Project Feasibility Study

Alex Foote

1 Aim

The project aims to develop a plant phenotyping framework that constructs 3-dimensional (3D) models of plants from images and extracts useful metrics from the models to enable the monitoring of plant traits over time. Additionally, methods for segmentation of the 3D plant models will be investigated.

2 Objectives

The aim of the project will be achieved through the following objectives:

- Obtain at least one publicly available dataset containing images of plants taken from multiple views and across part of their lifespans
- Obtain appropriate open source software that produces 3D representations given the obtained dataset as input
- Determine biologically relevant plant traits to measure from the 3D models, based on existing phenotyping literature
- Develop software that can extract the required traits from the 3D models and track trait changes over time
- Investigate methods for plant segmentation, in particular the possibility of applying deep learning techniques
- Stretch target: Develop software for 3D plant model segmentation
- Obtain or create a dataset containing the constructed 3D plant models and ground-truth for the required traits
- Evaluate the accuracy of the software using the ground-truth data
- Analyse the software quantitatively, using the accuracy data, and qualitatively, in the context of other similar research
- Produce recommendations for future research

3 Literature review and project scope

The construction of 3D plant models is challenging due to the complex architecture of many plants [1]. This project intends to focus on software that constructs models from visible light images taken of an object from multiple viewpoints, as the equipment needed (one or more standard digital cameras) is more accessible and therefore potentially more scalable than active light approaches such as LiDAR, laser scanning, or depth imaging [2]. The construction of 3D models from multi-view images is well researched and a number of approaches exist, including space carving, shape-from-silhouette (SFS), structure-from-motion (SFM), multi-view stereo, and others. The current intent, therefore, is to investigate the potential to use the open source software VisualSFM [3, 4] to construct a 3D point cloud from multi-view images. The possibility of extracting metric information directly from the point cloud will be investigated, with an alternative option of using open source Canopy Reconstruction software [1] to construct a 3D mesh from the point cloud and analysing the mesh.

As mentioned earlier, a dataset containing multi-view images of plants is required to construct the 3D models. Creating this dataset as part of the project would present a number of challenges, in particular the growing of plants in a controlled environment and the regular capture of images. As there are relevant datasets publicly available, these will be used as an alternative to creating a dataset.

Once a 3D plant model has been constructed, useful phenotype information must be extracted. A wide range of plant indexes have been proposed and used in the literature, and so part of the project will be choosing suitable metrics to extract that are biologically relevant, i.e., useful for plant breeding. Certain metrics, such as height, width, convex hull area, and biomass can be obtained directly from a 3D model, and are sometimes referred to as non-complex traits. In contrast, traits such as leaf area, leaf number, and stem length require the segmentation of the model into its organs, i.e., leaves, stem, fruit, etc., and as such may be referred to as complex traits [5]. The initial focus for the project will be to develop software to extract non-complex traits, and plant segmentation will be investigated to determine its feasibility as part of the project.

Ideally, plant phenotyping software would operate independent of plant species, i.e., able to extract traits across a variety of diverse plant species. Due to the general nature of the 3D reconstruction software, this part of the proposed framework should naturally fulfill this requirement. The extraction of non-complex traits should not require narrow species-specific algorithms either, and so should also fulfill this requirement. However, plant segmentation often uses algorithms that are not robust to diverse plant species. Deep learning techniques have been shown to outperform hand-crafted algorithms for 2D segmentation, and are generally significantly more robust. There is a small amount of early research applying deep learning techniques to 3D plant segmentation [6]. The possibility of implementing a deep learning based plant segmentation algorithm will be investigated, and will serve as an ambitious stretch target for the project that will only be attempted if the project progresses faster than expected.

The publicly available datasets contain multi-view images taken across time, which presents the possibility of tracking the desired traits over time. This is known as 4-dimensional (4D) phenotyping and is useful for measuring growth, as well as detecting plant stress [5]. As part of the project, software will be developed that tracks trait changes over time and a suitable method of visualising this data, for example through growth curves, will be determined.

An important aspect of the project is the evaluation of the software. There are likely to be two primary parts of the software to evaluate - the accuracy of the construction of the 3D model and the accuracy of trait extraction from a model. As part of the development of the Canopy Reconstruction software [1], the authors evaluated its accuracy, so if this software is used their results may be suitable to evaluate the 3D model accuracy. Publicly available datasets that contain 3D models and ground truth for desired metrics seem to be uncommon. Further investigation into their availability will be conducted. If no such datasets are available, the creation of such a dataset will likely be necessary to evaluate trait extraction accuracy. Determining the best way to accomplish this will be investigated further, with a current possibility being the use of software such as LabelMe [7] to manually extract the desired traits. Additionally, if the development of plant segmentation software is included in the project, a method for evaluating its accuracy will need to be investigated as well.

4 Justification

Improving the future of global food security - the availability of sufficient quantities of sufficiently nutritious food to the world's population - will require significant increases in food production, due to population growth, the increased demand for biofuels causing competition for crops, and other factors [8]. Food production is estimated to need to increase by 25%-70% by 2050 to meet increased food demand [9], which will require the development of higher yield food crops. Additionally, climate change is enhancing the environmental stresses faced by crops [10], necessitating the development of increasingly resilient crop species.

The development of plant species with desirable traits, such as those just described, is achieved through selective breeding and genetic engineering. An important step in both of these methods is the measurement of various physical characteristics of plants, to quantify the response of a selected plant genotype to its environment. This is known as plant phenotyping, and has historically been a bottleneck in plant breeding [11]. To address this bottleneck, high-throughput, automated phenotyping systems are being developed. These systems are progressing towards large-scale breeding programs that grow plants in controlled environments, automatically capture data such as images, and extract useful metrics from this data, such as plant height, leaf number, leaf area, and many others.

The majority of prior research in plant phenotyping has focused on 2D methods, with a recent increase in interest in 3D methods. Accurate 3D plant models are particularly valuable for plant phenotyping, allowing for the extraction of a wider variety of traits at a greater accuracy [2]. Additionally, as mentioned earlier, general phenotyping frameworks that perform well over a diverse range of plant species are more valuable than narrow frameworks. For these reasons, the project will focus on the development of a 3D plant phenotyping framework, with an emphasis on maintaining the generality of the framework to allow it to operate on a broad range of plant species with varying architectures and images taken under differing conditions.

Previously, plant phenotyping relied upon measurements taken by hand, and often used destructive methods, i.e., the plant was killed in the process of measurement, meaning measurements could only be taken at a single time point. The development of non-destructive image-based phenotyping means that measurements can be taken across a plant's lifespan and the change in traits can be analysed. The ability to track changes in plant traits over time presents the opportunity to monitor plant growth and stress response [12], information that is valuable for plant breeding. As such, part of the project will be to develop software that monitors changes in extracted plant traits over time and quantifies plant growth.

The measurement of complex plant traits requires the segmentation of 3D plant models into their constituent organs. The measurement of complex traits allows for more detailed information about plant growth, as traits such as leaf number and leaf area can be tracked over time [5]. Much of the research on plant trait measurement using 3D models has focused on the extraction of non-complex traits due to the challenge of 3D plant segmentation [6]. Part of the project will therefore involve researching possible methods for segmentation. The development of segmentation software is set as a stretch target for the project if the conducted research suggests it is feasible.

5 Risk assessment

Table 1 presents possible risks that could impede progress on the project. Risks are given a probability (P) on a scale of 0-1, and a severity (S) on a scale of 1-10. Potential ways to help mitigate the risk are discussed, and a final risk rating of low (L), medium (M), or high (H) is given.

Table 1: Assessment of possible risks. Key: P - probability; S - severity; L - low; M -medium; and H - high.

Risk	P	S	Methods for mitigation	Risk rating
Unable to obtain a suitable multi-view image dataset	0.4	8	Other datasets obtained using LiDAR, depth imaging, etc., could be used instead	M
Unable to obtain suitable 3D model construction software	0.1	8	The focus of the project could shift to analysing 2D images	L
Unable to obtain or create labelled 3D model dataset for software evaluation	0.4	7	2D image labelling to create ground truth data is highly likely to be possible and 2D images can be derived from a 3D model, therefore 2D evaluation could be used as a proxy	L
Large datasets could be challenging to work with due to equipment constraints	0.7	5	For large datasets, a small subset will be used during software development	M
Project work could be lost, e.g., through computer malfunction	0.2	8	All work will be backed up in a physical location and in the cloud	L
Project delays caused by build up of project and other course work	0.4	6	Weekly supervisor meetings will help ensure consistent progress, university resources are available to improve time management ability if needed	L
Balancing time spent on project and other commitments could cause stress, reducing ability to work on project	0.3	5	Wellbeing will be monitored through logbook, many resources for help dealing with stress if needed	L
Illness or other circumstances that cannot be controlled could temporarily reduce ability to work on project	0.4	4	Any such circumstances will be logged in Tabula to provide evidence of mitigating circumstances	L

5.1 Health and safety

As this is a software development project, the majority (if not all) of the project will involve the use of a computer. Computer use for long periods can put the user at risk of pain in the neck, wrists, and other areas [13]. A properly set up workstation can help reduce this risk, so a workstation assessment will be completed using the display screen equipment workstation checklist [14]. Additionally, regular short breaks will be taken during sustained periods of work at a computer. To ensure these breaks are taken, a timer will be set during sustained periods of computer use to as a reminder to take a five minute break every hour, as recommended in [13].

Another key health and safety consideration is personal wellbeing. The project has the potential to cause stress, due to workload, deadlines, and competition with other committments. To ensure maintained wellbeing throughout the project, a number of measures will be taken. Wellbeing will be monitored in the online project logbook, to allow for the early identification of problems. Any issues can then be addressed through support services provided by the University, through consultation with the project supervisor, or through talking to family and/or friends. Additionally, a wellbeing plan will be created at the start of the project, that considers preventative measures that can reduce the risk of excess stress or wellbeing problems, such as regular exercise, a consistent sleep schedule, good time management, etc.

6 Project plan

Table 2 presents an initial timeline for the project. It covers the project objectives and the major pieces of work that make up the project submission, and provides an estimate of when work will start on each task, as well as how long it will take.

6.1 Progress monitoring

During the project, progress will be monitored through the online logbook. This section will briefly address the relevant aspects of the logbook.

6.1.1 Meeting minutes

Project supervisor meetings take place every Wednesday. During the meetings, minutes are taken that contain brief summaries of the topics discussed and actions to be taken before the next meeting. These minutes are then uploaded to the logbook.

6.1.2 Time management and planning

This section of the logbook is used to create and review more detailed plans made throughout the project. After the weekly project supervisor meeting, a Gantt chart [15] planning the week ahead is created. This chart takes each action agreed in the meeting and plans the time over which it will completed. During the week, progress is updated in the Gantt chart, and necessary modifications are made as new tasks arise or circumstances change. At the end of each week, the initial and final Gantt charts are uploaded to the logbook, and a brief review of the discrepancies is conducted, to enable the improvement of time management throughout the project.

6.1.3 Risk registers

This section is for project risk management. As new risks arise throughout the project, they will be added to the initial risk assessment shown in Table 1. This will allow for their overall risk to be quantified, and inform any actions that may need to be taken to help reduce the risk.

6.1.4 Health and safety

Similarly to the risk register section, as new health and safety risks arise, they will be recorded in this logbook section. Methods for dealing with the risk will be discussed, and if necessary actions will be taken to reduce the risk.

6.1.5 Wellbeing

In this section, a full wellbeing plan will be created, expanding on the information presented in this report. Evidence of actions taken to fulfill the proposals in the wellbeing plan will be uploaded.

6.1.6 Budget

In this section, any project expenses will be recorded. Each expense will be justified in terms of its necessity to the completion of the project.

6.1.7 Self reflection and learning plan

In this section, an initial learning plan that discusses currently known deficits in the skills or knowledge required to complete the project will be created. Methods for addressing these deficits will be presented, and evidence of completion of said methods will be recorded. Additionally, throughout the project the same steps will be taken whenever a new deficit is identified. Finally, at the end of the project the learning that has taken place will be

Table 2: Initial project timeline

Milestone Description	Start	End	Duration
Dataset choice	29/10/19	12/11/19	14
3D software choice	29/10/19	12/11/19	14
Develop trait extraction software	12/11/19	22/12/19	40
Investigate segmentation options	02/01/20	20/01/20	18
Create evaluation dataset	20/01/20	05/02/20	16
Write draft Technical Report	29/10/19	11/02/20	105
Modify Technical Report based on supervisor feedback	19/02/20	04/03/20	14
Write Project Evaluation	12/02/20	04/03/20	21
Final Report and Evaluation adjustments and submission	04/03/20	11/03/20	7
Monitor progress through Logbook	29/10/19	13/03/20	136

reflected upon. The longer-term impact on the future that the skills learnt during the project could have will be discussed.

6.1.8 General progress records

This section will contain any information on project progress that does not fit into the other pages. Regular progress updates will be recorded, to provide evidence of consistent progress throughout the project. Additionally, the initial project timeline presented in Table 2 will be reviewed and updated monthly. This will ensure that the large-scale objectives are being met in a timely manner, and allow decisions to be made about whether stretch targets will be attempted, or whether the objectives of the project need to be modified to reflect difficulties caused by unexpected obstacles, in accordance to the below back up plan.

6.2 Back up plan

Although this report has attempted to define a realistic project scope, and has considered possible risks and methods to mitigate these risks, there is still the possibility of unexpected events disrupting the project. Any reaction to these events will depend significantly on the nature of the event itself, and so specific plans cannot be put into place at this time. However, general possibilities for reducing the scope of the project can be considered. These could then be implemented in the case of significant disruption to the project.

The main possibility would be to remove the 3D modelling aspect of the project, and instead extract plant traits from 2D images. This would reduce the complexity of the project whilst still allowing work of value to occur. 2D plant phenotyping is an active area of research, and so there are still interesting problems that would fulfill the project justification presented earlier, as well as the overall requirements for a 3rd Year Project.

This would be an option if early on into the project significant challenges regarding the construction of 3D plant models were encountered.

If the above is not an option, due to the point at which problems have arisen, then another option would be to remove some later objectives and focus on fulfilling the earlier objectives. For example, if problems arose during the development of the trait extraction software, focus could shift back to the earlier stage of 3D model construction, perhaps instead trying to add a novel element to the construction software. This would again allow for the project justification and the project requirements to be met, even in the face of unexpected challenges.

7 Ethics

Ethical approval for a project is required when data involving humans is to be gathered. This project will only use data collected from plants, and additionally it is unlikely that new data will be gathered - rather, existing publically available datasets will be re-analysed. As such, ethical approval for the project is not required, and this has been confirmed through the completion of the online ethical flowchart.

References

[1] M. P. Pound, A. P. French, E. H. Murchie, and T. P. Pridmore, "Automated recovery of three-dimensional models of plant shoots from multiple color images," *Plant Physiology*, vol. 166, no. 4, pp. 1688–1698, 2014.

[2] J. Gibbs, M. Pound, A. French, D. Wells, E. Murchie, and T. Pridmore, "Approaches to three-dimensional reconstruction of plant shoot topology and geometry," *Functional Plant Biology*, vol. 44, 2016.

[3] C. Wu, "Towards linear-time incremental structure from motion," *Proceedings - 2013 International Conference on 3D Vision, 3DV 2013*, pp. 127–134, 2013.

[4] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz, "Multicore bundle adjustment," in *CVPR 2011*, 2011, pp. 3057–3064.

[5] S. Paulus, "Measuring crops in 3d: using geometry for plant phenotyping," *Plant Methods*, vol. 15, no. 103, 2019.

[6] W. Shi, R. Zedde, H. Jiang, and G. Kootstra, "Plant-part segmentation using deep learning and multi-view vision," *Biosystems Engineering*, vol. 187, pp. 81–95, 2019.

[7] B. Russell, A. Torralba, and K. Murphy, “Labelme: A database and web-based tool for image annotation,” *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 157–173, 2008.

[8] FAO, “The future of food and agriculture – trends and challenges,” 2017.

[9] M. C. Hunter, R. G. Smith, M. E. Schipanski, L. W. Atwood, and D. A. Mortensen, “Agriculture in 2050: Recalibrating Targets for Sustainable Intensification,” *BioScience*, vol. 67, no. 4, pp. 386–391, 2017.

[10] S. S. Myers, M. R. Smith, S. Guth, C. D. Golden, B. Vaitla, N. D. Mueller, A. D. Dangour, and P. Huybers, “Climate change and global food systems: Potential impacts on food security and undernutrition,” *Annual Review of Public Health*, vol. 38, no. 1, pp. 259–277, 2017.

[11] F. Fiorani and U. Schurr, “Future scenarios for plant phenotyping,” *Annual Review of Plant Biology*, vol. 64, no. 1, pp. 267–291, 2013.

[12] N. Fahlgren, M. Gehan, and I. Baxter, “Lights, camera, action: high-throughput plant phenotyping is ready for a close-up,” *Current Opinion in Plant Biology*, vol. 24, pp. 93–99, 2015.

[13] Working safely with display screen equipment. Accessed October 27, 2019. [Online]. Available: https://www.hse.gov.uk/msd/dse/

[14] Display screen equipment (dse) workstation checklist. Accessed October 27, 2019. [Online]. Available: https://www.hse.gov.uk/pubns/ck1.pdf

[15] W. Clark, W. Polakov, and F. Trabold, *The Gantt chart, a working tool of management.* New York, The Ronald press company, 1922, accessed October 27, 2019. [Online]. Available: https://archive.org/details/cu31924004570853/page/n7