Studentship Project: Annual Progress Report DEC/2023 to OCT/2024

Student Name:	James	AHDB Project Number:	SF/TF 170/a
Project Title:	High-throughput phenotyping of fruit traits for automatic strawberry harvesting		
Lead Partner:	Berry Gardens		
Supervisor:	Grzegorz Cielniak and Daniel Sargent		
Start Date:	01/10/2021	End Date:	31/12/2025

1. Project aims and objectives

Strawberries are a berry crop of major economic importance, grown commercially throughout the temperate and subtropical zones of the world. Labour, or rather the lack of it, poses a huge challenge to the soft fruit industry. In a survey conducted by British Summer Fruits, the industry body representing 95% of all British grown soft fruits), a loss of £36.5 million was reported in 2021, equating to 7709 tonnes of fruit, solely due to berry waste from lack of labour (British Summer Fruits Members' Survey, 2021).

A promising technological solution to these labour problems is the introduction of fruit-picking robots (Ducket et al, 2018). Considering the strawberry industry alone, a substantial amount of research has been conducted towards developing strawberry harvesters, spanning both academic institutions and commercial start-ups around the world (Kondratieva et al, 2022).

Automated phenotyping of strawberry plants has the potential to revolutionise the breeding process by supplying high-throughput, quantitative measurements which are infeasible to complete through current manual methods (Zheng et al, 2021). Furthermore, different plant architectures have the potential to pose a greater or lesser challenge to automated harvesters and capturing this through a complexity measure as a phenotypic trait during selection has the potential to result in future varieties which simplify the harvesting task.

The objectives of this PhD centre around developing and evaluating methods for automated phenotyping of strawberry and linking these phenotypes to measures of structural complexity.

This will be approached through four objectives, as follows:

O1. Determine the current status of automation of relevant phenotypic traits and the potential for automation of unautomated traits

O2. Investigate and implement algorithms necessary to conduct standard holistic and component phenotyping tasks

O3. Investigate how structural complexity can be linked to both phenotypes and harvest success

2. Key messages emerging from the project

In the first year, we conducted and published a review to establish the current status of automated phenotyping in strawberry (James, 2022). This review highlighted that very little research has been done into automating the phenotyping of strawberry, as the only traits which have been automated relate to fruit and phenology, and research has not been focussed on the automation of traits such as peduncle length or truss complexity.

Access to relevant datasets (annotation rich 3D point clouds of strawberry) is also a challenge. We collected a dataset in the first year of this project to address this, but even a high-quality dataset poses challenges to phenotyping methods which must be considered by developers.

Partial point clouds, due to occlusion, pose a challenge to accurate phenotyping and need research to develop methods which suitably handle this to ensure that measurements are biologically relevant.

3. Summary of results from the reporting year

In this, the third year of the PhD, the focus of work has been distributed across three themes; datasets, skeletonisation (converting a point cloud to a thin representative structure through its centre) and harvest complexity.

Theme 0: Data

Access to relevant datasets is one of the largest challenges facing the automated phenotyping research community. This year I presented our dataset paper of high-quality point clouds associated with rich annotations of individual organs at the Computer Vision Problems in Plant Phenotyping and Agriculture Workshop at ECCV in Milan. This paper included benchmarks for stem skeletonisation, which is essential to making measurements relating to the topology of the plant and interpreting accurate measurements of peduncles, petioles and other stem-like structures in the plant.

Synthetic data has the potential to further address the lack of annotated data available. To this end, I have refined a 3D procedurally generated strawberry model which can produce annotated point clouds with controlled phenotypic measurements. This allows the generation of a sufficient quantity of data to pretrain networks on for segmentation of point clouds into individual organs, and also allows for control over plant architecture variations.

Data from breeding trials is also essential for confirming the success of automated methods compared to manual methods. This year, with collaboration from Edward Vinson Ltd, we obtained data from a breeding trial along with real phenotypic measurements. This will form a valuable resource for further automated phenotyping research.

Theme 1: Skeletonisation

Existing skeletonisation algorithms suitable for the skeletonisation of plant structures in 3D were identified to be those of Xu (2007), Huang (2013) and Magistri (2020). The former produces plant graphs that branch too much, the second produces highly fragmented graphs. While the third produces reasonable skeleton points, these are often not centred within the stem in the cases of occlusion on one side and topological connections go awry if using the same classes as the authors (leaf and stem).

Our own research into skeletonisation has consisted of development and refinement of our variant of unsupervised k-means clustering, in which a clustering is performed around k-cylinder centres. This results in a set of skeleton points, at the top and bottom of each cylinder, forming a small segment which can then be connected using least cost to form a skeleton. This method with shown through an ablation study to deliver substantially lower error in stem length measurements at higher levels of occlusion compared to existing methods and was further demonstrated to maintain this performance on a general dataset if occlusion is the main challenge in the data. This work was presented at the Towards Autonomous Robotic Systems Conference (TAROS) in London and was awarded the best paper in field robotics.

Theme 2: Complexity

Towards developing measures of complexity that link phenotypes to ease of harvest, we produced a synthetic dataset using L-Systems in LPy. This allows for each phenotypic trait to be varied in a controlled way, facilitating exploration of the effect of each trait on complexity measures. This dataset will form the basis of work in which we hope to define measures of complexity which indicate the difficulty the structure of the plant poses to harvest as a function of phenotype. Work has commenced in defining this complexity metric.

4. Key issues to be addressed in the next year

- Define harvester agnostic harvest complexity index using the synthetic data
- Reliably compute the complexity index from a point clouds input
- Incorporate gripper geometry into the complexity measure

5. Outputs relating to the project
(events, press articles, conference posters or presentations, scientific papers):

Output	Detail
Review paper in New Phytologist Foundation Journal	James, K., Sargent, D., Whitehouse, A., & Cielniak, G. (2022). High-throughput phenotyping for breeding targets – current status and future directions of strawberry trait automation. Plants, People, Planet, 4 (5), 432–443. doi:10.1002/ppp3.10275.
TedXBrayfordPool Talk	"How agri-robotics will change the food you eat" A TedX talk presented on 3 September 2022 at TedXBrayfordPool. Abstract: Agricultural robotics and artificial intelligence are changing how we farm, from day-to-day practices to selecting new crop varieties. In this talk, Katherine James gives an overview of how agri-robotics is changing the way we interact with crops, how this will change the crops themselves and, ultimately, how this will allow humans to focus more on things which require the 'art-of-being-human'.
CTP Summer/Winter presentations	Bi-annual meeting at which a presentation of research highlights is given.
TAROS paper and presentation (best paper in field robotics)	James, K.M.F and Cielniak, G., (2024), Unsupervised clustering with geometric shape priors for improved occlusion handling in plant stem phenotyping}, In Proceedings of the 25th Towards Autonomous Robotic Systems (TAROS) Conference, Brunel, London, August 2024, 334-345, (in press).
CVPPA paper and presentation	James, K.M.F, Heiwolt, K., Sargent, D.J. and Cielniak, G., (2024), Lincoln's Annotated Spatio-Temporal Strawberry Dataset (LAST-Straw), In Proceedings of the 9th workshop on Computer Vision Problems in Plant Phenotyping and Agriculture, Milan, 29 September (in press)

6. Partners (if applicable)

Scientific partners	
Industry partners	Berry Gardens (Richard Harnden)
Government sponsor	