Proceedings of the 7th International Conference on Functional-Structural Plant Models, Saariselkä, Finland, 9 - 14 June 2013. Eds. Risto Sievänen, Eero Nikinmaa, Christophe Godin, Anna Lintunen & Pekka Nygren. http://www.metla.fi/fspm2013/proceedings. ISBN 978-951-651-408-9.

An automated image-processing pipeline for high-throughput analysis of root architecture in OpenAlea

J. Diener[1]*, **P. Nacry**[2], **C. Périn**[3], **A. Dievart**[3], **X. Draye**[4], **F. Boudon**[1], **A. Gaujon**[2], **C. Godin**[1]
[1]*Virtual Plants, INRIA, CIRAD, INRA, 34095 Montpellier France,* [2] *Biochimie et Physiologie Moléculaire des Plantes, INRA, CNRS, UM2, 34060 Montpellier, France,* [3] *UMR DAP, CIRAD, 34398 Montpellier, France,* [4] *Unité d'Ecophysiologie et d'Amélioration végétale, Université catholique de Louvain, B-1348 Louvain la Neuve, Belgium.*
**correspondence:* julien.diener@inria.fr

Highlights: FSPM analysis of root systems requires structural data obtained on large data set. This paper describes a processing pipeline developed to extract automatically the architecture of root system from images databases.

Keywords: Root architecture, Image processing, Analysis pipeline, High-throughput, OpenAlea

INTRODUCTION

Automated acquisition systems of Root System Architecture (RSA) are now readily available for developmental research and provide high-throughput image data of roots. Existing acquisition systems provide many types of data, from images of dispersed root pieces to full 3d scans of underground root systems. Here we consider RSA grown in Petri plates. This is a traditional experimental protocol for which image data can be acquired easily and in large amount. Their analysis is thus a major challenge for researches on root development.

Existing tools range from the automatic estimation of global traits distribution to details extraction of the root architectures through tedious manual work. Software such as Rootreader2d (Clark et al., 2012), EZ-Rhizo (Armengaud et al., 2009), Rootfly (Zeng et al., 2008) and RootTrace (Naeem et al., 2011) extract in a single image a limited part of the architecture data such as the main axes length and the number of lateral roots. Similarly, programs such as WinRhizo (Arsenault et al., 1995) or GiaRoot (Galkovskyi et al., 2012) extract representative value of the whole RSA such as total root length, area, or branching number. These are suitable for processing high-throughput data but do not provide architectural data, as they do not order the detected root segments in an arborescent structure.

To extract the full root system architecture, the available software are either entirely manual such as DART (Le Bot et al., 2010), or semi-automated such as SmartRoot (Lobet et al., 2011) or a high-end version of WinRhizo (WinRhizo Pro, 2012b). Those are suitable for detail inspection of architectural trait of a few root systems. But because of the time required by user interaction they cannot be used for high-throughput analysis.

In this study, we present a solution for extracting the full root architecture automatically and on a large-scale data set. Because our framework is included in the OpenAlea platform, in addition to traditional data analysis such as the ones provided by existing software, the extracted data can also be used seamlessly as input of all the architectural analysis packages already contained in OpenAlea.

MATERIAL AND METHOD

For the purpose of development and validation, two data sets have been used. The first set studies the growth of nine genotypes of *Arabidopsis thaliana* in three different nitrate concentrations (see fig. 2). For each of these experimental modalities, four Petri boxes containing each five root systems are daily imaged during six days. The second data set contains images of rice (*Oryza sativa* of the nipponbarre genotype) of eight Petri boxes, each containing five root systems, which were scanned once per day avec a 6 days period (see fig. 3). In total both data sets represent 688 images and 3440 root systems in total. Comparisons with expert measurements have been done for around 300 plants.

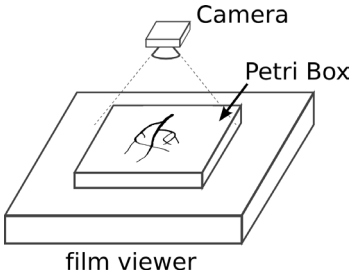


Fig. 1. Schema of the acquisition setup

Image acquisition is done either using a high-resolution scanner (~100Mp) with back-lighting for the rice data, or with a camera setup associated with a x-ray film viewer that provide back-lighting (see fig. 1)

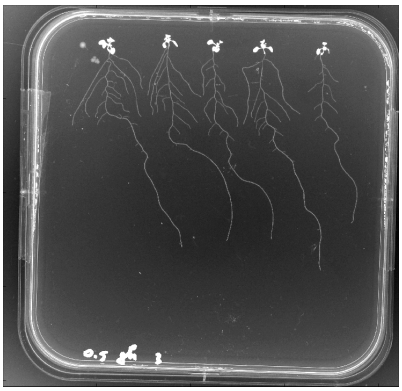


Fig. 2. One image of the Arabidopsis data set

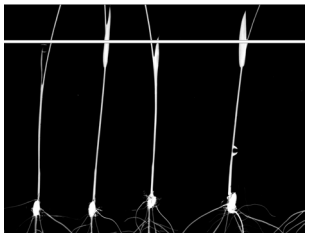


Fig. 3. One image of the Rice data set

To allow processing and analysis of large experimental data, the proposed framework organises images data sets in a two level structure. At the bottom, *root sequences* contain a time sequence of images where each image contains several root systems for a specific experimental modality (e.g. one genotype in one environment condition). On top, root sequences are organized in *projects* that first automate common processing of image sequences, and second allow comparative analysis of different experimental modalities.

All images of the project are processed through a three steps pipeline (see fig 4):

1. Image segmentation

Initial root images have to be segmented, i.e. each pixel should be classified as either background or root objects. This is done in two passes. First the smooth but non-uniform background lighting is estimated based on minimum pixel intensity with large-scale distribution. It is then removed from the original images. Second, the pixels are classified by fitting a gaussian mixture model of the remaining background noise and of the root pixels intensity.

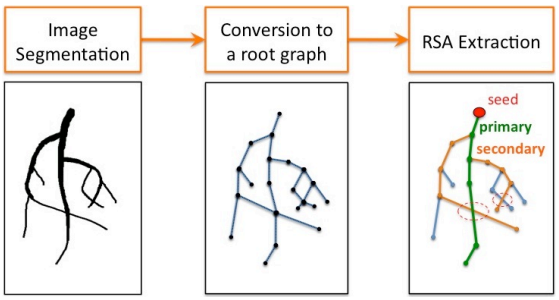


Fig. 4. Image processing pipeline

In addition to this binary classification, seeds and leaves areas are detected using specific characteristics of the observed plants. In the case of Arabidopsis, leaves are difficult to describe by their shape and they are segmented as pixels with higher intensity. Arabidopsis roots being more transparent than leaves, these last one appear whiter (see fig. 2). For rice, it is the seeds that are detected based on their radius, which is much larger than for the roots (see fig. 3).

2. Extraction of a graph representing of the observed root systems

In this step, the segmented image pixels are clustered in root *segments*, i.e. root pieces of linear shape, which contains no branching or crossing. This is done by first applying a thinning algorithm (skeletonization) to the binary images. Then each linear curve of the obtained skeleton is further divided in a set of line segment by fitting a 1st order spline on the curve pixels. A graph is then created containing the set of extracted segments as vertices, and the set of all pairs of segments that are contiguous as edges. At this stage, the constructed graph is not necessarily a tree structure as it usually contains loops.

3. Estimation of a root architecture

The root architecture is obtained by converting the root graph to an *axial tree* structure (Godin et al., 1998). This is done in two passes. First the graph is converted to a tree rooted at the detected seeds or leaves. The tree is computed using the shortest path algorithm that select the parent of each segment such that the sum of the edge cost along the path from the segment to the seed is minimal. The edge cost (i.e. the distance between two segments of the graph) is defined as the turning angle. The algorithm thus generates a tree structure with the straighter possible paths. This conversion breaks graph loops, meaning that it solves the axe crossing ambiguities (as shown by orange circles in the right image of fig. 4). The second pass then identifies branching relationship by selecting for each parent segment its direct child, i.e. that belongs to the same root axe. Again, the selection is done in order to minimize the axes curvature.

RESULTS

After the image-processing step, the project contains the sequences of the extracted arborescent structure (see fig 5).

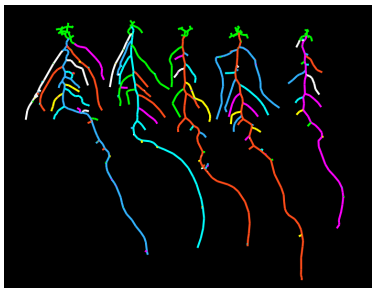
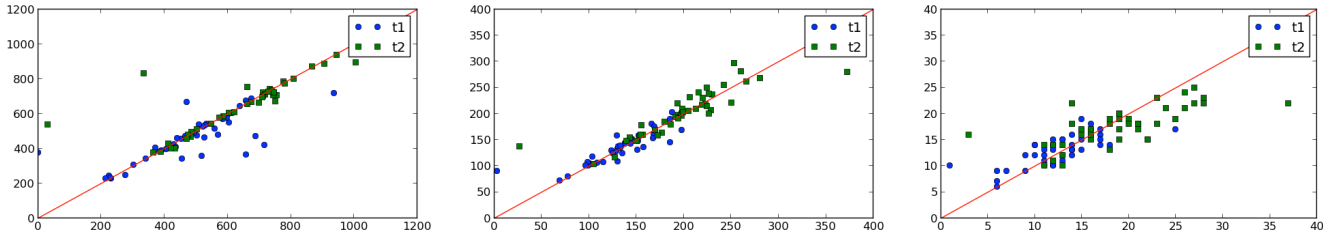


Fig. 5. Extracted RSA from the image in figure 2. Each root axe is drawn with a random colour.

In order to quantify the accuracy of our reconstruction, reference data were obtained by manual annotation of the analysed root systems. Experts manually specified the root architectures on a series of images. We then compared these reference data with the structures automatically reconstructed with our pipeline. A comparison has been done on a first data set for different RSA scales. Selected results are shown in figure 6.

We also have used the structural validation method of (Boudon et al., 2013). Two indices are provided by this method. For the first one, the reference and tested structured are mapped one onto each other to find similarities and differences. This gives the percentage of correctly identified elements. The second characterises the similarity of organisation of these elements in the two structures i.e. the percentage of correct connections between elements. This validation procedure has given promising results on an initial data set (between 80 and 90% correspondence). Results showing more complete analysis and over the whole data sets will be presented and discussed at the conference.



a) Smallest convex hull area containing each root system

b) Total root length

c) Number of secondary root axes per plant

Fig. 6: selected comparison with ground truth data. For all subfigures, the y-axis is the ground truth and x-axis the measures obtained with our method. The red line indicates exact match. The compared data are from 40 plants of *Arabidopsis thaliana* over 2 time steps.

LITERATURE CITED

Armengaud P, Zambaux K, Hills A, Sulpice R, Pattison RJ, Blatt MR, Amtmann A. 2009. EZ-Rhizo: integrated software for the fast and accurate measurement of root system architecture. *The Plant Journal* **57**:945–956

Arsenault JL, Pouleur S, Messier C, Guay R. 1995. WinRHIZO™, a root-measuring system with a unique overlap correction method. *HortScience* **30**: 906

Boudon F, Preuksakarn C, Ferraro P, Diener J, Nikinma E, Godin C. 2013. Quantitative assessement of automatic reconstructions of branching systems. *Submitted*

Clark R T, Famoso A N, Zhao K, et al. 2012. High-throughput 2D root system phenotyping platform facilitates genetic analysis of root growth and development. *Plant Cell & Environment* **36**:454-466

Galkovskyi T, Mileyko Y, Bucksch A, et al. 2012. GiA Roots: software for the high throughput analysis of plant root system architecture. *BMC Plant Biology* **12**:116

Godin C, Caraglio Y. 1998. A multiscale model of plant topological structures. *Journal of Theoretical Biology* **191**

Le Bot J, Serra V, Fabre J, Draye X, Adamowicz S, Pagès L. 2010. DART: a software to analyse root system architecture and development from captured images. *Plant and Soil* **326**:261-273

Lobet G, Pagès L, Draye X. 2011. A Novel Image Analysis Toolbox Enabling Quantitative Analysis of System Architecture. *Plant Physiology* **157**:29-39

Naeem A, French AP, Wells DM, Pridmore TP. 2011. High-throughput feature counting and measurement of roots. *Bioinformatics* **27**:1337-1338

Preuksakarn C. 2012. *Reconstructing plant architecture from 3D laser scanner data*. Ph.D. thesis, University of Montpellier 2, France, 126 p.

WinRhizo, Pro, 2012b. http://www.regent.qc.ca/assets/winrhizo_software.html *Regent Instruments Inc., Canada*.

Zeng, G., Birchfield, S.T., Wells, C.E. 2008. Rapid automated detection of roots in minirhizotron images. *Machine Vision and Applications* **21:**309–317.