

3D Point Cloud Reconstruction from UAV Images through Incremental Structure from Motion (SfM) for Plant Phenotyping

Andreas Ming

Institute of Electrical Engineering

Lucerne University of Applied Sciences and Arts

Horw, Switzerland

andreas.ming@stud.hslu.ch

Richard Green

Department of Computer Science and Software Engineering

University of Canterbury

Christchurch, New Zealand

richard.green@canterbury.ac.nz

Abstract—The paper proposed a method for constructing a 3D point cloud of plants from overhead images taken by a unmanned aerial vehicle (UAV). The structure-from motion (SfM) algorithm constructs a point cloud through SIFT feature extraction, matching, and geometric verifications in correspondence search, and next best view selection, triangulation, local bundle adjustment, retriangulation, observation merging, and global bundle adjustment in a incremental SfM approach. Comparing the algorithm to a commercially available solution, shows that the commercial solution generates 150% more voxels. Furthermore, by reconstructing the point cloud with the high resolution images and finally applying multi-view-stereo (MVS) for point cloud densification, 950% and 9000% more voxels are generated respectively. Future research into incorporation of geolocation data, MVS, and deep learning approaches like neural radiance fields (NeRF) for dense point cloud reconstruction and real-time processing are recommended.

Index Terms—3D reconstruction, phenotyping, structure from motion, incremental structure from motion, point cloud processing

I. INTRODUCTION

To overcome the up to 56% increase in food demand by 2050, serious measures have to be applied, as current crop production is insufficient to lower the risk of hunger. Furthermore, an overall more stable crop production, able to deal with pests, pathogens, heat waves, and other climate change induced weather extremes, must be achieved [1]. A key point of breeding new resilient crops are double haploidy, genomics-assisted breeding, breeding data management, and high-throughput and precise phenotyping [2], [3]. Acquisition of large-scale phenotype data in traditional research is mainly based on manual measurements and thus become a major bottleneck for big-data [4]. Unmanned areal vehicles (UAV) - drones are used in agriculture for pesticide spraying and crop monitoring. Equipped with various sensors and cameras of different spectra, drones are a useful tool for crop monitoring, either for pesticide spraying, fertiliser estimation, or crop monitoring [5].

In plant phenotyping specifically, there is a need in the accurate capture of individual plants, or the recording of individual characteristics depending on the phonological stage

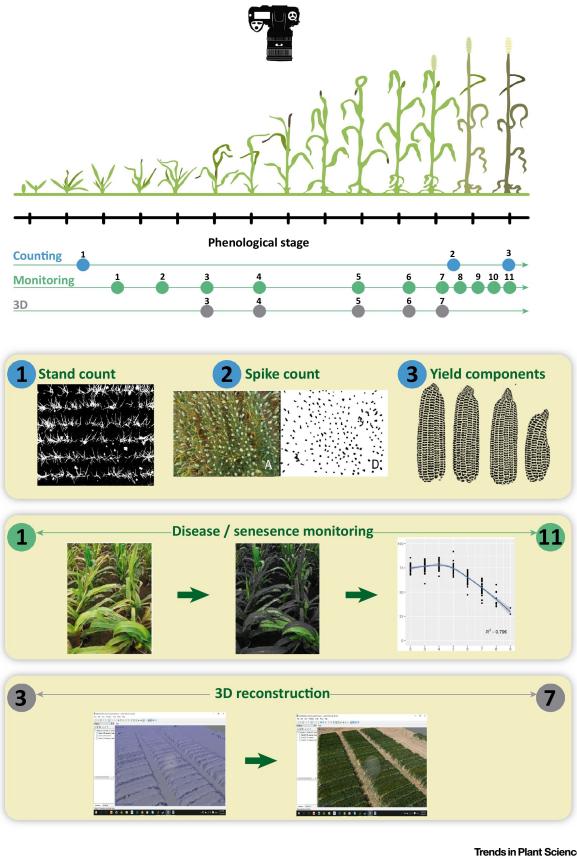


Figure 1: Examples of potential applications of field phenotyping with RGB-images. Different phenotyping needs in respect to the corresponding phenological stage. [3]

of a plant (see Figure 1) [3]. Traditionally ground based systems are used for this task, as the environmental variables can be controlled much easier [4]. If drones could be used instead, this would increase effectiveness and reduce the direct environmental impact, e.g. through ground pressure. Drones seem to be especially useful for this task, based on

the low price and ease-of-use [6]. If needed, drones can be combined with ground robots to capture more data from angles obstructed for the drone (e.g. under the canopy) [7]. 3D reconstruction is an important part of phenotyping, and if accurate enough, could work as a new base for algorithms to work out other plant details like correlations between biomass and plant constituents [8]. Furthermore, with the acquisition of precise 3D models of plants over time, functional structural plant models (FSPMs) can be constructed and conclusions drawn [9].

3D reconstruction through Structure-from-Motion (SfM) is usually featured as lightweight and cheap to implement, as just one or more standard monochrome or colour cameras need to be used to capture the desired images. Furthermore, a complete reconstruction can be done on unstructured images, without the use of references e.g. global navigation satellite system (GNSS). With increasing dataset size a higher resolution can be achieved and SfM algorithms are typically able to provide mm-accuracy [10].

II. BACKGROUND

This section describes the state-of-the-art implementation of SfM for 3D reconstruction and is heavily based on J. Schönebergers [11], [12] work, and the implementation in COLMAPS [13].

SfM is constructing a scene through projections of images taken from different viewpoints. To incorporate a series of images, a sequential processing pipeline with a iterative reconstruction component forms the incremental SfM (see Figure 2). [11]

After completing the iterative SfM, we acquired a set of poses \mathcal{P} and reconstructed points \mathcal{X} . These points X_k represent the triangulation of matched key points between image pairs.

A. Correspondence Search

The first stage identifies projections of the same points in overlapping input images (key points) and then outputs a set of geometrically verified image pairs as well as a graph of the image projections for each point. [11]

1) *Feature Extraction:* For each input image I_i , sets of local features $\mathcal{F}_i = \{(x_j, f_j) | j = 1 \dots N_{F_i}\}$ represented as a appearance descriptor f_j at location $x_j \in \mathbb{R}^2$ are processed. To be uniquely recognisable in different images, the features should be invariant under radiometric and geometric changes [14]. Several feature detectors based of Scale Invariant Feature Transform (SIFT) [15], [16] appear to be especially performing in terms of robustness. [11]

2) *Matching:* To work out scene overlaps between images for later triangulation, the features \mathcal{F}_j of all image pairs are tested. It searches for corresponding features by finding the most similar feature in image I_a for each feature in image I_b through the similarity metric f_j . The output of this are potential overlapping image pairs $\mathcal{C} = \{I_a, I_b\}$ and their respective feature correspondences \mathcal{M}_{ab} .[11]

3) *Geometric Verification:* Since matching is only based on appearance and not location, it can not be guaranteed that the potential image pairs \mathcal{C} are actually overlapping. To verify this, SfM estimates a transformation matrix that maps feature points between images. As a result, a homography \mathbf{H} for planar scenes, a essential matrix \mathbf{E} , or a fundamental matrix \mathbf{F} for moving cameras is estimated. Because of heavy outlier-contaminated matching correspondences, robust estimators like RANSAC [17] are required. A potential image pair \mathcal{C} becomes a verified image pair $\bar{\mathcal{C}}$ if enough inliers can be described as a fundamental matrix \mathbf{F} . Besides the verified image pair $\bar{\mathcal{C}}$, their respective inlier correspondence $\bar{\mathcal{M}}_{ab}$, and a description of their geometric relation \mathcal{G}_{ab} are outputted by this stage. To compute the geometric relation \mathcal{G}_{ab} , GRIC [18] is used as a decision criterion. The output of GRIC is a scene graph with images as nodes and verified pairs of images as edges. [11]

B. Incremental Reconstruction

The reconstruction stage computes the pose estimates \mathcal{P} and the reconstructed scene as a set of points \mathcal{X} with $X_k \in \mathbb{R}^3$. For that, the scene graph from the last scene is used and one image after the other is added to the model. In a iterative process, all images are incorporated in the reconstruction of the scene. [11]

1) *Next Best View Selection:* To work out what images to take next, and especially with what pair of images to start with, a robust view selection has to be applied. This is done through a scoring approach which is applied to each image, more precisely to the matching key points of image pairs. More visible points and a more uniform distribution of these points results in a higher score \mathcal{S} (see Figure 3) [19]. Because of more points to triangulate, images with a high score are, especially in early iterations, better suited for a robust reconstruction.

Each image is split into a grid of cells which are either *empty* if no point is present, or *full* if one or more points are present. This gives a higher score to equally distributed points. This scheme might not capture distribution well if few partitions are present, as every partition might be activated by a point (e.g. most bottom left configuration in Figure 3). Therefore a multi resolution approach with $l = 1 \dots L$ levels of partitioning with incrementing resolution $K_l = 2^l$ is used. The score is then accumulated over all levels. [11]

2) *Image Registration:* Through the Next Best View Selection approach, we get the image which sees the most triangulated points of the current model. To estimate the pose \mathbf{P}_c of the newly registered image, we further maximise the score and thus the viewable points. This is commonly known as the Perspective-n-Pose (PnP) [17] problem. The estimated pose \mathbf{P}_c is then added to the set of poses \mathcal{P} . [11]

3) *Triangulation:* With the registered image and the scene graph of the newly present image pairs, further points X_k can be triangulated and added to the scene set \mathcal{X} . This point X_k is valid as soon as one or more images covering the scene from a different viewpoint are registered. For multi-view triangulation, and especially handling outlier contamination,

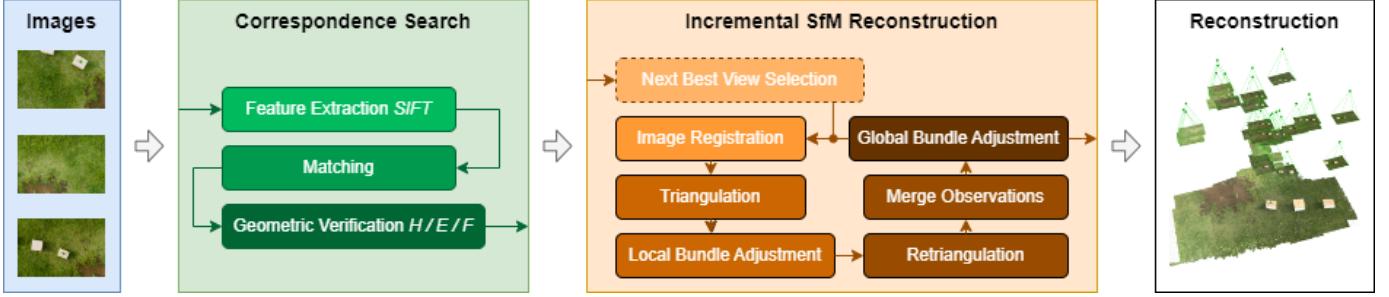


Figure 2: Incremental SfM pipeline (adapted from [11])

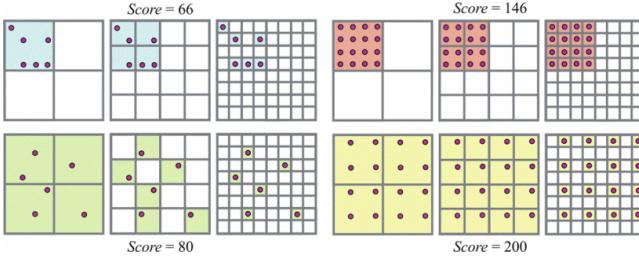


Figure 3: Scores for different number of points (left and right) with different distributions (top and bottom) in the image for $L = 3$ [11]

RANSAC is used. A feature track \mathcal{T} consisting of normalised image observations $\bar{x}_n \in \mathbb{R}^2$ and a corresponding pose \mathbf{P}_n . The goal is then to maximise the support of measurements conforming with a triangulation

$$X_{ab} \sim \tau(\bar{x}_a, \bar{x}_b, \mathbf{P}_a, \mathbf{P}_b) \text{ with } a \neq b$$

where τ is the DLT [20] triangulation method and X_{ab} the triangulated point. A well-conditioned model satisfies a sufficient triangulation angle α and positive depths $d_{a/b}$ with respect to the image poses $\mathbf{P}_{a/b}$. [11]

4) Local Bundle Adjustment: As image registration and triangulation are separate processes, their outputs are highly correlated. Still there are uncertainties propagating from camera poses to triangulation and similarly the reverse is true. To further suppress such induced errors, we perform bundle adjustment. Because incremental SfM affects models only locally, we only perform a local BA on a set of the most connected images. BA is a iterative process, with the goal to reduce a loss-function, in our case a Cauchy function ρ_j . The optimisation is performed through a Ceres Solver [21]. [11]

5) Filtering: After a local BA, observations which do not conform with the model are filtered out. These are characterised through a large reprojection error [22]. [11]

6) Retriangulation: To account for drift effects induced through the local BA, a retriangulation is performed. This helps in improving completeness of reconstruction by continuing tracks of points, that previously might have failed to triangulate. [11]

7) Global Bundle Adjustment: After the model has grown by a certain percentage, a global BA is performed, resulting in

Table I: Drone Specifications [23]

Component	Specification
Drone Body	GD-X8 V2 (1000 mm)
Motors	T-motor P60 $\times 4$
Flight Controller	CUAV X7
GNSS	RTK-GNSS: ZED-F9P module
Gimbal	GREMSY S1V3
Camera	Canon PowerShot G7 X Mark 2
Image Resolution	5472 x 3638
Micro-computer	JETSON Nano

a amortised linear run-time of SfM. Without this refinement, SfM tend to drift into a non-recoverable state. Similarly, to the local BA, a Cauchy function ρ_j is optimised with the Ceres Solver [21]. [11]

III. PROPOSED METHOD

This section describes the proposed method for 3D reconstruction of plants through a UAV for phenotyping. First, the imaging platform, which consists of a drone equipped with a high resolution camera for imaging the plants; second, the computer vision system, which builds a 3D point cloud of the ground and its plants.

A. Drone Platform

As a basis for SfM reconstruction a series of pictures acquired by a Drone with a built-in synchronised camera and RTK-GNSS System is used, the dataset is publicly available by ref. [23]. The camera and RTK-GNSS are synchronised by a microcomputer on the drone (see Figure 4), while the images and their positions are monitored from the ground base. The RTK-GNSS data is not used for 3D reconstruction in this paper. For evaluation of the algorithm, a fake plant, and three boxes of known size are placed on the ground, one of the boxes with a plant on it. [23]

Images are captured in a resolution of 5472 x 3638 pixels from a height of 10m and 20m in a overflight configuration. For the further drone specifications, refer to Table I. [23]

B. Computer Vision System

To retrieve a 3D point cloud for phenotyping of plants, we propose the use of a incremental SfM approach as described in Section II. Due to computational limits (see Table II for



Figure 4: Drone used to acquire the data (left) and the drone in the testing environment (right) [23]

Table II: System Specifications

Component	Specification
OS	Linux Mint 21.1 Cinnamon
CPU	12 th Gen Intel Core™ i7-12700x12
IDE	Visual Studio Code
Interpreter	Python 3.10.6
CUDA Support	No
Image Resolution	2472 x 1648
COLMAP Version	3.9.1

system specifications), images are restricted to a resolution of 2472 x 1648. A comparison with a commercially available 3D reconstruction tool PIX4Dmapper, utilising the full resolution images, is done. Furthermore, the reconstruction with PIX4Dmapper incorporates multi-view stereo (MVS) to further condense the point cloud.

IV. RESULTS

The reconstruction algorithm is run on a system defined in Table II. The algorithm is run on the downsampled dataset acquired with the drone platform described in Section III-A. All images are collected in one run and then given to the algorithm as an unordered dataset. The algorithm then worked out a 3D point cloud. For comparison, a 3D reconstruction with the commercially available PIX4Dmapper is done on the downsampled and original datasets. Additionally, a densification of the point cloud through MVS is done in PIX4Dmapper on the high resolution data set.

If we compare the number of points in each point cloud in Table III, it can be observed, that simple incremental SfM produces the least amount of points, whereas PIX4Dmapper works out 150% more points on the same dataset. With a higher resolution the increase grows to 950% and 9000% if MVS is applied.

To work out the plant heights, a peak finding algorithm is used. For this, the point cloud is filtered for outliers, and a plane is fitted to represent the ground. To find the plants, the point furthest away from the plane is taken, and its surrounding points deleted for the next iteration. This is repeated until the

Table III: Tool Performance

Tool	Reconstructed Points
COLMAP (SfM, Low-Res)	85749
PIX4Dmapper (SfM, Low-Res)	217153
PIX4Dmapper (SfM, High-Res)	901223
PIX4Dmapper (SfM+MVS, High-Res)	7859904

height falls under a certain threshold. As seen in Figure 6, bounding boxes are set around each found peak, to visualise the separate plants in the point cloud. It has been identified, that this algorithm can not distinguish between plants, ground, nor different objects, as no segmentation is performed. The algorithm is also incapable of adapting to different plant sizes, as the bounding boxes are a fixed size. Furthermore, fitting the ground to the plane leads to bumps in the ground to be recognised as plants. The plane approximation only works for flat grounds or flat patches in bigger point clouds.

V. CONCLUSION

The paper proposed a method for constructing a 3D point cloud of plants from overhead images taken by a UAV. The incremental SfM algorithm, processing the unordered downsampled dataset, successfully generates a 3D point cloud. A comparison with the commercially available PIX4Dmapper revealed, that the commercial solution generates 150% more voxels. Furthermore, by reconstructing the point cloud with the high resolution images, and finally applying multi-view stereo (MVS) for point cloud densification, 950% and 9000% more voxels are generated respectively.

For plant determination, a peak finding algorithm is applied to the point cloud. After filtering the point cloud for outliers, a plane is fitted, and plants in respect to the plane searched for. A fixed bounding box is created for each plant. Because of a missing segmentation, the algorithm can not distinguish between plants and surrounding points. Additionally, the fixed-size bounding boxes fail to adapt to plant dimensions, further inducing errors.

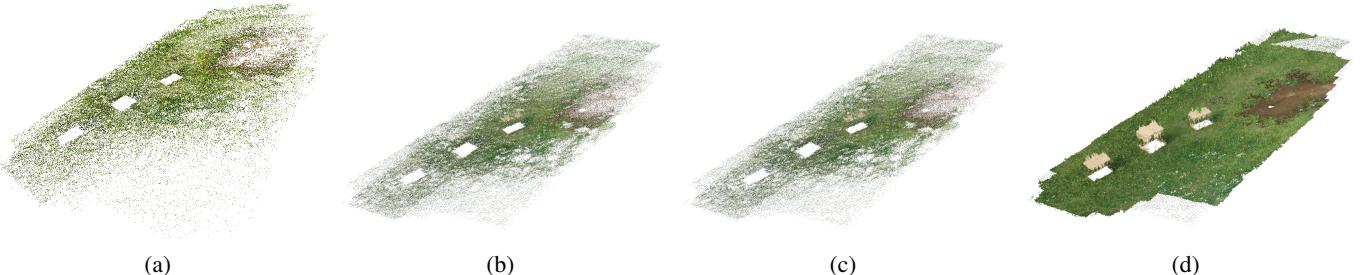


Figure 5: The reconstructed point clouds. In (a) through incremental SfM described in Section II, visualised in Open3D, (b) and (c) are computed with the commercial PIX4Dmapper and shows all triangulated points from matched image features (ties) on a low- and high-resolution dataset, respectively. In (d) MVS is applied in PIX4Dmapper for a denser point cloud.

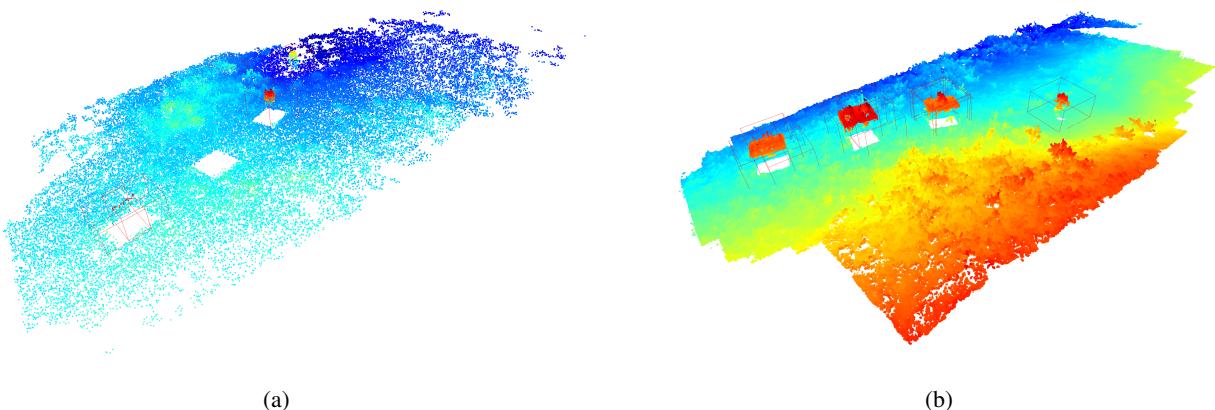


Figure 6: The found plants represented as red bounding boxes in the point cloud. The colour of each voxel is mapped on the z-coordinate of the respective voxel. (a) is the point cloud acquired through incremental SfM (Figure 5a). (b) is the point cloud acquired through MVS in PIX4Dmapper (Figure 5d)

A. Future Research

While this paper has provided valuable insights and advancements in the field of 3D reconstruction through incremental SfM, there are several points of improvements to enhance the 3D reconstruction and the plant phenotyping itself. The comparison in Section IV showed, that the commercial solution computes denser point clouds than the open-source implementation of the incremental SfM. To improve this, multi-view stereo (MVS) [24] and its refined variation pixelwise view selection for multi-view stereo [12] could be applied. However it must be noted, that a computation of a denser point cloud has to be viewed critically, as the computational effort rises exponentially with more voxels and thus a trade-off between the accuracy and computational effort should be analysed [10], [25].

Another way to improve point clouds while still keeping the required computing power low, is to incorporate a RTK-GNSS signal, which provides spatial information about a taken image. Taking the location data into account, a stereo matching and reconstruction can be done through calculating the baseline of two images from the location data, after each acquired image. With this approach a real-time reconstruction

can be achieved [23], [26]. To work out a more accurate location of each image, and thus a higher accuracy baseline, ground control points (GPC) either as physical markers in the field or through feature matching should be investigated. Through the alignment of each new image, the accuracy of the position, thereby of the baseline and the stereo matching can be improved [27].

Recently 3D reconstruction through deep learning approaches like neural radiance fields (NeRF) [28], which is already used in the agricultural domain in the form of Zero NeRF [29], or Gaussian splatting [30] have gained considerable attention for its ability to capture detailed spatial information in submillimeter accuracy from 2D images.

After retrieving a sufficient 3D point cloud of a plant, different approaches for phenotyping of plants should be looked into. For growth analysis, the measurement of plant heights can be done through disparity maps [23]. In most cases, more information about plants, such as width, convex hull, projected leaf area, leaf density, number of leaves, and the respective leaf lengths can be retrieved through deep learning segmentation. Through repeated acquisition over time, a time variable can be induced and thus a functional structural plant model is retrieved [9], [10], [25].

REFERENCES

- [1] Michiel van Dijk, Tom Morley, Marie Luise Rau, et al. “A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050”. In: *Nature Food* 2.7 (July 2021). Publisher: Nature Publishing Group, pp. 494–501. DOI: 10.1038/s43016-021-00322-9. URL: <https://www.nature.com/articles/s43016-021-00322-9>.
- [2] Boddupalli M. Prasanna, Jill E. Cairns, P. H. Zaidi, et al. “Beat the stress: breeding for climate resilience in maize for the tropical rainfed environments”. In: *Theoretical and Applied Genetics* 134.6 (June 1, 2021), pp. 1729–1752. DOI: 10.1007/s00122-021-03773-7. URL: <https://doi.org/10.1007/s00122-021-03773-7>.
- [3] José Luis Araus, Shawn C. Kefauver, Mainassara Zaman-Allah, et al. “Translating High-Throughput Phenotyping into Genetic Gain”. In: *Trends in Plant Science* 23.5 (May 1, 2018), pp. 451–466. DOI: 10.1016/j.tplants.2018.02.001. URL: <https://www.sciencedirect.com/science/article/pii/S1360138518300207>.
- [4] Wanneng Yang, Hui Feng, Xuehai Zhang, et al. “Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives”. In: *Molecular Plant* 13.2 (Feb. 3, 2020). PMID: 31981735, pp. 187–214. DOI: 10.1016/j.molp.2020.01.008.
- [5] Umamaheswara Mogili and B. Deepak. “Review on Application of Drone Systems in Precision Agriculture”. In: *Procedia Computer Science* 133 (Jan. 1, 2018), pp. 502–509. DOI: 10.1016/j.procs.2018.07.063.
- [6] Wei Guo, Matthew E. Carroll, Arti Singh, et al. “UAS-Based Plant Phenotyping for Research and Breeding Applications”. In: *Plant Phenomics* 2021 (June 10, 2021). Publisher: American Association for the Advancement of Science. DOI: 10.34133/2021/9840192. URL: <https://spj.science.org/doi/10.34133/2021/9840192>.
- [7] Xiang Gao, Lingjie Zhu, Hainan Cui, et al. “Complete and Accurate Indoor Scene Capturing and Reconstruction Using a Drone and a Robot”. In: *IEEE Sensors Journal* 21.10 (May 2021). Conference Name: IEEE Sensors Journal, pp. 11858–11869. DOI: 10.1109/JSEN.2020.3024702. URL: <https://ieeexplore.ieee.org/document/9201133>.
- [8] Shunfu Xiao, Honghong Chai, Ke Shao, et al. “Image-Based Dynamic Quantification of Aboveground Structure of Sugar Beet in Field”. In: *Remote Sensing* 12.2 (Jan. 2020). Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 269. DOI: 10.3390/rs12020269. URL: <https://www.mdpi.com/2072-4292/12/2/269>.
- [9] Soualiou Soualiou, Zhiwei Wang, Weiwei Sun, et al. “Functional–Structural Plant Models Mission in Advancing Crop Science: Opportunities and Prospects”. In: *Frontiers in Plant Science* 12 (Dec. 23, 2021). Publisher: Frontiers. DOI: 10.3389/fpls.2021.747142. URL: <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.747142/full>.
- [10] Stefan Paulus. “Measuring crops in 3D: using geometry for plant phenotyping”. In: *Plant Methods* 15.1 (Sept. 3, 2019), p. 103. DOI: 10.1186/s13007-019-0490-0. URL: <https://doi.org/10.1186/s13007-019-0490-0>.
- [11] Johannes Schönberger and Jan-Michael Frahm. *Structure-from-Motion Revisited*. DOI: 10.1109/CVPR.2016.445. June 1, 2016.
- [12] Johannes Schönberger, Enliang Zheng, Marc Pollefeys, et al. *Pixelwise View Selection for Unstructured Multi-View Stereo*. Vol. 9907. DOI: 10.1007/978-3-319-46487-9_31. Oct. 1, 2016.
- [13] Johannes Schönberger. *colmap/colmap*. original-date: 2014-08-16T00:55:31Z. colmap, Mar. 26, 2024. URL: <https://github.com/colmap/colmap>.
- [14] Chris McGlone. “Manual of photogrammetry”. In: (Jan. 1, 2004). URL: https://www.academia.edu/69009469/Manual_of_photogrammetry.
- [15] D.G. Lowe. “Proceedings of the Seventh IEEE International Conference on Computer Vision”. In: Kerkyra, Greece: IEEE, 1999, 1150–1157 vol.2. DOI: 10.1109/ICCV.1999.790410. URL: <http://ieeexplore.ieee.org/document/790410/>.
- [16] Tinne Tuytelaars and Krystian Mikolajczyk. “Local Invariant Feature Detectors: A Survey”. In: *Foundations and Trends® in Computer Graphics and Vision* 3.3 (June 14, 2008). Publisher: Now Publishers, Inc., pp. 177–280. DOI: 10.1561/0600000017. URL: <https://www.nowpublishers.com/article/Details/CGV-017>.
- [17] Martin A. Fischler and Robert C. Bolles. “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”. In: *Communications of the ACM* 24.6 (June 1, 1981), 381–395. DOI: 10.1145/358669.358692. URL: <https://dl.acm.org/doi/10.1145/358669.358692>.
- [18] P.H.S. Torr. “An assessment of information criteria for motion model selection”. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1997). Conference Name: IEEE Computer Society Conference on Computer Vision and Pattern Recognition ISBN: 9780818678226 Place: San Juan, Puerto Rico Publisher: IEEE Comput. Soc, pp. 47–52. DOI: 10.1109/CVPR.1997.609296. URL: <http://ieeexplore.ieee.org/document/609296/>.
- [19] Arnold Irschara, Christopher Zach, Jan-Michael Frahm, et al. *From structure-from-motion point clouds to fast location recognition*. Journal Abbreviation: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009 Pages: 2606 Publication Title: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009 DOI: 10.1109/CVPR.2009.5206587. June 1, 2009.

- [20] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. 2nd ed. DOI: 10.1017/CBO9780511811685. Cambridge: Cambridge University Press, 2004. URL: <https://www.cambridge.org/core/books/multiple-view-geometry-in-computer-vision/0B6F289C78B2B23F596CAA76D3D43F7A>.
- [21] Sameer Agarwal, Keir Mierle, and The Ceres Solver Team. *Ceres Solver*. Version 2.2. Oct. 2023. URL: <https://github.com/ceres-solver/ceres-solver>.
- [22] Changchang Wu. “2013 International Conference on 3D Vision - 3DV 2013”. In: ISSN: 1550-6185. June 2013, pp. 127–134. DOI: 10.1109/3DV.2013.25. URL: <https://ieeexplore.ieee.org/document/6599068>.
- [23] Yuta Matsuura, Zhang Heming, Kousuke Nakao, et al. “High-precision plant height measurement by drone with RTK-GNSS and single camera for real-time processing”. In: *Scientific Reports* 13.1 (Apr. 18, 2023). Publisher: Nature Publishing Group, p. 6329. DOI: 10.1038/s41598-023-32167-6. URL: <https://www.nature.com/articles/s41598-023-32167-6>.
- [24] Christian Bailer, Manuel Finckh, and Hendrik Lensch. *Scale Robust Multi View Stereo*. Vol. 7574. Pages: 411 DOI: 10.1007/978-3-642-33712-3_29. Oct. 7, 2012.
- [25] Yuchao Li, Jingyan Liu, Bo Zhang, et al. “Three-dimensional reconstruction and phenotype measurement of maize seedlings based on multi-view image sequences”. In: *Frontiers in Plant Science* 13 (Sept. 2, 2022). PMID: 36119622 PMCID: PMC9481285, p. 974339. DOI: 10.3389/fpls.2022.974339. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9481285/>.
- [26] Seyyed Meghdad Hasheminasab, Tian Zhou, and Ayman Habib. “GNSS/INS-Assisted Structure from Motion Strategies for UAV-Based Imagery over Mechanized Agricultural Fields”. In: *Remote Sensing* 12.3 (Jan. 2020). Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, p. 351. DOI: 10.3390/rs12030351. URL: <https://www.mdpi.com/2072-4292/12/3/351>.
- [27] Aijing Feng, Chin Nee Vong, Jing Zhou, et al. “Developing an image processing pipeline to improve the position accuracy of single UAV images”. In: *Computers and Electronics in Agriculture* 206 (Mar. 1, 2023), p. 107650. DOI: 10.1016/j.compag.2023.107650. URL: <https://www.sciencedirect.com/science/article/pii/S0168169923000388>.
- [28] Kyle Gao, Yina Gao, Hongjie He, et al. “NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review”. In: () .
- [29] Casey Peat, Oliver Batchelor, Richard Green, et al. *Zero NeRF: Registration with Zero Overlap*. DOI: 10.48550/arXiv.2211.12544. Nov. 22, 2022.
- [30] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, et al. “3D Gaussian Splatting for Real-Time Radiance Field Rendering”. In: () . DOI: 10.48550/arXiv.2308.04079.