Towards Digital Twin of Crops for Growth Modelling using Virtual Reality

Karanvir Singh
karanvir.21csz0016@iitrpr.ac.in
Indian Institute of Technology Ropar
India

Mukesh Saini
mukesh@iitrpr.ac.in
Indian Institute of Technology Ropar
India

ABSTRACT

A major problem that a farmer faces, while adopting a new crop variety in the farm; is the uncertainty associated with its growth. Farmers working on real farms are not aware of the growth models, even for the existing crops. Hence, there is a need for more accessible and intuitive models. This work is a step towards the realization of another promising model, which is the digital twin of a crop. A primary requirement of the digital twin is the digital representation of the crop itself. Extending that notion, the work discusses the development of 3D assets of crops and their temporal alignment. It also describes the methodology involved in the development of a VR framework, which stores the ideal growth of a crop. This framework could be useful to farmers who want to confirm the growth of their crops. Furthermore, it also proposes a quantitative metric to evaluate the VR framework. The consistency of this proposed metric is further backed by a user study which is based on a qualitative method.

CCS CONCEPTS

• **Applied computing** → *Agriculture*; • **Human-centered computing** → *Visualization design and evaluation methods*; • **Computing methodologies** → *Mesh models.*

KEYWORDS

Virtual Reality, Photogrammetry, Agriculture

ACM Reference Format:
Karanvir Singh and Mukesh Saini. 2023. Towards Digital Twin of Crops for Growth Modelling using Virtual Reality. In *ACM Multimedia Asia 2023 (MMAsia '23), December 06–08, 2023, Tainan, Taiwan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3595916.3626368

1 INTRODUCTION

The new crop varieties that are invented in agricultural laboratories are generally not adopted directly by the farmers [1]. One of the reasons for this non-adoption could be the crop growth trajectory of the new crop, which could be different from any crop they might have grown in the past. There would be uncertainties associated with these new varieties. As a result, the growth models for these crops would not be readily available.

Furthermore, even the presence of growth models for old crops like cauliflower [11], peanuts [7], and sweet pepper [8]; are also of no direct use to the farmers because the farmers work in fields and generally don't understand those mathematical models. Hence, they need something more visual and perceptive than just those models. This situation calls for the need for virtual reality (VR) in the agriculture domain. As a prequel to that, there is a need for a digital representation of a plant. Yu et al. [15] propose a virtual plant and its related applications in VR, but that work lacks implementation in the real field. Extending the same to crops, the digital representation of crops is the first step in the fruition of the digital twin of crops. This representation could be composed of internal characteristics and external appearance.

Our research focuses on the external appearance of a crop. We propose a VR method to enlighten the farmers about how the ideal crop should look on a daily basis if the appropriate care is provided to the crop. The idea is to acquire the 3D assets of the new crop daily in lab conditions via photogrammetry. It is followed by the alignment of those assets and the model integration in Unity for a thorough inspection by human eyes in VR. VR has a strong reputation for making demonstration-related applications very effective, owing to its concept of immersion and presence [2]. Furthermore, specific visual effects related to various environmental stimuli can also be induced on these 3D assets. Some of the possible effects could be water stress, lack of sunlight, and pest infestation.

In our current work, we developed a prototype VR framework through which one can navigate the timestamped assets of the crop to visualize its growth. We also proposed a quantitative metric to describe the realism of our framework. The metric is based on the alignment error between the pairs of iterative 3D assets involved in our case. We also conducted a user study to validate this metric. The crop selected for our framework was cauliflower. Later, it could be extended to other crops by following the same steps in Section 3 and consulting the agricultural expert for the growth inputs needed for the desired crop.

2 RELATED WORKS

To the best of our knowledge, there is not enough literature available on the 3D assets of real crops. However, the research is already progressing at the image level. A common theme in agriculture is the use of Generative Adversarial Networks (GANs) to create realistic-looking diseased plant images. Then these synthetic images are used for the augmentation of existing datasets. Afterward, various tasks like disease classification, plant labeling, and phenotyping. are performed using the augmented data. To name a few,

MMAsia '23, December 06–08, 2023, Tainan, Taiwan

ACM ISBN 979-8-4007-0205-1/23/12…$15.00
https://doi.org/10.1145/3595916.3626368

work done by Chen et al. [4] focuses on the disease classification for apples. Nazki et al. [12] have done the disease classification for the case of tomatoes. Divyanth et al. [6] discuss the identification of maize from using conditional GANs (cGANs).

Another active sub-field is 'Image to Image Translation' [13], where one can induce various simulation effects by converting images from one domain to another. Xu et al. [14] introduced instance segmentation to transfer the disease effects on various tomatoes in a single image. Then, there is LeafGAN [3] which added a segmentation module to the cycleGAN to transfer the disease effect solely to the plant, keeping the background intact. Furthermore, Cui et al. [5] have performed even the disease induction across the cross-species of the plants also; by which, we mean inducing a disease from a specific crop to some other crop.

Contrary to the earlier works, our work focuses on the modelling and alignment of 3D temporal assets of the crops. The temporal dimension is a key aspect to consider for the progress to be made toward the digital twin of crops.

3 METHODOLOGY

The framework developed in our work had four aspects namely the *Cultivation*, *Photogrammetry*, *Alignment of 3D crop assets & Model Integration*. We cultivated the cauliflower crop in our laboratory itself. Along with that, we also carried out the photogrammetry process. These two steps involved giving specific daily input to the crop and capturing photographs from different camera poses. After gathering all the 3D assets, the next step was to align all of them so as to visualize the crop growth. Finally, we imported the aligned assets in the Unity and created scripts to navigate the assets in VR. The corresponding flowchart is shown in Figure 1.

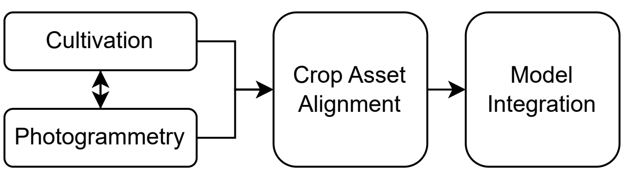


Figure 1: The Overall Process

3.1 Cultivation

We brought a cauliflower sapling from a nearby field. Then we transplanted it into a pot with an appropriate amount of cocopeat soil as prescribed by the gardeners. Considering the climatic conditions, we provided it with 100 watts of LED exposure for about 8 hours daily. We watered the crop on alternative days with the apt amount to keep the soil wet at an optimum level and avoid leaching. The grown plant can be seen in the Figure 2.

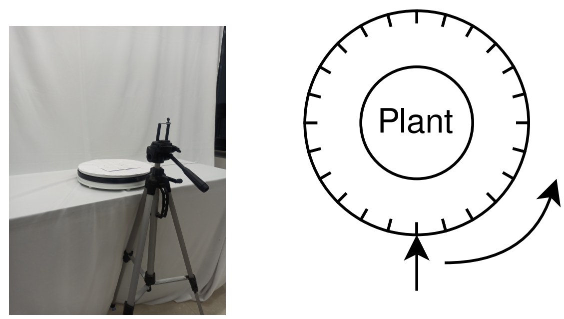
3.2 Photogrammetry

Our photogrammetry setup had a tripod, a turntable, and a white backdrop cloth. We used this setup to capture photographs of the crop from different poses. To reduce measurement errors, we cut out a circle shape from a white chart to cover the turntable. Then, we marked 24 points on the boundary at equal angles. We also traced

(a) Day 10 **(b) Day 11** **(c) Day 12**

Figure 2: Cauliflower Growth from Day 10 to Day 12 after the transplantation to the pot

the boundary of the pot at the center of the cutout, to always keep it in the same position. It can be seen in Figure 3b. We also placed 3 tape marks on the floor to keep the tripod position fixed for the entire experiment. The tripod legs were given two full extensions. The entire setup can be seen in Figure 3a. Some of the example images taken from this setup can be seen in Figure 2. Furthermore, we associated different levels with different heights of the camera. The height was increased by rotating the center column crank of the tripod by the angles as shown in Table 1.



(a) Setup **(b) Turntable Cover**

Figure 3: 3a shows the tripod, turntable, and white cloth. 3b is the schematic of chart paper placed on turntable

Table 1: Tripod Height Adjustment w.r.t Knob Rotation

Height Level	Knob Rotation
Level 0	0π
Level 1	$+3\pi$ from Level 0
Level 2	$+3\pi$ from Level 1
Level 3	$+3\pi$ from Level 2
Level 4	$+2\pi$ from Level 3

We used the 'Redmi Note 8 Pro' smartphone camera to capture the photographs. For each level, we clicked 24 photographs. The 24 markers on the turntable helped to vary the poses and to capture the crop from all directions. Overall, 120 photographs were given as input to the MeshroomCL [9] software.

The MeshroomCL runs a pipeline that outputs a 3D textured mesh using all the captured photographs. The pipeline [9] runs Structure from Motion (SfM) and Multi View Stereo (MVS) for the 3D dense reconstruction of the mesh. The SfM stage estimates the camera poses and reconstructs the environment. The MVS stage uses that information to reconstruct the 3D shape in the images and outputs a dense mesh as shown in Figure 4. Afterward, the Mesh Filtering stage and Texturing stage run. The output of the pipeline consists of an obj file along with several texture files.

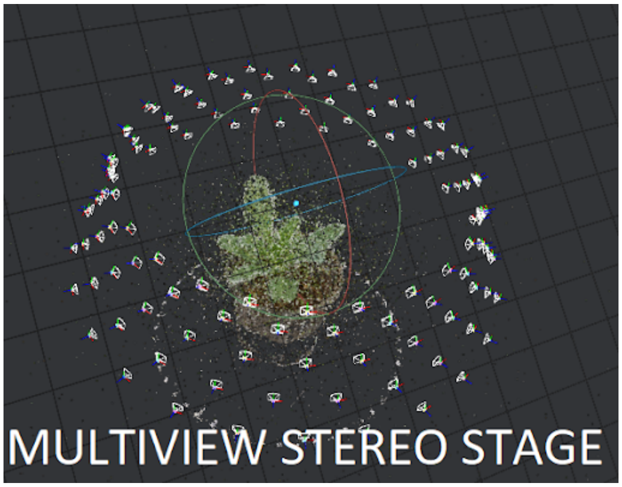


Figure 4: Multi View Stereo(MVS) Stage of MeshroomCL (3D Dense Reconstruction Using 120 Images)

3.3 Alignment of Crop Assets

The meshes outputted by the photogrammetry are not aligned even if the physical setup remains fixed because the actual crop grows and as a result, the geometry of the mesh changes daily. To resolve that, we did pairwise aligning of the assets. The already aligned $(n-1)^{th}$ day asset was considered as the reference for the n^{th} day asset. In other words, we aligned the second-day asset with respect to the first-day asset, the third-day with respect to the second-day, and so on. The Iterative Closest Point (ICP) algorithm was used which needed 4 initial points from both assets for each alignment process. It aligned the second asset about these 4 selected points ensuring that the distance between all other corresponding points of the two assets was minimal.

Those 4 points must not vary in time for the assets to be aligned. Hence, we selected 4 points on the visible soil surface near the bottom of the shoot system of the crop. There is a scope to improve the process of selection of these points, which could be dealt with in future work. The asset alignment that we got was sufficiently accurate for the human eye to perceive the crop pot as a static entity as per our requirement. Furthermore, we proposed an error metric to evaluate this alignment in a quantitative way which can be found in the Section 4.

3.4 Model Integration

The last step was the integration. All the assets were collected and imported to Unity software. They were already aligned among themselves, as the alignment step was already done earlier. Hence, they were all added as meshes under a single parent 'GameObject'. Now this parent object was aligned in the unity scene. Our logic was to activate the corresponding mesh on button clicks. To achieve that, two C# scripts were attached to the HTC Vive Controllers (left and right) that changed a static variable responsible for the activation of the corresponding mesh.

4 EVALUATING THE FRAMEWORK

The validation of VR frameworks is usually done at a qualitative level. It considers several concepts like presence, immersion, and a typology of fidelity [10]. However, a reliable quantitative measure is still not proposed yet. We believe that this numerical measure should be specific to each framework.

As in our framework, the alignment of the 3D assets is a prominent component because it controls the realism of the crop growth depiction. So, we propose an error metric that calculates the average alignment error ('*AAE*'). A lower '*AAE*' implies better overall asset alignment and higher realism. But prior to that, some terms that are used later have been explained below.

4.1 Representative Point of a 3D Asset

The property that this point must have is that it should be static over time. Hence, no point on the crop could be the representative point of a 3D asset because the crop grows. Therefore, our focus shifts to the static part which is the pot having a frustum geometry. This point could be any point on our pot provided it doesn't move while we switch the asset in our framework. Considering all this, it makes sense to consider the center of the upper circle of our geometry as the representative point. However, finding the 3D coordinate of that point is difficult because the triangular mesh of each of our 3D assets has about 0.4 million vertices on average in our case. Our approach to finding that point is as follows:-

- We reduce the original mesh (say A) to a new mesh (say A') i.e. $A \rightarrow A'$.
  - First, we remove the vertices that are associated with the crop and the lower pot portions from the asset.
  - Then we remove the remaining vertices to get two arcs on the upper circular periphery as in Figure 5. This can be done easily with visual inspection.

  These arcs should ensure that at least one diameter of that periphery passes through them, so they could be a little longer than the ones shown in Figure 5.
- We find a pair of vertices that have the maximum distance between them in the mesh A', which would be equal to the length of any diameter of the upper periphery circle. We processed the mesh A' using the OpenMesh library in Python and outputted a pair of vertices.
- We calculate the centroid of those two vertices, which would also be the center of the diameter passing through those vertices. This point is considered the representative point of the 3D asset.

The number of vertices was reduced from about 0.4 million on average for the original meshes to just about 1800 on average for the final meshes ($\approx$ 99.5% reduction). The sole purpose of this reduction was to reduce the computations that were required for the last two steps.

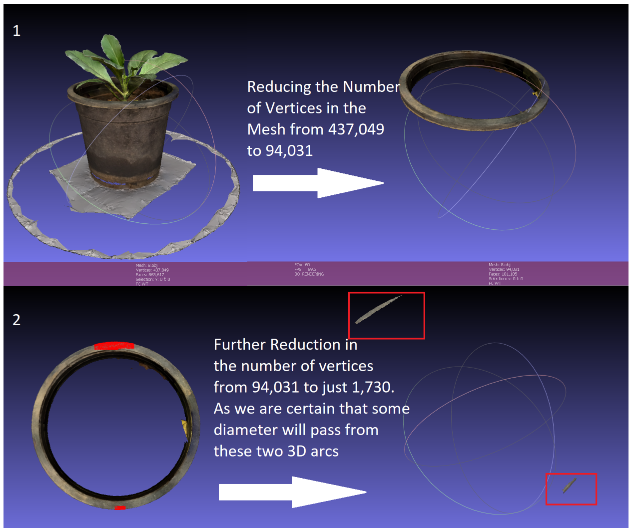


Figure 5: Removing unnecessary vertices from the 3D asset

4.2 Average Alignment Error

First of all, we define the align error ('*AE*') of two aligned meshes as the Euclidean distance between their representative points. Then the entire alignment process was done in an iterative pairwise manner as in Section 3.3. The final '*AAE*' for the entire scenario is attained after taking the arithmetic mean of all the *AE*s. The total number of alignment steps is $(N-1)$, where N is the number of timestamped assets. The final calculations involved are as follows:

- Find the representative point of each 3D asset as Rep_i for the i^{th} asset.
- Find the alignment error for each consecutive pair

$$AE_{(i+1,i)} = ||Rep_{i+1} - Rep_i||_2 \tag{1}$$

- Find the average of all the *AE*s to get the final error.

$$AAE = \sum_{1 \le i < N} AE_{(i+1,i)}/(N-1) \tag{2}$$

This metric considers the relative spatial alignment of the temporal assets. A lower *AAE* is highly desirable for more realism in our VR framework.

5 CALCULATIONS AND RESULTS

We had 13 assets representing the growth of the cauliflower for 13 consecutive days. The representative points of all these 13 assets were calculated as mentioned above. Then the alignment errors (*AE*s) were also calculated for the consecutive pairs. These values are contained in Table 2. Using these *AE*s, the final *AAE* was computed as 0.0656, which quantifies the realism of our framework.

6 USER STUDY

We also performed a user study to confirm whether the proposed quantitative metric was consistent with the qualitative metric or not. For the qualitative metric, a rating was given by each user from 1 (Very Bad) to 5 (Very Good). This rating was considered for each asset transition where an 'asset transition' implies swapping any day's asset to the next day's asset. The users were asked to give ratings based on the change in the pot's characteristics namely its position, orientation, and size. It means the higher the change in the pot characteristics, the lower the rating. We had 17 users in this study. Each user was invited into our lab and given the HTC Vive Headset to experience our VR framework. The ratings for all 12 transitions were collected. Since the ratings and the error metric would follow an opposite trend; for a comparative observation, the average ratings for each transition were converted to normalized inverse average ratings ('*NIAvgRatings*').

$$NIAvg\ Rating_{(i+1,i)} = (5 - \overline{Rating_{(i+1,i)}})/5 \tag{3}$$

where $\overline{Rating_{(i+1,i)}}$ is the average of the ratings given by all users for the transition i^{th} to $(i+1)^{th}$ asset. The calculated '*NIAvgRatings*' can be found in Table 2. Furthermore, the Pearson Correlation Coefficient was also calculated between the two quantities as 0.7564. The value implies that there is a strong correlation between the quantitative metric and the *NIAvgRatings* given by the user study.

Table 2: *AE*s and *NIAvgRatings* between the consecutive pairs

Between	*AE*	*NIAvgRating*
1 and 2	0.0198	0.2000
2 and 3	0.0234	0.3750
3 and 4	0.1670	0.5250
4 and 5	0.1308	0.4625
5 and 6	0.0539	0.3250
6 and 7	0.0287	0.2250
7 and 8	0.0555	0.2875
8 and 9	0.0840	0.2625
9 and 10	0.0520	0.1375
10 and 11	0.0685	0.2750
11 and 12	0.0648	0.2750
12 and 13	0.0393	0.2750

7 CONCLUSION & FUTURE SCOPE

In this research, we have developed a novel digital twin framework that is useful to farmers. This framework can be used to create growth models of crops using VR technology. We have also proposed an evaluation metric for our VR framework and validated it with a user study. In the future, we want to automate the development process completely so that growth models of new crops can be added with minimal effort. Furthermore, we also aim to increase the number of assets per crop and animate the transitions between them. Afterward, we intend to simulate various effects of the change in abiotic (water, sunlight, and temperature) and biotic (pests) factors on the crop in this framework.

ACKNOWLEDGMENTS

This work was supported by the Department of Science and Technology (DST), Government of India, for the Technology Innovation Hub at IIT Ropar within the framework of the National Mission on Interdisciplinary Cyber-Physical Systems.

REFERENCES

[1] Maricelis Acevedo, Kevin Pixley, Nkulumo Zinyengere, Sisi Meng, Hale Tufan, Karen Cichy, Livia Bizikova, Krista Isaacs, Kate Ghezzi-Kopel, and Jaron Porciello. 2020. A scoping review of adoption of climate-resilient crops by small-scale producers in low- and middle-income countries. *Nature Plants* (October 2020). https://doi.org/10.1038/s41477-020-00783-z

[2] Doug A. Bowman and Ryan P. McMahan. 2007. Virtual Reality: How Much Immersion Is Enough? *Computer* 40 (July 2007). https://doi.org/10.1109/MC.2007.257

[3] Quan Cap, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. 2020. LeafGAN: An Effective Data Augmentation Method for Practical Plant Disease Diagnosis. *IEEE Transactions on Automation Science and Engineering* (December 2020), 10 pages. https://doi.org/10.1109/TASE.2020.3041499

[4] Yiping Chen, Jinchao Pan, and Qiufeng Wu. 2023. Apple leaf disease identification via improved CycleGAN and convolutional neural network. *Soft Computing - A Fusion of Foundations, Methodologies and Applications* 27 (July 2023), 13 pages. https://doi.org/10.1007/s00500-023-07811-y

[5] Xiaohui Cui, Yongzhi Ying, and Zhibo Chen. 2021. CycleGAN based confusion model for cross-species plant disease image migration. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology* 41 (2021), 11 pages. https://doi.org/10.3233/JIFS-210585

[6] L.G. Divyanth, D.S. Guru, Peeyush Soni, Rajendra Machavaram, Mohammad Nadimi, and Jitendra Paliwal. 2022. Image-to-Image Translation-Based Data Augmentation for Improving Crop/Weed Classification Models for Precision Agriculture Applications. *Algorithms* (October 2022). https://doi.org/10.3390/a15110401

[7] R. A. Gilbert, K. J. Boote, and J. M. Bennett. 2002. On-Farm Testing of the pnutGRO Crop Growth Model in Florida. *Peanut Science* 29 (January 2002). https://doi.org/10.3146/pnut.29.1.0011

[8] C. Giménez, M. Gallardo, C. Martínez-Gaitán, C. O. Stöckle, R. B. Thompson, and M. R. Granados. 2012. VegSyst, a simulation model of daily crop growth, nitrogen uptake and evapotranspiration for pepper crops for use in an on-farm decision support system. *Irrigation Science* 31 (February 2012). https://doi.org/10.1007/s00271-011-0312-2

[9] Carsten Griwodz, Simone Gasparini, Lilian Calvet, Pierre Gurdjos, Fabien Castan, Benoit Maujean, Gregoire De Lillo, and Yann Lanthony. 2021. AliceVision Meshroom: An open-source 3D reconstruction pipeline. In *Proceedings of the 12th ACM Multimedia Systems Conference - MMSys '21*. ACM Press. https://doi.org/10.1145/3458305.3478443

[10] David J. Harris, Jonathan M. Bird, Philip A. Smart, Mark R. Wilson, and Samuel J. Vine. 2020. A Framework for the Testing and Validation of Simulated Environments in Experimentation and Training. *Frontiers* 11 (March 2020). https://doi.org/10.3389/fpsyg.2020.00605

[11] Guangqing Li, Zhujie Xie, Xueqin Yao, and Xuehao Chen. 2011. Study on the mathematical model of the effects of NPK on winter cauliflower. *Mathematical and Computer Modelling* 54 (August 2011). https://doi.org/10.1016/j.mcm.2010.11.045

[12] Haseeb Nazki, Jaehwan Lee, Sook Yoon, and Dong Sun Park. 2019. Image-to-Image Translation with GAN for Synthetic Data Augmentation in Plant Disease Datasets. *Smart Media Journal* 8 (June 2019), 11 pages. https://doi.org/10.30693/SMJ.2019.8.2.46

[13] Ebenezer Olaniyia, Dong Chenb, Yuzhen Lua, and Yanbo Huangc. 2022. Generative Adversarial Networks for Image Augmentation in Agriculture: A Systematic Review. *Computers and Electronics in Agriculture* (September 2022). https://doi.org/10.1016/j.compag.2022.107208

[14] Mingle Xu, Sook Yoon, Alvaro Fuentes, Jucheng Yang, and Dong Sun Park. 2022. Style-Consistent Image Translation: A Novel Data Augmentation Paradigm to Improve Plant Disease Recognition. *Frontiers in Plant Science* (February 2022). https://doi.org/10.3389/fpls.2021.773142

[15] Feng Yu, Jun feng Zhang, Yousen Zhao, Ji chun Zhao, Cuiping Tan, and Ru peng Luan. 2010. The Research and Application of Virtual Reality (VR) Technology in Agriculture Science. In *Computer and Computing Technologies in Agriculture III*. https://doi.org/10.1007/978-3-642-12220-0_79