

	Soybean		Maize	
	LAIE (↓)	LAIPE (↓)	LAIE (↓)	LAIPE (↓)
S-C	17.85	1.42	52.12	6.74
C-C	53.03	6.13	46.96	5.91
Ours	0.69	0.15	0.97	0.26

Table 4. **Procedural model comparison. Bold: best.**

B.3. Procedural Model Comparison

We test the performance of our method with different, simpler procedural models to investigate the importance of the procedural model’s design. We implemented the Spherical Crowns and Conical Crowns models from Helios [1] and fit them using our inverse modeling pipeline. The results (Tab. 4) show that while these models can roughly model the visible surfaces, they vastly overestimate the LAI by assuming the space underneath is densely filled with leaves. This confirms the need for grounded morphology models that capture the right assumptions about leaf arrangement.

B.4. Bayesian Optimization Performance Analysis

We visualize the distributions of the optimized parameter values before averaging in Fig. 2. The values include the 10 runs for all 18 soybean scenes and are normalized by the averaged result. We observe that there is a moderate amount of variance in the optimization results. We then examine how reconstruction performance changes as we average across different numbers of random initializations for Bayesian optimization in Tab. 5. Although there is no significant improvement in the leaf angle metrics, we observe that the leaf area estimates improve overall when averaging over more runs, eventually plateauing around 10-20 runs. Note that the constraints of the procedural model generally ensure that all possible combinations of parameters yield plausible 3D reconstructions, so there is no real concern over the average giving an unreasonable result.

We also examine how performance changes as the number of Bayesian optimization iterations varies in Tab. 6. Here, we are still averaging over 10 runs. In addition to the previous metrics, we report the average value of the minimized loss function and time per run. Performance does tend to improve with more iterations, but takes increasing amounts of time per iteration due to larger and larger kernel

# of runs	Soybean			
	LAIE (\downarrow)	LAIPE (\downarrow)	AME (\downarrow)	ASDE (\downarrow)
1	0.76	0.23	12.18	7.81
2	0.79	0.20	12.18	7.74
5	0.78	0.16	12.13	7.74
10	0.69	0.15	12.07	7.39
20	0.69	0.14	12.23	7.77

Table 5. **Effect of averaging solutions over runs. Bold: best.**

computations. In addition, beyond a certain point, decreases in the loss function does not necessarily translate to better reconstruction quality.

B.5. Wind Simulation

The procedurally generated meshes from our model can also be used for dynamics simulations. We show rendered frames of a wind simulation using NVIDIA PhysX in Fig. 3. Visually, the results appear to show a physically plausible animation of leaves blowing in the wind. Video visualizations can be found in the supplementary video.

C. Limitations

One limitation of our method is that the inverse procedural modeling is dependent on the NeRF reconstruction performance, which may degrade due to wind-induced leaf motion. Another limitation is that the procedural generation models used cannot model certain details, e.g. damaged leaves and non-leaf organs of the plants. A final limitation is that our RANSAC-based row-fitting method is sensitive to hyperparameters such as the inlier threshold and segmentation threshold, which may need to be tuned per dataset. Improving the robustness by leveraging learning-based methods is a potential direction for future work.

References

- [1] Brian N Bailey. Helios: A scalable 3d plant and environmental biophysical modeling framework. *Frontiers in Plant Science*, 10:1185, 2019. [3](#)
- [2] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [2](#)
- [3] Binxiang Qian, Wenjiang Huang, Donghui Xie, Huichun Ye, Anting Guo, Yuhao Pan, Yin Jin, Qiaoyun Xie, Qunjun Jiao, Biyao Zhang, et al. Coupled maize model: A 4d maize growth model based on growing degree days. *Computers and Electronics in Agriculture*, 212:108124, 2023. [1](#)
- [4] Qingfeng Song, Venkatraman Srinivasan, Steve P Long, and Xin-Guang Zhu. Decomposition analysis on soybean productivity increase under elevated co2 using 3-d canopy model reveals synergistic effects of co2 and light in photosynthesis. *Annals of botany*, 126(4):601–614, 2020. [1](#)
- [5] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin,

Kamyar Salahi, Abhik Ahuja, et al. Nerfstudio: A modular framework for neural radiance field development. In *ACM Transactions on Graphics (SIGGRAPH)*, 2023. [1](#), [2](#)

- [6] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3d: A modern library for 3d data processing. *arXiv preprint arXiv:1801.09847*, 2018. [2](#)

# of iterations	Soybean					
	LAIE (\downarrow)	LAIPE (\downarrow)	AME (\downarrow)	ASDE (\downarrow)	Average L (\downarrow)	Time (\downarrow)
300	0.73	0.18	12.32	7.77	0.0172	3m
400	0.77	0.17	12.18	7.77	0.0133	9m
500	0.69	0.15	12.07	7.39	0.0110	16m
600	0.70	0.13	12.56	8.01	0.0096	26m

Table 6. Effect of number of Bayesian optimization iterations. Bold: best.

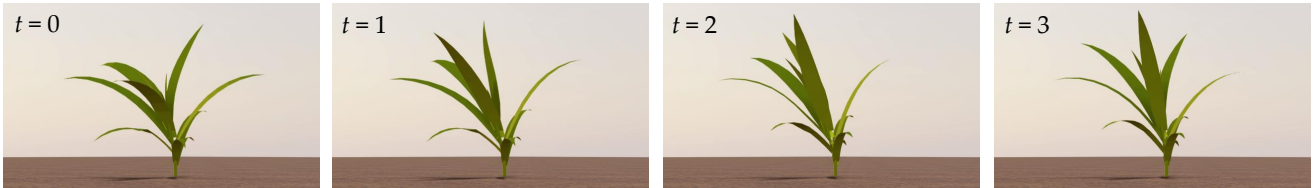


Figure 3. **Wind simulation.** We show that the procedurally generated mesh outputs can be used for physically realistic dynamics simulations. Here, a uniform wind force is applied directly from the right. Video visualizations are included in the supplementary video.