

Washington University in St. Louis

Washington University Open Scholarship

All Computer Science and Engineering
Research

Computer Science and Engineering

Report Number:

2019-4

Feature Extraction form CT Scan of Plant Root

Chunyuan Li

Roots are vital for plant by absorbing water and nutrients and providing anchorage from beneath the soil. These roles are closely related to the roots' architecture, which describes the geometry of individual roots and their branching structure. We proposed a pipeline to efficiently annotate root architecture. My contribution focus on building an interactive tool to visual and annotate root architecture. Besides, we come up with heuristics to automate the annotation process.

Follow this and additional works at: https://openscholarship.wustl.edu/cse_research



Part of the [Computer Engineering Commons](#), and the [Computer Sciences Commons](#)

Recommended Citation

Li, Chunyuan, "Feature Extraction form CT Scan of Plant Root" Report Number: (2019). *All Computer Science and Engineering Research*.

https://openscholarship.wustl.edu/cse_research/1175

Department of Computer Science & Engineering - Washington University in St. Louis
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

Feature Extraction form CT Scan of Plant Root

Chunyuan Li

Abstract—Roots are vital for plant by absorbing water and nutrients and providing anchorage from beneath the soil. These roles are closely related to the roots architecture, which describes the geometry of individual roots and their branching structure. We proposed a pipeline to efficiently annotate root architecture. My contribution focus on building an interactive tool to visual and annotate root architecture. Besides, we come up with heuristics to automate the annotation process.

I. INTRODUCTION

Roots are important for plants for which they grow downward and hold the plant in place. They absorb water and minerals for use by the plant and store food reserves. Root architecture is an important feature to describe and understand root system. It includes shape of individual roots, their branching hierarchy, and time of appearance. Being able to quantitatively describe the architecture of a root system is a prerequisite for statistical and computational study of roots. For example, this will benefit discovering relationship between genetic basis and root traits. Biologists can make use of such relationship and breeding crops with favorable traits, therefore increasing crop yield while reducing adverse environmental effects. Different species have different root architecture. Our work focusing on corn root, while it can be extended to understand architecture of other species as well.

The root architecture consists of a stem in the center (also called primary root, see figure 1. The nodes locate on the stem where there are multiple branches reach out. The branching hierarchy composed of the brace root from node above ground, crown root, seminal root which grown from the seed, etc. Those branches that originate from stem are grouped as primary branches. The lateral roots which branch out from primary branches are noted as secondary branches.

The objective of this project is to identify root traits from CT scan. Our original input is an X-ray CT scan of a root (image slices). A common process is to apply An intensity threshold on the 3D image such that voxels whose intensity is above the threshold are considered to be part of the segmented shape (segmentation). Then a persistence-based method is applied to remove topological noises and a graphbased algorithm is used to compute a curve skeleton skeleton. We use the curve skeleton to annotate root structure. A pipeline for the entire work is shown in figure 2.

The root traits we are interested including Overall traits, Architectural traits, Geometric traits. Overall traits include volumes (number of voxels), depth, convex hull size, etc. Architectural traits are branching hierarchy, types, number and species-specific features e.g. nodes for corn root vs nodules for soybean. Geometric traits like branch length, tortuosity, branching angle, etc.

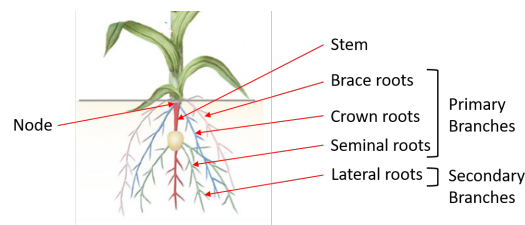


Fig. 1. Architectural traits of corn root. Root illustration from [1]

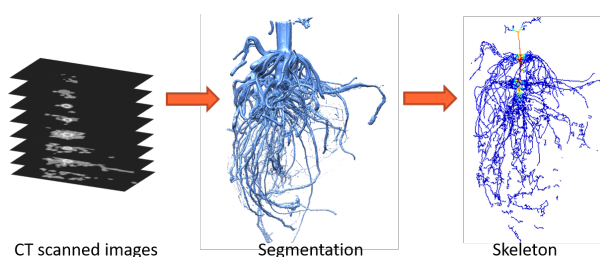


Fig. 2. Pipeline from CT scanned of plant root to skeleton.

Overall traits are easy to obtain directly from segmentation, while architectural or geometric traits requires understanding of the structure and shape of the segmentation. This is challenging because segmentations are often imperfect, which causing erroneous structures or shape. Besides, the root system is highly complex. It is difficult to annotate root traits in both CT scanned images and segmentations. In scanned image volume, it is time consuming to identify the root traits by looking at each image slices. In the segmentation, it's inefficient to annotate thousands of voxels of each root traits.

To solve the problem, we convert segmentation to skeleton and use it as the input to our tool for traits identification. The thinned 1d representation (graph representation) of 3D shape, described by vertices and edges, is far less complex than segmentation and thus reduced effort for annotating.

Our contribution for this project is to develop a software for interactive and automatic root traits analysis. We take in a segmentation and geometric skeleton of root. Example of the input data is shown in figure 3 left. The blue shape with transparency is the segmentation, and the red edges represents the skeleton. Our work is operated on skeleton while we use segmentation as a reference to quickly identify traits. My work focus on creates annotations of the skeleton (branches, types, etc.) and develop heuristics to compute root traits.

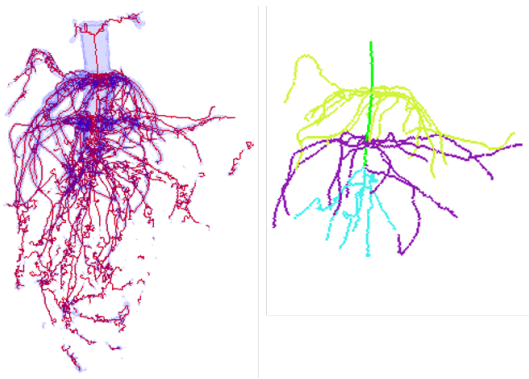


Fig. 3. Left: Example of input skeleton and segmentation. Right: Annotated root structure.

II. PRIOR WORKS

There are a few semiautomated methods to describe root structure by skeletonizing. RootReader2D [2] draws 2D root skeleton image using interactively selected root endpoints and stored shortest path information. SmartRoot, developed by Lobet et al. [3], enables quantitative analysis of root system architecture. Such techniques using simple heuristics are sensitive to segmentation errors and tend to invite incorrect branching structures. Besides, user intervention is required to infer individual roots, their hierarchy and types. As far as we know, there is no automated solution capable of producing a complete representation of root architecture.

III. SKELETON-BASED ROOT ANALYSIS

A. Interactive Tool

We designed graphical tool for modeling 3D root architecture from segmented images. This tool aims to visualize the root structure and allow interactive operation on the architecture. It takes in a segmentation and geometric skeleton as input.

In editing aspect, the tool can clean up the topological errors of the skeleton such as loops and disconnections.

To identify root architecture, we start from annotating the stem, then primary nodes on the stem and finally, the primary branches. The operation is simple and straight forward, which requires around twenty minutes to annotate entire root structure for a six week corn root like in figure 3.

B. Automatic Algorithm

We first identify the stem (Figure 4) using the radius measure along the medial axis, which is the closest distance to the surface for each skeleton vertex. The basic idea is to find the longest path from a group of vertices that have high radius. Detailed procedure is listed as follow.

- 1) The stem is the thickest region of the root, so naturally we start finding the stem from the point of the skeleton with the highest radius. This point, corresponds to the thickest part of the stem, is chosen as a seed point. See the yellow point in figure 4 (a).
- 2) Since the stem is thicker than primary roots, a lower threshold on radius can be used to determine a possible

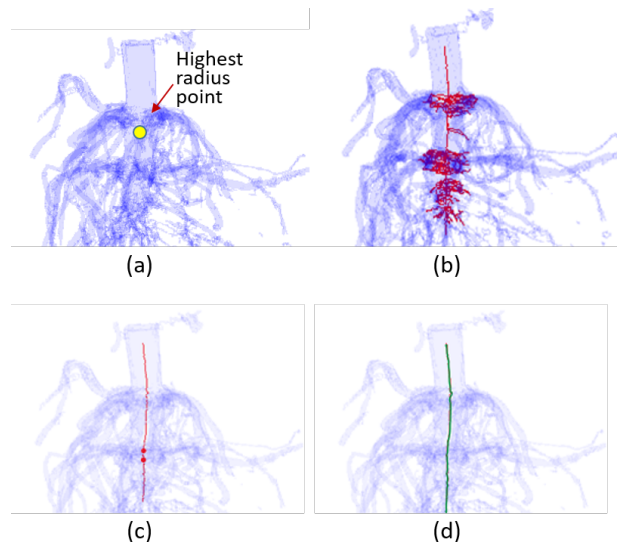


Fig. 4. Stem Identification. (a) The algorithm start from the vertex with the highest radius. (b) A region of the skeleton is expanded to include a neighborhood around the stem. (c) The region is shrunk to a single point or edge while each vertices burn time is recorded. (d) Points are added back two at a time to produce a high-radius chain of skeleton vertices.

region for the stem. Given a low radius threshold T_l , neighboring vertices are added to a set S over several iterations until a reached vertex falls below T_l . Figure 4 (b) shows a example of resulting regions. This step make sure S contains all vertices of the stem but not too many of the neighboring points.

- 3) The next task is to find a subset of S containing the stem. Kruskals algorithm is applied on S to obtain a minimum spanning tree S_{mst} which is devoid of cycles. Then, several iterations of endpoint removal are applied to S_{mst} until only a single point or a single edge are left - see figure 4 (c). During each iteration, all current endpoints of S_{mst} are removed, and the iteration number is recorded to be the burn time for each removed vertex.
- 4) The set S_{stem} , the final estimation of the stem, is initialized with either the single point or two points remaining in S_{mst} . For each endpoint in S_{mst} , the neighboring point from S with the highest radius is added to S_{mst} . This addition is performed until the burn time of an added point is 0, or equivalently when one of the endpoints of S is reached. The resulting S_{stem} is a chain of high radius vertices and edges which approximate the main stem.

Based on identified stem, the primary nodes are estimated using the idea the mean shift clustering. Mean shift is a hill climbing algorithm which involves shifting node position iteratively to a higher density region until convergence. The algorithm is detailed next.

- 1) As it shown in figure 5 (a), we start from an identified stem and a list of junction points on the stem.
- 2) For each junction point, we find a set of neighboring points within a look distance.

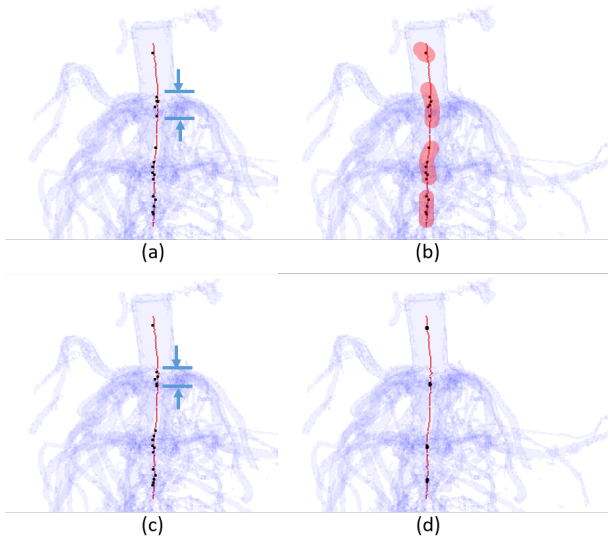


Fig. 5. Node Identification. (a) Start with a set of interaction points on the stem. (b) For each vertex, find a group of neighboring vertices within predefined range. (c) Update position of each vertex by the weighted mean of neighboring points. Go to step (b). (c) When the position of each vertex no longer change or change within tiny range, terminate the algorithm.

- 3) Then updating the position of each point by shifting it to the weighted mean of the neighboring points. The weight is calculated by a Gaussian kernel, where we use bandwidth to defined shape of the Gaussian kernel. The updated vertices tend to converge to dense region, see figure 5 (c).
- 4) Repeat the updating process until the position no longer change or change within a small range. Figure 5 (d) shows the example of identified nodes of a six week corn root. The final clustering result doesn't reduce number of vertex. the grouped vertices converge to the same location, which looks like a single vertex.

IV. EXPERIMENTAL RESULT

Due to the complexity of the root skeleton, the resulting image in figure 6 only shows identified stem on trimmed skeleton. From the figure 6, the estimated stem (in white) are very similar to ground truth (in green). The only difference is on the top the stem, identified stem include partial branch. The branches that grow from top of the stem are very thick and the radius are very high. Therefore, the stem algorithm cannot distinguish the upper endpoint of the stem. It takes several seconds to compute stem depends on age of the root. On the flowering root, the most complicated root that contains thousands of vertices and edges, stem identification costs approximate three second.

Figure 7 shows two example of node identification on four week root (left) and six week root (right). In each age group, the first image shows ground truth primary node with arrow highlight the location of the nodes. The red points indicate that there are branches grow out. The middle one is identified node marked by cubes. The third image contains partial branches around stem. On average, our algorithm identifies

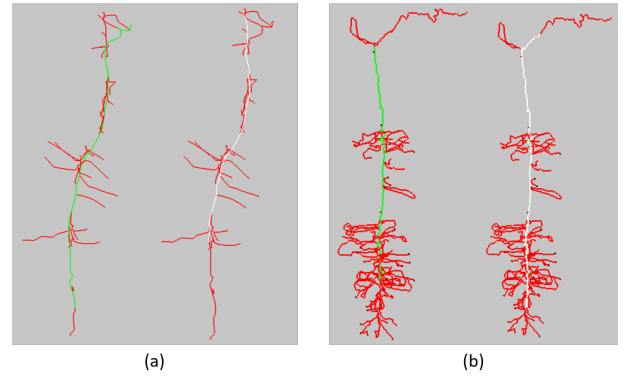


Fig. 6. Result of automatic stem identification. (a) a four week corn root. (b) six week corn root. In each age group, the ground truth stem is colored in green on the left and the automatically identifies stem is colored in white on the right.

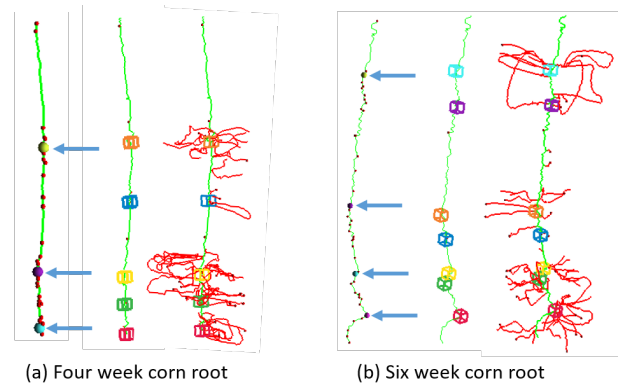


Fig. 7. Result of automatic node identification on four week and six week corn root. In each age group, ground truth primary nodes are shown on the left, where the red points represent interaction on the stem and the larger points indicated by arrows are primary nodes. The middle one is our identification result, where the nodes are marked by cubes. The right image shows branches around stem, which helps to analyze algorithm performance.

2.7 nodes more than ground truth. The euclidean distance between identified nodes and corresponding nearest ground truth nodes is less than 0.5 distance unit by average. (The skeleton is converted from segmentation, which composed of voxels. One distance unit in skeleton correspond to length of one voxel in segmentation.)

The imperfectness of node identification mainly caused by loops and noise. (1) When the branches grow downward and touch the stem, it forms a intersection. Since the clustering algorithm is based on intersections, it might be identified as a primary nodes. A example of this case is the second node in purple on Figure 7 (b). We can solve this problem by using existing loop information to exclude fake intersection on the stem. (2) The noise originate from the transformation of segmentation to skeleton, which introduce perturbation edges on stem. They are not actual branches. The current algorithm cannot distinguish real branches and unreal edges. These unreal edges are very short and lies within the radius range. Therefore, removing such edges by comparing edge length with radius will improve node identification process.

V. CONCLUSION

In this project, we built interactive tool to annotate root structures and developed heuristics to automatically identify stem and node. The node identification can be improved according to analysis in above section. In addition, we will automate primary branches identification. Our software will be freely distributed upon completion.

APPENDIX

Figure 8 shows the user interface. A video demo of manual process is available [here](#)

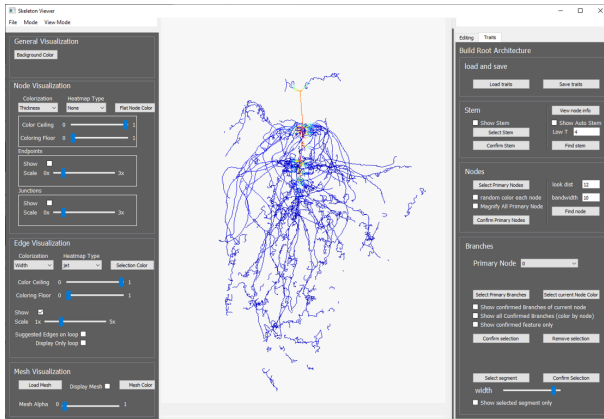


Fig. 8. User interface. Left: visualization options - controls dataset loading, color mode option, size/thickness of vertices/edges etc.; center: show root skeleton and segmentation; right: editing operation to clean topological errors on the skeleton and architecture annotation operation to build root structure.

ACKNOWLEDGMENT

I send my special thanks to Professor Tao Ju for his guidance. I would like to thank Dan Zeng and other lab members.

REFERENCES

- [1] Fangping Gong, Xiaolin Wu, Huiyong Zhang, Yanhui Chen, and Wei Wang, "Making better maize plants for sustainable grain production in a changing climate," *Frontiers in Plant Science*, vol. 6, pp. 835, 2015.
- [2] Randy T Clark, Adam N Famoso, Keyan Zhao, Jon E Shaff, Eric J Craft, Carlos D Bustamante, Susan R McCouch, Daniel J Aneshansley, and Leon V Kochian, "High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development," *Plant, cell & environment*, vol. 36, no. 2, pp. 454–466, 2013.
- [3] Guillaume Lobet, Loïc Pagès, and Xavier Draye, "A novel image-analysis toolbox enabling quantitative analysis of root system architecture," *Plant physiology*, vol. 157, no. 1, pp. 29–39, 2011.