

Combining Close Range UAVs and Gaussian Splatting for High-Speed Tomato Leaf Size Estimation and Counting

Stefan Baar, Yosuke Kobayashi, and Shinya Watanabe
Muroran Institute of Technology
Hokkaido, Japan
Email: {sbaar@, ykobayashi@, sin@muroran-it.ac.jp}

Abstract—Automation of greenhouse monitoring and harvesting tasks using robots and Unmanned Aerial Vehicles (UAVs) has gained attention in recent years. However, automated methods remain slow compared to manual (human) monitoring methods due to high complexity constraints imposed by the environment and the tasks at hand. We present an UAV based approach to monitor tomato plant growth from close ranges. We focus on precise leaf counting, as well as leaf area and curvature measurements. We combine fast UAVs and modern Structure from Motion (SfM) methods, such as Gaussian Splatting, to reveal complex plant morphologies on reasonable time scales and with sufficient measurement precision. We compare our results to traditional (manual) approaches, with only minor divergence.

Index Terms—Leaf area estimation, Neural Radiance Fields, instance segmentation, photogrammetry

I. Introduction

Monitoring plant growth and quality is a crucial factor for optimising the environmental conditions in large scale green-house environments. Unfortunately, modern monitoring methods are either slow or imprecise, due to the high morphological complexity of tomato plants. Further, statistically estimated physical quantities such as the Leaf Area Index (LAI) are subject to high variability and only approximate the actual leaf population [1]. The power law presented in Equation 1 approximates the relationship between Light Absorption (L).

$$L \sim L_0(1 - e^{-A \times LAI}) \quad (1)$$

where L_0 is the intensity at the top of the canopy and A is the normalised extinction coefficient. Note, that uncertainties arise from determining the value of A , but also by unknown variables, such as leaf orientation, leaf curvature etc. Further, the precise light absorption rate (L), depends on area, curvature and extinction (through partial concealment) [2]. A direct measure for LAI is presented by Equation 2:

$$LAI = \frac{\sum_n^N A_{leaf}(n)}{A_{ref}} \approx \frac{\mu(A_{leaf}) \times N}{A_{ref}} \quad (2)$$

where $A_{leaf}(n)$ is the individual area for each leaf indexed by n , N is the number of leafs and A_{ref} is the

projected reference area spanned by a one or more plants. Note, that the computing the mean (μ) requires the leaf area population to be normal distributed [3].

Recently, progress in the field of photogrammetry (2D image to 3D scene reconstruction), has made it possible to resolve and measure three dimensional structures with millimetre precision using Neural Radiance Fields [4] and gaussian splats [5] to realistically approximate scenes and individual objects. Various photogrammetry have been used in previous studies for fruit size estimation for e.g. apples, tomatoes, etc. [6], [7] However, mapping individual leaves remains challenging, because of their flat and curved morphology, which requires high precision.

In addition, recent breakthroughs in deep reinforcement learning have enabled the development of sophisticated autonomous navigation systems for Unmanned Aerial Vehicles (UAVs), allowing them to operate at reasonable velocities and navigate complex environments with increased precision and efficiency [8], [9], [10]. SfM in combination with UAV based data acquisition has been performed in forest sciences and outdoor environments [11]. Still, due to complexity of indoor navigation, many agricultural studies focus their UAV supported research on outdoor environments, using top-down imaging or ranging [12], [13]. Especially, greenhouse environments are challenging to navigate autonomously for the following reasons:

- plants (e.g. tomato) exhibit a sparse morphology (thin branches and leafs)
- high plant density
- sparse and thin metal rods
- free hanging wires and threads
- populated with staff
- obstacles: e.g. pipes, occasional flooding, etc.

Lightweight and small UAVs present a safe and valuable alternatives to ground based robots [14], [12]. Challenges, in navigating UAVs in a greenhouse environment have been elaborated extensively in previous studies [15]. Therefore, this manuscript focuses on the reconstruction process and its limitations. Following, we propose an Unmanned

Areal Vehicles (UAV) based approach, combined with a combination of modern Structure from Motion (SfM) methods including Gaussian Splatting to characterise tomato plant growth and precisely measure the area of each leaf within a reasonable time frame. Supplementary information, such as videos and data are available on Github (<https://github.com/StefanBaar/Combining-FPV-and-Gaussian-Splatting>).

II. Materials and Methods

Frame Accumulation - UAV Trajectory

UAV-video frames

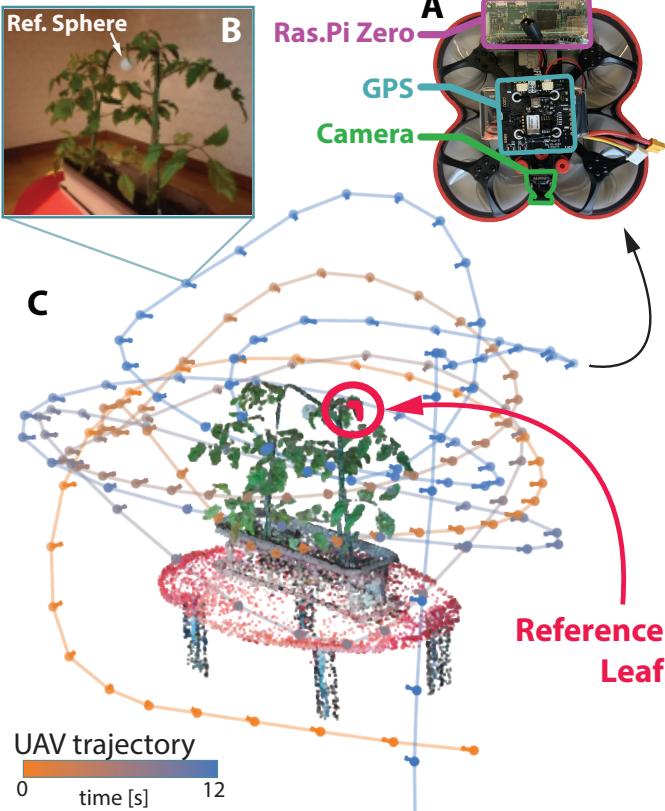


Fig. 1. Frame Accumulation, Model and Trajectory: A) Shows an image of the UAV and its components. B) Example of a randomly chosen sample image taken by the UAV. C) Presents the 3D-RGB point cloud, recovered from the 3D Splats produced by Gaussian Splatting. Further, the UAV trajectory is recovered from the set extrinsic camera parameters and are displayed as arrows with temporal colour coding.

In this study we analyse the morphology of two tomato plants by scanning them using a costumed, high speed Unmanned Areal Vehicle (UAV). The UAV (modified BetaFPV X95V3) and its components, its trajectory as well as the sample plants are presented Figure 1 A,B and C, respectively. The UAV is controlled from a remote server via Wifi, which is provided by a raspberry Pi Zero, which is connected via UART to the Flight-controller. RGB frames are apprehended simultaneously by two cameras with different properties, as presented in Table I.

TABLE I
Measurement Methods

camera	resolution	FPS	bit rate	method	velocity
insta360 Go2	2K	50	50 bps	UAV	12m/s
Caddx vista	720p	30	50 bps	UAV	12m/s
iPhone X	720p	24	50 bps	hh. video	6 m/s
iPhone X	4K	1	-	hh. still	-

Reference data is taken in the form of hand held video and leaf-by-leaf portraits, where each leaf is photographed in front of a regular grid with grid spacing of one centimeter, as later shown in Figure 4 A.

3D Leaf segementation - Methods

A) Frame accumulation

Semi-automated UAV → Spiral Scan

B) 3D - Reconstruction

1 Features

hloc-super glue → Tasks:
feature extraction
feature matching
bundle adjustment

2 SfM model

Gaussian-Splatting → Results:
3D Splats
Reference Sphere

3 Metric Calibration

Reference Sphere → 3D Splats
Camera parameters

C) Leaf Segmentation

1 Point-Cloud + Gaussian features

4 - 3D Leaf instances

2 Probability mapping

3 Connected Components

Fig. 2. UAV based leaf segmentation pipeline: A: The UAV is directed on spiral path around the target plant. B: The Structure from Motion pipeline to produce the camera trajectories and SfM model containing the point-cloud and GS information.C: The individual leaf instances are computed from the SfM model components.

Both plants were recorded from various angles to produce a variety of locally independent images. The frames, apprehended by the UAV are send to a dedicated server, there the data is analysed in post. The Structure from Motion (SfM) pipeline is presented in Figure 2 B. The Hierarchical Localisation Toolbox (hloc) in combination with Gaussian Splatting (GS) is used to compute the intrinsic and extrinsic parameters for each frame and to compute the combined point-cloud and the 3D gaussian features (splats) associated with each point within the point cloud. The retrieved spacial information and its metric is represented in camera (image) based coordinates (associated with pixel coordinates). Therefore, metric calibrations is necessary to retrieve real-world coordinates

in meters. A spherical reference object with unique colour and known dimensions (diameter $d_r = 0.04m$) is therefore placed in the scene (between the two tomato plants), and calibrated against.

Once, the point cloud and its corresponding features are computed and calibrated, we perform heretical segmentation by separating the plant from the environment via clustering the channel representing the colour information of the corresponding HSB (Hue Saturation Brightness) space (in this scenario, we assume no or only limited green components in the surrounding environment and sufficient white balancing accuracy) as presented in Figure 1 C. The gaussian distributions for each point are converted into wave-like probability functions, mapping the leaf surface. The individual leaves are separated from the plant body through watershed segmentation and are subsequently labeled.

III. Results

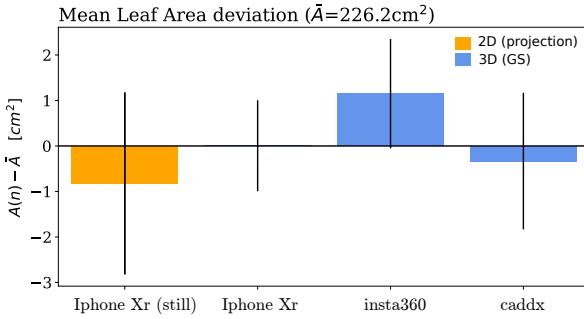


Fig. 3. Leaf Area Precision: The residual to the mean of the Leaf area and its in centimetre for four different measurement (from left to right, projected leaf foto (iPhone Xr), handheld (iPhone Xr) video and UAV based insta360 Go2 and Caddx vista. The corresponding image for the projected leaf area (orange) is presented in Figure 4 A.

Following, we present the surface measurement of a randomly selected leave with index B4, indicated in Figure 1 C. The Leaf Area (LA) of the the leaf, as the mean of the four different measurement methods, was determined to be $LA = (22.6 \pm 1) \text{ cm}^2$. The residual Leaf Area from the mean of the four different methods is presented in Figure 3. The errors are estimated to incorporate resolution and curvature. Gaussian splatting was used for all three runs. The corresponding reference image and its segmentation mask for estimating the 2D projection area is presented in Figure 4 A.

For comparison, the point cloud and splats for Leaf B4 are presented in Figure 4 C and D. The 3D models are based on the same image apprehension and pre-processing methods (images taken with insta360 and feature extraction based on hloc). The narrow leaf morphology is more precisely reproduced by splatfacto than by nerfacto as indicated by the dashed circle in Figure 4 A C and D. The metric calibration has been performed by evaluating the shape of the reference sphere by fitting an

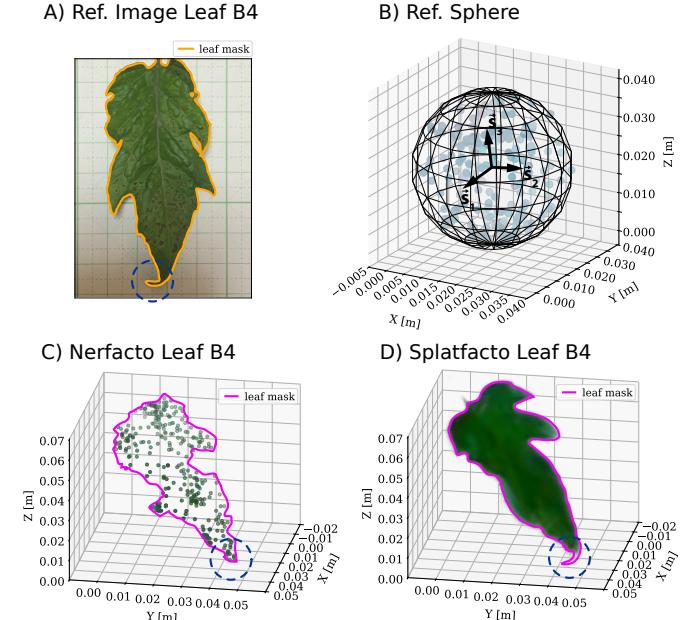


Fig. 4. Photogrammetry Results: A: Reference image of Leaf B4 and its segmentation mask (orange outline). B: Point cloud of reference sphere and fitted ellipsoid for metric calibration and determination of basis vectors. C and D: Calibrated Point cloud (via insta360) and splats for Leaf B4 using nerfacto and splatfacto, respectively. The dashed circle indicates the bend leaf tip.

ellipsoid and computing its independent axis components $\mathbf{S} = (\vec{s}_1, \vec{s}_2, \vec{s}_3)$ to compute the basis vectors of the coordinate system $\mathbf{B} = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ via SVD (Singular Value Decomposition). We assume that the coordinate system exhibits no curvature. The trace of metric g_{ij} representing the length of the basis vectors $b_i = |\vec{b}_i|$ for both nerfacto and splatfacto based approaches is presented in Table III. The metric calibration, as well as the influence of leaf curvature on the estimated area will be elaborated elsewhere.

method	g_{11}	g_{22}	g_{33}
splatfacto	1.56	1.49	1.45
nerfacto	1.58	1.52	1.44

IV. Discussion and Conclusion

The results presented above show, that it is possible to efficiently monitor individual leafs within a reasonable time frame ($\sim 10s$). The measurements do not deviate more than five percent from each other. One has to note, that producing an accurate real-world ground truth is not possible due to leaf curvature and complexity. As expected, the photogrammetry based approaches appear to produce similar results (higher LA). Photogrammetry based approaches can capture the curvature of the individual leaf, while providing a reasonable resolution even at relatively high camera velocities. UAV based measurements were significantly more time efficient than

handheld methods. Photographing each leaf individually while aligning a reference scale is especially time consuming.

In future studies, a complete set of leafs has to be analysed and compared with the methods above. Further, daily measurements in comparison with environmental monitoring (light absorption, temperature, humidity, etc.) would be especially interesting when characterising effects and causes in relation to leaf curvature on a large scale.

In this study, we focused on showing that precise LA measurements are possible even at higher velocities. In future studies, we aim to further elaborate on the implementation, viability and limitations of automated unsupervised, close range UAVs. Recent advancements in combining Gaussian Splatting and SLAM are promising approaches to map complex environments on reasonable timescales, aiding autonomous navigation [16], [17], [18].

Acknowledgments

This research was commissioned and supported by the NICT, JAPAN, and A-STEP of JST (Grant Number: JPMJTM20A1). This work was also supported in part by JSPS KAKENHI (Grant Number No. 20K11968 and No. 23K23728).

References

- [1] S. Baar, Y. Kobayashi, T. Horie, K. Sato, H. Suto, and S. Watanabe, “Non-destructive leaf area index estimation via guided optical imaging for large scale greenhouse environments,” *Computers and Electronics in Agriculture*, vol. 197, p. 106911, 2022.
- [2] J. M. Chen and T. Black, “Defining leaf area index for non-flat leaves,” *Plant, Cell & Environment*, vol. 15, no. 4, pp. 421–429, 1992.
- [3] B. Thomas, *Encyclopedia of applied plant sciences*. Academic Press, 2016.
- [4] Alex Yu and Sara Fridovich-Keil, M. Tancik, Q. Chen, B. Recht, and A. Kanazawa, “Plenoxels: Radiance fields without neural networks,” 2021.
- [5] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, “3d gaussian splatting for real-time radiance field rendering,” *ACM Transactions on Graphics*, vol. 42, no. 4, July 2023. [Online]. Available: <https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/>
- [6] J. Gené Mola, R. Sanz Cortiella, J. R. Rosell Polo, A. Escolà i Agustí, and E. Gregorio López, “Apple size estimation using photogrammetry-derived 3d point clouds,” in *9th Annual Catalan Meeting on Computer Vision*. September 19, 2022, Universitat Autònoma de Barcelona, (<http://acmcv.cat/>), 2022.
- [7] J. Rong, Y. Yang, X. Zheng, S. Wang, T. Yuan, and P. Wang, “Three-dimensional plant pivotal organs photogrammetry on cherry tomatoes using an instance segmentation method and a spatial constraint search strategy,” Available at SSRN 4482155, 2023.
- [8] E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, “Champion-level drone racing using deep reinforcement learning,” *Nature*, vol. 620, no. 7976, pp. 982–987, 2023.
- [9] A. Famili, A. Stavrou, H. Wang, and J.-M. Park, “Pilot: High-precision indoor localization for autonomous drones,” *IEEE Transactions on Vehicular Technology*, vol. 72, no. 5, pp. 6445–6459, 2022.
- [10] P. Foehn, D. Brescianini, E. Kaufmann, T. Cieslewski, M. Gehrig, M. Muglikar, and D. Scaramuzza, “Alphapilot: Autonomous drone racing,” *Autonomous Robots*, vol. 46, no. 1, pp. 307–320, 2022.
- [11] R. Mlambo, I. H. Woodhouse, F. Gerard, and K. Anderson, “Structure from motion (sfm) photogrammetry with drone data: A low cost method for monitoring greenhouse gas emissions from forests in developing countries,” *Forests*, vol. 8, no. 3, p. 68, 2017.
- [12] M. F. Aslan, A. Durdu, K. Sabancı, E. Ropelewska, and S. S. Gültekin, “A comprehensive survey of the recent studies with uav for precision agriculture in open fields and greenhouses,” *Applied Sciences*, vol. 12, no. 3, p. 1047, 2022.
- [13] D. S. Komarchuk, Y. A. Gunchenko, N. A. Pasichnyk, O. A. Opryshko, S. A. Shvorov, and V. Reshetiuk, “Use of drones in industrial greenhouses,” in *2021 IEEE 6th International Conference on Actual Problems of Unmanned Aerial Vehicles Development (APUAVD)*. IEEE, 2021, pp. 184–187.
- [14] J. J. Roldán, G. Joosse, D. Sanz, J. d. Cerro, and A. Barrientos, “Mini-uav based sensory system for measuring environmental variables in greenhouses,” *Sensors*, vol. 15, no. 2, pp. 3334–3350, 2015.
- [15] Y. Khosiawan and I. Nielsen, “A system of uav application in indoor environment,” *Production & Manufacturing Research*, vol. 4, no. 1, pp. 2–22, 2016.
- [16] H. Matsuki, R. Murai, P. H. Kelly, and A. J. Davison, “Gaussian splatting slam,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 18039–18048.
- [17] T. Wu, Y.-J. Yuan, L.-X. Zhang, J. Yang, Y.-P. Cao, L.-Q. Yan, and L. Gao, “Recent advances in 3d gaussian splatting,” *Computational Visual Media*, pp. 1–30, 2024.
- [18] M. Li, J. Huang, L. Sun, A. X. Tian, T. Deng, and H. Wang, “Ngm-slam: Gaussian splatting slam with radiance field submap,” *arXiv preprint arXiv:2405.05702*, 2024.