DIRT/μ: Automatic root hair measurement in maize (*Zea mays ssp.*) from microscopy images

Peter Pietrzyk[a], Neen Phan-Udom[b], Chartinun Chutoe[b], Patompong Saengwilai[b], Alexander Bucksch[a,c,d]

[a]Department of Plant Biology, University of Georgia, Athens, GA, USA; [b]Department of Biology, Faculty of Science, Mahidol University, Thailand; [c]Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, USA; [d]Institute of Bioinformatics, University of Georgia, Athens, GA, USA

ABSTRACT

Improving nutrient and water uptake in crops is one of the major challenges to sustain a fast-growing population that faces increasingly nutrient-limited soils. Root hairs, which are specialized epidermal cells, are important drivers of nutrient and water uptake from the soil. Microscopy provides a mean to record root hairs as digital images. However, due to their geometry and complex spatial arrangements quantifying root hairs in microscopy images manually remains a bottleneck. Manual selection of representative root hairs can result in inaccurate estimations of root hair traits and misrepresentation of root hair functions. We present a method to quantify phenotypes automatically by measuring all individual root hairs in digital microscopy images. Our method uses random forests classification to separate root hair from the parent root and the image background. We define metrics to evaluate segments of root hairs that intersect or form blobs of two or more root hairs. Using simulated annealing for combinatorial optimization, we reconstruct individual root hairs by resolving intersections in a globally optimal way. As a result, we measure the root hair length, its distribution, and root hair density in each image. We demonstrate our method on examples of three maize cultivars under phosphorus, nitrogen, and potassium stress. Results show that our measurements of root hair traits strongly correlate with manually measured data in mean root hair length (R^2: 0.72 to 0.85, p<.001), as well as in root hair density (R^2: 0.38 to 0.66, p<.001). We show that our method computes reliable estimates of root hair length, density and their distributions along the root on complex root hair arrangements in maize. We believe that our study paves a way towards identifying the genetic control of root hair traits and increased agricultural production.

Keywords: root hair, abiotic stress, phenotyping, machine learning, simulated annealing, trait distribution

1. INTRODUCTION

Branching patterns in biology occur at all spatial scales, organismal levels and for many physiological, protective or reproductive reasons and are most prominent in the plant kingdom [1]. Hair-like structures with a high length to width ratio are specific objects that branch off and extend from the organism's surface and have diverse functions across biology [2]. To study their function it is necessary to quantify their shape and arrangement accurately, which, despite of their simple shape—in comparison to a multi-level branching architecture—remains a challenge. Modern imaging tools can capture digital images of these hair-like structures, but extracting all of them individually from the image is ambiguous if they occlude each other partially. The occlusion is especially prevalent in root hairs, which are elongated epidermal cells extending from the root surface of a plant. By increasing the root surface area and extending away from the root surface into the soil, root hairs can increase water and nutrient uptake from the soil [3].

We demonstrate an algorithm to resolve occlusion in 2D microscopy images of root hairs. Researcher studying morphological traits of root hairs traditionally use microscopy images and quantify root hair density and length manually. Not only is this manual trait measurement of root hair traits extremely tedious

but the interpretation of 2D images representing root hairs is also subjective. If measurements are done automatically, only the total area or a profile of root hair length along the root can be extracted [4, 5]. Other studies used X-ray computed tomography (CT) to scan roots and root hairs in soil, but still used manual tracing to segment root hairs from 3D X-ray CT scans [6-8].

The challenge to extract individual root hairs from microscopy images is that all intersections of root hairs must be resolved in the 2D projection of the image. A single intersection of two root hairs can be instantly resolved by determining the straightest solution from a small number of possible combinations. By increasing the number of root hairs and intersections, however, the number of potential combinations increases in real scenarios to trillions of possible outcomes. We present an approach to strategically resolve intersections and extract individual root hairs in feasible computational time. As such, our approach allows to measure root hair traits, like length and density, and their distributions within a root sample at a much finer resolution than previous 2D approaches.

2. MATERIAL AND METHODS

2.1 Datasets

To validate our method we selected fifteen microscopy images of rice, maize and common bean roots taken at different root classes and grown under different treatments. All roots grew in a hydroponic system and were stained with toluidine blue before image acquisition. For each image, we counted and traced all individual root hairs to be used as a validation set. To demonstrate the usability of our method we used a dataset of three Thai inbred maize lines: Tak Fa 1, Tak Fa 2 and Tak Fa 3. For each genotype three to four replicates were grown in control, reduced nitrogen, reduced phosphorus, and reduced potassium conditions, respectively. For each replicate, five to seven images were taken along the root; for each image, five representative root hairs were traced manually to measure their length and root hair density was measured along a representative area of the root.

2.2 Image classification

We used the software ilastik [9] to classify all image pixels as either root hair, root or background. The software provides a user interface to create training sets by manually labelling pixels according to their class and to look interactively at the results. For the classification, we used a random forest classifier with features obtained from morphological filters (validation dataset: σ=0.3-10; demonstration dataset: 0.3-50). In the validation dataset, each image was trained and classified separately, while for the demonstration dataset we classified each genotype separately using ten to eleven images from the corresponding genotype.

2.3 Root hair extraction

Based on the classified image we extracted the medial axis from pixels, which were classified as root hair. We then identified all termination points, which have only one neighboring pixel on the medial axis, and junction points, which have more than two neighboring pixels on the medial axis. We connected all pairs of these points, if they were within a certain proximity, by traversing paths along the medial axis. This resulted in a set with paths that either do or do not resemble a segment of a root hair correctly. To smooth out artifacts of the medial axis and to resemble actual root hairs more accurately we fitted a weighted spline to each path with weights calculated by Equation 1, where d_{MA} is the measured diameter of the medial axis and d_{RH} is an estimate of the expected root hair diameter.

$$w = 1/\max(1.0, d_{MA} - d_{RH}) \quad (1)$$

We then used simulated annealing to combinatorically find the set of splines that resolves intersections and find root hairs in a globally optimal way [10]. At each iteration of the annealing process, either a spline is removed or a new spline is added to the set. During the annealing process we try to minimize a predefined cost function, which we calculate as the weighted root mean square of three metrics: (1) The average of the smallest distances of all root hairs to the root, (2) the residual strain energy, i.e. a metric to reduce the

curvature of splines and (3) the fraction of remaining root hairs based on unresolved medial axis segments. At each iteration we restore the feasibility of the solution by ensuring that no spline branches of another spline and all intersections of splines are valid [11]. Two splines overlapping at one end are merged to a single entity to ensure that longer root hairs can be extracted in their full length. The annealing process stops once the cost function converges to a minimum and no more changes to the set are accepted.

2.4 Root hair traits

As a result, of the root hair extraction, we were able to compute traits based on the shape and distribution of individual root hairs as shown in Figure 1. As such, we calculated the length of each individual root hair and the density of root hairs on both sides of the root as the number of root hairs per root edge length. Root hairs that were too short, too far from the root or that emerged from a secondary root were not included into the calculations. We excluded images where the algorithm failed to identify the root edge correctly.

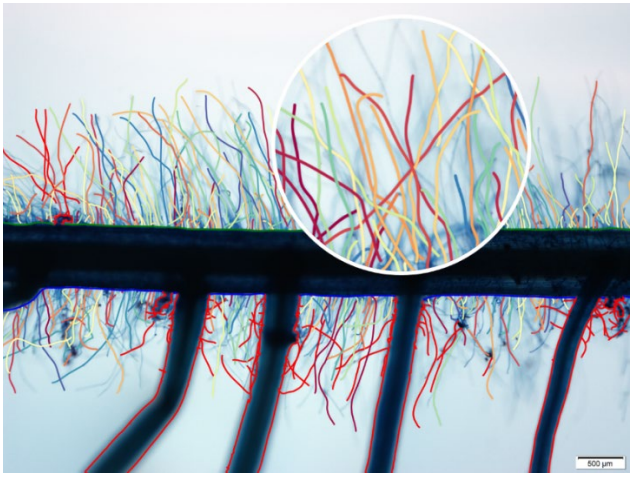

Figure 1: Extracted root hair and edges of root are shown as colored lines. Root hair on main root are shown in random colors, other root hairs in red. The root's top edge (green), bottom edge (blue), and edges of secondary roots and noise (red) are classified.

3. RESULTS

3.1 Validation

To validate our method we extracted root hairs in all fifteen images of the validation data set and compared the results to the manual measurements. As illustrated in Figure 2, the distributions of root hair length obtained with the automatic approach matches the distributions of manually extracted root hair in all fifteen images. The analysis of the results showed a strong correlation between automatic and manual results in mean root hair length (R^2=0.93, p<.001), maximum root hair length (R^2=0.90, p<.001) and number of root hairs (R^2=0.63, p<.001).

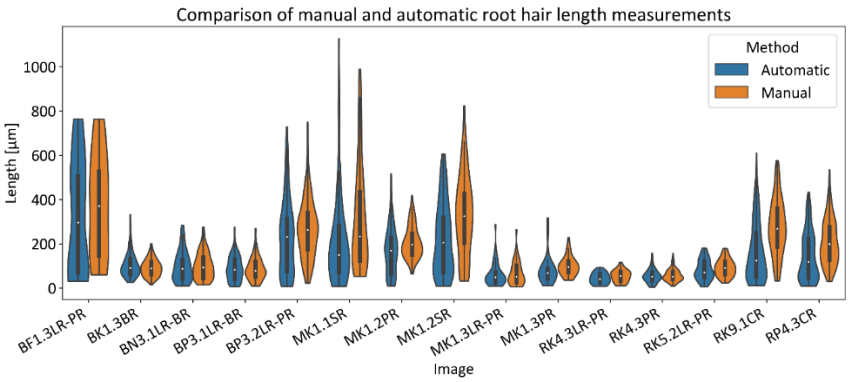

Figure 2: The distribution of root hair length extracted with our approach in comparison to manual measurements for 15 images. Automatic measurements are shown in blue and manual in orange.

3.2 Root hair traits in maize

We further computed the root hair length and density for Tak Fa 1, Tak Fa 2 and Tak Fa 3 under all four treatments (Figure 3, columns a,b,c). In each replicate, which had more than one image, we calculated the mean root hair length as the mean of all root hairs from all images and the density as the total number of root hairs from all images divided by the total length of the root edge from all images. Despite root hairs in the Tak Fa 3 control group being on average 384.1µm long and thus, 73.7µm and 58.0µm longer than Tak Fa 1 control and Tak Fa 2 control, respectively, the difference was not statistically different. We further did not observe significant differences in density between all three genotypes, despite Tak Fa 2 being 69% and Tak Fa 3 44% denser than Tak Fa 1. We did not observe any significant differences in density between all treatments to the control group in Tak Fa 1 and Tak Fa 3. Only in Tak Fa 2 we observed a decrease in root hair length by 112.0µm under reduced nitrogen with weak significance (p=0.057).

To show how our results compare to standard manual root hair measurements, for each image we correlated the mean root hair lengths of all automatically extracted root hairs to the mean of the five representative root hairs (Figure 3, column d). For all three genotypes we obtained high correlations (Tak Fa 1: R^2=0.78, p<.001; Tak Fa 2: R^2=0.85, p<.001; Tak Fa 3: R^2=0.72, p<.001). Similarly, we correlated root hair density from the automatic method to manual measurement and obtained strong correlations as shown in Figure 3, column e (Tak Fa 1: R^2=0.66, p<.001; Tak Fa 2: R^2=0.46, p<.001; Tak Fa 3: R^2=0.38, p<.001).

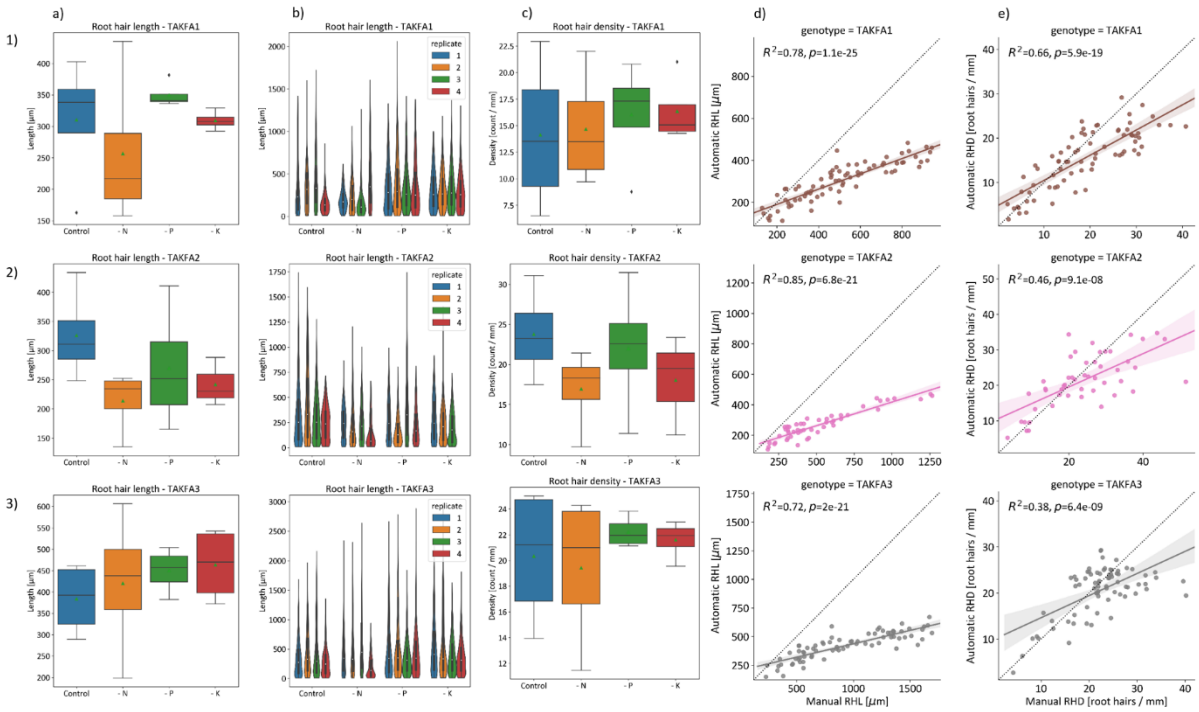

Figure 3: Overview of results. Results for Tak Fa 1, Tak Fa 2 and Tak Fa 3 are shown in rows 1), 2) and 3). Column a) shows mean root hair length by treatment; column b) shows distribution of root hair length by replicate and treatment; column c) shows root hair density by treatment; columns d) and e) show correlations between automatic and manual measurements of root hair length and density, respectively.

4. CONCLUSIONS

We demonstrated an algorithm to resolve and extract occluded hair-like structures on the example of root hairs captured in 2D microscopy images. The high correspondence of our measurements to the validation data shows that we are able to accurately determine the length and density of root hairs in microscopy images. We observed that correlations between automatic and manual measurement are lower in the demonstration dataset with maize images than in our validation dataset. Higher correlations in our validation set suggests that improving our classification step could result in overall enhanced performance and that manual measurements of root hairs are potentially not representative. We further observed that manually measured root hair length resulted in longer root hairs compared to automatically measured root hairs, which suggests that selecting root hairs manually is biased towards longer root hairs. We tested for difference between groups in root hair density and mean root hair length per replicate (i.e. per plant). While we could not determine significant differences between genotypes and treatments in this dataset, we believe that our method paves the way towards better identifying the genetic control of root hair traits and an improvement in breeding programs for these traits. We anticipate that incorporation of the measured trait distributions (Figure 3, column b) into future statistical analysis will provide more insight into root hair response to abiotic stresses. Further development of our algorithm will allow to resolve intersections in branching architectures with several orders of branching hierarchy to accurately determine traits at larger organismal levels in the future.

DATA AVAILABILITY STATEMENT

The code for extracting root hairs and a sample image are available on GitHub: https://github.com/Computational-Plant-Science/DIRTmu.

ACKNOWLEDGMENTS

The research was supported by the NSF CAREER Award No. *1845760 and* USDOE ARPA-E ROOTS Award Number DE-AR0000821 to A.B. Any Opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the funders.

REFERENCES

[1] Bucksch, A., "A practical introduction to skeletons for the plant sciences," Applications in Plant Sciences 2(8), (2014).

[2] Seale, M., Cummins, C., Viola, I. M., Mastropaolo, E., and Nakayama, N., "Design principles of hair-like structures as biological machines," Journal of the Royal Society Interface 15(142), (2018).

[3] Gilroy, S., and Jones, D. L., "Through form to function: root hair development and nutrient uptake," Trends in Plant Science 5(2), 56-60 (2000).

[4] Vincent, C., Rowland, D., Na, C., and Schaffer, B., "A high-throughput method to quantify root hair area in digital images taken in situ," Plant and Soil 412(1-2), 61-80 (2017).

[5] Guichard, M., Allain, J. M., Bianchi, M. W., and Frachisse, J. M., "Root Hair Sizer: an algorithm for high throughput recovery of different root hair and root developmental parameters," Plant Methods 15(1), (2019).

[6] Keyes, S. D., Daly, K. R., Gostling, N. J., Jones, D. L., Talboys, P., Pinzer, B. R., Boardman, R., Sinclair, I., Marchant, A., and Roose, T., "High resolution synchrotron imaging of wheat root hairs growing in soil and image based modelling of phosphate uptake," New Phytologist 198(4), 1023-1029 (2013).

[7] Daly, K. R., Keyes, S. D., Masum, S., and Roose, T., "Image-based modelling of nutrient movement in and around the rhizosphere," Journal of Experimental Botany 67(4), 1059-1070 (2016).

[8] Koebernick, N., Daly, K. R., Keyes, S. D., George, T. S., Brown, L. K., Raffan, A., Cooper, L. J., Naveed, M., Bengough, A. G., Sinclair, I., Hallett, P. D., and Roose, T., "High-resolution synchrotron imaging shows that root hairs influence rhizosphere soil structure formation," New Phytologist 216(1), 124-135 (2017).

[9] Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A., and Kreshuk, A., "ilastik: interactive machine learning for (bio)image analysis," Nature Methods 16(12), 1226-1232 (2019).

[10] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., "Optimization by simulated annealing," Science 220(4598), 671-680 (1983).

[11] Abramson, D., Dang, H., and Krishnamoorthy, M., "A comparison of two methods for solving 0-1 integer programs using a general purpose simulated annealing algorithm," Annals of Operations Research 63, 129-150 (1996).