

Building a Plant Atlas from Real Images

Report Name	Outline Project Specification
Author (User Id)	Siôn Griffiths (sig2)
Supervisor (User Id)	Hannah Dee (hmd1)
Module	CS39440
Degree Scheme	G600 (Software Engineering)
Date	February 5, 2016
Revision	1.0
Status	Release

1 Project description

1.1 Background

The ability to determine the specific growth stage of a plant is of key interest to many plant scientists. The researchers at the National Plant Phenomics Centre (NPPC) [5] are especially interested in the accurate determination of the growth stage of a given plant in order to help inform the phenotyping experiments they carry out at their facility.

These plant growth stages are usually defined using some form of numeric scale, such as the 0-100 scale developed by Zadoks et al [10] to describe the distinct stages of cereal plants. In these scales, typically a seed will be growth stage 0 and a fully matured and dying plant will be growth stage 100. Often, these growth stage scales are presented along with hand drawn illustrations to indicate the distinct characteristics of a given feature or characteristic that defines a growth stage. The growth stage of a plant can be decided through comparison with the illustration and an associated text descriptions.

The term atlas was initially used to describe similar illustrated growth stage maps in medical fields, such as those used to describe and predict bone development/growth stages in infants by Tanner et al [7].

1.2 Project Description

This project will deliver a web-based application that presents a visualisation of oat growth stages using image data captured by the NPPC as part of an ongoing phenotyping experiment. Oat plants were chosen as target over other plants currently in experiments at the NPPC since they have fairly visually distinct, well understood growth stages and are interesting since they are a commercially viable crop of some importance to general food security.

One of the goals of this project is to provide a web-based visual reference of growth stages in a way that will assist plant scientists in digesting the vast quantity of image data produced during experiments at the NPPC. This will be achieved by presenting the images in an interactive carousel type interface which will enable side-by-side comparison of time-serialised images of a given plant and also allow quick navigation between series of images for other individual plants. Associated metadata (for example age in days, growth stage, height, weight etc) will be presented in an easily digestible format alongside each image. Some sorting of the image series on these metadata attributes will be enabled to aid with comparison of these attributes between individual plants.

In delivery of this project, the system can be used as a visual reference for growth stages accessible on many any web-capable devices. It is hoped that the system will provide a more convenient and accurate method for estimating the growth stages of oat plants by visual inspection when compared to the traditional paper-based atlases with their hand drawn sketches.

Alongside the core visualisation part of the system, there should be time to implement a machine learning approach to prediction of a growth stage for a plant from the metadata collected by staff at the NPPC. This would likely use an ID3 decision tree based approach as detailed by Quinlan [6] although exact implementation will be researched during the course of the project. Inferring the growth stage from metadata in an automated way can provide useful information to the plant scientists in the form of corroboration with true observations and potentially providing insight into previously unseen relationships between certain attributes and growth stages in a plant.

2 Proposed tasks

1. **Domain Research** - Initial research needs to be undertaken in order to understand the domain. This research will be based around both the plant science and trying to better understand growth stages and the plants in general and also around the way the work conducted at the NPPC
2. **Prototyping & Implementation Technology Evaluation** - The choice of technologies and languages used for the implementation needs to be examined and evaluated. Suitable frameworks and/or webserver technologies will need to be determined and prototyping of certain system aspects conducted.
3. **Development Environment & Build configuration** - To facilitate implementation and deployment a suitable environment configuration will need to be put in place along with certain automated processes such as continuous integration using systems such as TravisCI [8]. Github [2] will be used as remote version control repository. An investigation into suitable hosting environment for the system will be carried out as part of this task.
4. **Main Implementation** - Implementation of the core system using the chosen technologies will be carried out according to an agile, SCRUM-like approach with timeboxed iterations which will consist of design, testing and implementation with regular releases.
5. **Test Suite creation** - A thorough, automated testing strategy will be employed throughout the project stages with test driven development being employed during development and a behavioural/user-centred testing framework such as cucumber or selenium used to continuously verify front-end functionality and consistency.
6. **Stretch Goals** Given enough time after core implementation the following stretch goals are considered :
 - (a) **Infer Growth Stage from Metadata** - There's a wealth of metadata collected by the plant scientists at the NPPC, it would be useful to apply machine learning techniques to these data in order to attempt to discern the growth stage of a given plant.
 - (b) **Infer Growth Stage from Images** - Using computer vision and image processing techniques it is possible to estimate the growth stage of a plant. An example is work conducted by Boyle et al [1] at the NPPC which details an approach that fairly successfully detects flowering growth stages in oat plants.
7. **Final Report and Demonstration** - The final project documentation will be produced along with preparations for demonstrations of the final system at the end of the project.

3 Project deliverables

1. **Mid Project Demonstration** - A prototype system showcasing current project progress will be presented along with suitable documentation, likely in the form of notes or slides, during the demonstration.
2. **Implemented System** - A final implemented system will be produced as part of the project. This system will include the web server software, including all application code, necessary databases and third party software or frameworks that are required in order for the system to function as specified. Build, installation and deployment scripts will be included.

3. **Test Suite** - As part of the technical hand-in alongside the main implementation, a comprehensive test suite will be included in the form of unit test scripts, cucumber or selenium scripts and reports of manual usability testing of the front end.
4. **Design Documentation** - Documentation detailing the design choices made during the project should be produced. These will likely take the form of user stories supplemented with UML, UML-like diagrams or CRC cards.
5. **Final Report** - A final project report will be produced providing a detailed discussion of the produced application along with justification of techniques and approaches taken during implementation. An overview of research undertaken into the problem domain and discussion of results of any research based stretch goals at detailed in section 2.6

Annotated Bibliography

- [1] R. Boyle, F. Corke, and C. Howarth, "Image-based estimation of oat panicle development using local texture patterns," *Functional Plant Biology*, vol. 42, no. 5, p. 433, 2015. [Online]. Available: <http://www.publish.csiro.au/?paper=FP14056>

Paper detailing a technique used to detect oat panicles via computer vision techniques. Development of panicles can be directly correlated with certain growth stage (around GS55) in oats

- [2] "Build software better, together," GitHub, Inc. [Online]. Available: <https://github.com>

An online Git repository hosting service

- [3] D. Kendal, C. E. Hauser, G. E. Garrard, S. Jellinek, K. M. Giljohann, and J. L. Moore, "Quantifying Plant Colour and Colour Difference as Perceived by Humans Using Digital Images," *PLoS ONE*, vol. 8, no. 8, p. e72296, Aug. 2013. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0072296>

Paper detailing how a humans perception of colour in images of plants can affect judgements made about these images.

- [4] M. N. Merzlyak, A. A. Gitelson, O. B. Chivkunova, and V. Y. Rakitin, "Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening," *Physiologia Plantarum*, vol. 106, no. 1, pp. 135–141, May 1999. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1034/j.1399-3054.1999.106119.x/abstract>

Paper detailing senescence detection in plant images by analysis of colour

- [5] "National Plant Phenomics Centre," National Plant Phenomics Centre. [Online]. Available: <http://www.plant-phenomics.ac.uk/en>

National Plant Phenomics Centre

- [6] J. R. Quinlan, "Induction of Decision Trees," *Mach Learn*, vol. 1, no. 1, pp. 81–106, Mar. 1986. [Online]. Available: <http://link.springer.com/article/10.1023/A%3A1022643204877>

Paper detailing the ID3 decision tree algorithm

- [7] J. M. Tanner, R. H. Whitehouse, W. A. Marshall, M. J. R. Healty, and H. Goldstein, "Assessment of Skeleton Maturity and Maturity and Prediction of Adult Height (TW2 Method)," 1975. [Online]. Available: <http://core.tdar.org/document/125299>

Paper detailing the atlas approach used in this instance to predict adult height in human's from skeletal features in children

- [8] "Travis CI User Documentation," Travis CI GmbH. [Online]. Available: <https://docs.travis-ci.com/>

A continuous integration tool which can hook into other online resources such as GitHub

- [9] G. W. Williams, "Comparing the Joint Agreement of Several Raters with Another Rater," *Biometrics*, vol. 32, no. 3, pp. 619–627, 1976. [Online]. Available: <http://www.jstor.org/stable/2529750>

A paper describing the comparison of expert opinion with that of other experts or a group of experts

- [10] J. C. Zadoks, T. T. Chang, C. F. Konzak, and others, "A decimal code for the growth stages of cereals," *Weed res*, vol. 14, no. 6, pp. 415–421, 1974. [Online]. Available: http://old.ibpdev.net/sites/default/files/zadoks_scale_1974.pdf

Paper detailing the decimal growth stages of cereal plants