Root Segmentation Over Multiple Time Points

H. Allen, L. J. Cooper, G. Delius, M. T. Gallagher, T.D. Johnson,
F. Klimm, F. Randisi, T. Shearer and C. Ziegler

Introduction

Global food availability is a growing concern. In order to feed a growing global population it is essential that more food is produced [Keesstra et al., 2016]. All human food besides aquatic resources comes either directly or indirectly from plants which, in turn, gather their nutrients from soil via their roots. However, the majority of the soil that can be used for growing crops is already begin utilised and it is being eroded faster than new top soil can be formed. It is vital that crops take full advantage of resources in the soil to achieve the highest yield possible.

Fundamental to optimising plant growth, is the understanding of how root systems grow and adapt to the soil environment. One of the greatest challenges in this field is that plants grow in an opaque medium. This can be overcome thanks to advances in X-ray computer tomography (XCT) [Keyes et al., 2013]. These developments have made it possible to visualise root systems in three dimensions within the soil. This means that root traits, such as length density, tortuosity and transport characteristics, can be analysed *in situ*.

For an XCT scan of soil containing the roots of one plant, the soil appears as a heterogeneous mixture with high density minerals appearing white, organic material, water and clays appearing as various shades of grey and any air appears dark grey or black. The roots themselves are also a shade of grey, somewhere within the range for organic material, water and clays, but can be identified by eye due to edge effects and their long cylindrical geometry. This makes automatic segmentation difficult as traditional methods such as thresholding are not appropriate. Manual segmentation of the roots can be carried out by manually selecting the root on individual 2D slices from the 3D image stack and interpolating between the slices. This is time consuming and user dependent. and manual segmentation is not appropriate when a large number of samples need to be analysed, as is the case in biological studies, in a way that is not biased or user dependent.

Of particular interest is how plant roots adapt to their local environment, for example, to local sources of nutrients and soil structure. One approach to investigate this area is to XCT the same plant's root system over multiple time points under different conditions. This increases the number of scans that need to be analysed making manual method much less attractive. It also becomes necessary to correlate individual roots in each scan to a scan of the same plant at a different time point. The plant roots impact the surrounding soil both physically, by improving soil stability, and chemically [Barré and Hallett, 2009].

Twelve wheat plants in different soil conditions have been scanned at multiple time points [Ahmed et al., 2015]. A subset of this data was provided; 3 XCT reconstructions of a single wheat plant scanned at 6, 8 and 10 weeks. The aim was to design an algorithm to automatically segment the roots from the soil. This algorithm could then be applied to the whole data set. This will allows use to investigate the effect of the different soil conditions on the plant roots using an appropriate sized sample set to improve the statistical analysis and reduce the uncertainties of the biological population.

The three-dimensional data sets provided are challenging to segment automatically due to the fact that:

- the greyscale values of the features of interest (i.e. the plant roots) are very similar to those of the surrounding soil grains,

- the average greyscale value of the roots varies from slice to slice when the data set is represented by a stack of two-dimensional images,

- the roots are in contact with the soil grains in several locations.

To overcome these difficulties, we developed a methodology that automatically detects the roots in a slice based on the average greyscale value of the roots in the previous slice. Therefore, the only human input that may be needed is a manual segmentation of the top slice (which is often quick and straightforward). In fact, this step can often also be automated if the top slice is above soil level since there is sufficient contrast between the plant and air to allow the use of simple thresholding techniques. Once the top slice has been accurately segmented, the algorithm marches down through the stack image by image, detecting the roots in each slice. The alogrithm is described in detail below.

The algorithm

To illustrate how the algorithm works, we shall make reference to one particular stack of images, whose top, middle and bottom slices are shown in Figure 1a. These images are cropped versions of larger data sets, as shown in Figure 1b. The top slice is straightforward to segment; however, further down, identifying the roots is much more difficult and automatic segmentation cannot be easily achieved without some prior information. Therefore, we decided to use an algorithm that in a first step segments the top slice and then iteratively the slices below.

Most of the steps are straightforward and are available from various image analysis tools. We created implementations in Mathematica, Matlab, and ImageJ. Interestingly the results were very similar but not identical. Cause for this is most likely slight variants in the implementation of various algorithmic steps, e. g. the opening.
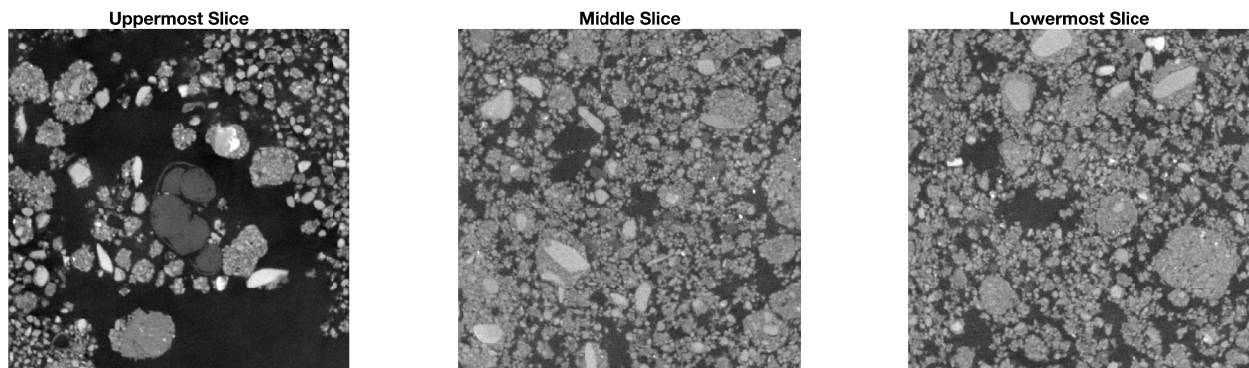
(a) Horizontal cross sections of the base of a wheat plant and its roots. The plant is the easily identifiable set of dark grey objects in the centre of the image in the top slice. As we move down through the stack of images, this object splits into many smaller objects (the roots). By the middle slice, identifying the roots is much more difficult and automatic segmentation cannot be easily achieved without some prior information.
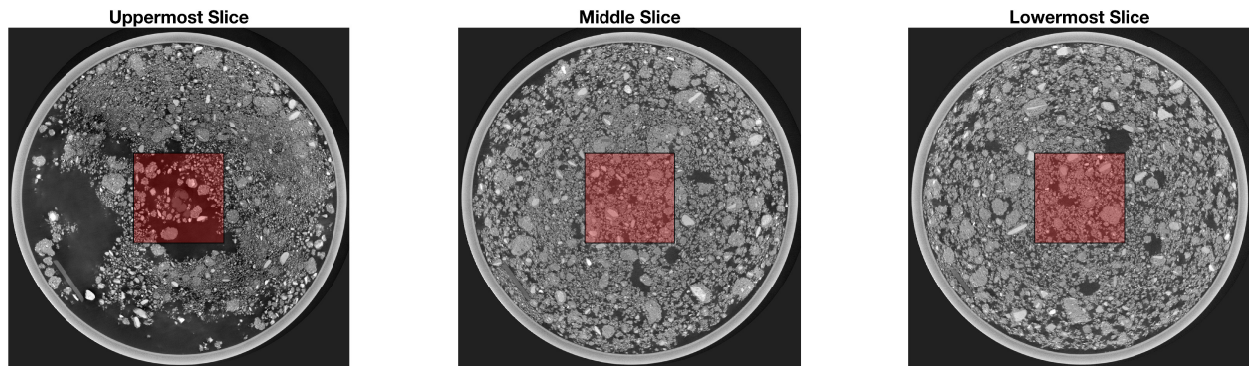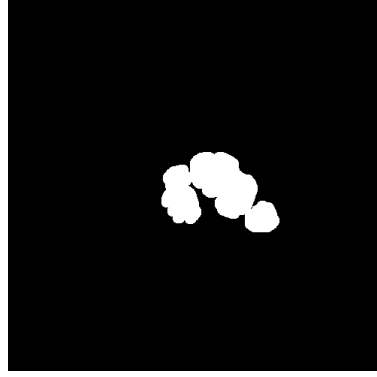

(b) The larger images from which the images in Figure 1a are taken.

Figure 1: Sample slices taken from an image stacks.
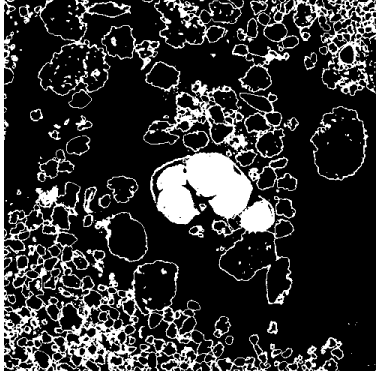
(a) Pixels from top slice of Figure 1a with a greyscale value ranging between 0.2 and 0.4 displayed as a binary image.

(b) A morphological opening using a disk matrix of radius 6, followed by a dilation of radius 1 were applied to the image in Figure 2a to remove the soil grain edges and to ensure that the edges of the plant were not excluded.

(c) The binary image from Figure 2b was multiplied by the original image to obtain this image.

Figure 2: Initial segmentation of top slice, steps 1.1 to 1.3.

Step 1: Segment the top slice

This step is the only step that requires human input. It can be achieved by any methodology that is satisfactory to the user; however, for the purposes of this report, the methodology that was used in this example is described in detail:

1. Firstly, a crude thresholding was applied to select all pixels with a greyscale value between 0.2 and 0.4 (where 0 corresponds to black and 1 to white - in an 8-bit image, this is equivalent to selecting pixels with greyscale values ranging between 51 and 102). The resulting binary image is shown in Figure 2a.

2. Next, a morphological opening using a disk matrix of radius 6 as a structuring element, followed by a dilation of 1 pixel were carried out to remove the soil grain edges and to ensure that the edges of the plant were not unintentionally excluded. The resulting image is shown in Figure 2b. This is a rough segmentation that forms the basis for the more precise segmentation described below.

3. The binary image from Figure 2b was then multiplied by the original image to obtain the image in Figure 2c.

4. The mean and standard deviation of the greyscale values of the non-black pixels from the image in Figure 2c were then calculated. In this case, the calculated mean, μ, and standard deviation, σ, of the greyscale values were
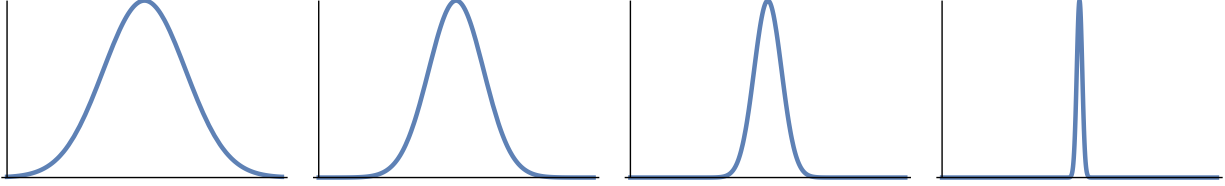
$$\mu = 0.31 \text{ and } \sigma = 0.05, \tag{1}$$

4

Figure 3: Pseudo-Gaussian mapping function plotted for $\mu = 0.5$ and, from left to right, $\sigma = 0.15$, $\sigma = 0.1$, $\sigma = 0.05$, $\sigma = 0.01$. For all of these plots, both the horizontal and vertical axes range from 0 to 1.

respectively.

5. The values calculated in step 4, above, were then used as the basis of an image transformation that took the original image and changed its greyscale values via a pseudo-Gaussian mapping using the following equation

$$y = e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{2}$$

where x is the greyscale value of a pixel in the original image and y is the greyscale value of the corresponding pixel in the transformed image. This equation is similar to a Gaussian probability distribution function, but always peaks at a value of 1. This function is plotted for $\mu = 0.5$ and several values of σ in Figure 3. The result of applying this transformation with the parameter values given by equation 1 is shown in Figure 4a. Figure 5 shows a histogram of the number of occurances of each greyscale level over the whole image in red, just the root region in blue and the applied pseudo-Gaussian mapping function in black.

6. Next, a median filter (over a range-5 neighbourhood) was applied to the transformed image, giving the image shown in Figure 4b. This reduced the greyscale value of the thin soil grain edges, whilst leaving a high greyscale value on the features of interest.

7. Since the features of interest are now of a much higher greyscale value than the rest of the image, simple thresholding can be used to select them. In Figure 4c, all pixels with greyscale value greater than 0.5 were selected.

8. Finally, another morphological opening was applied using a disk matrix of radius 2 as the structuring element in order to remove the small patches of noise. The resulting selection is shown in Figure 4d. This concludes step 1.

Substeps 1-8, above, are summarised in Figure 6.

Step 2: Dilate, then image multiply the next slice

Once the top slice has been successfully segmented, the resulting binary image is dilated using a radius 10 disk matrix as a structuing element. The dilated binary image is then
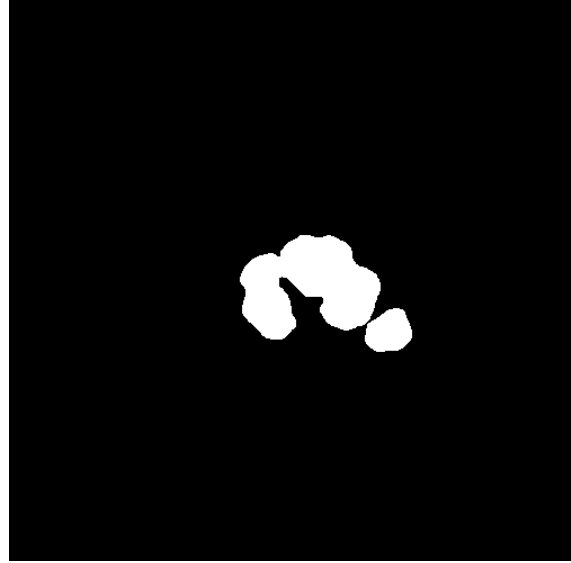
5
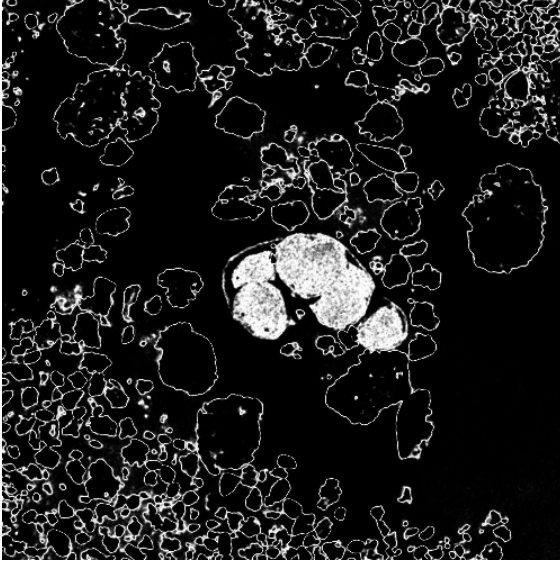
(a) The result of applying the transform given by equation (2) with the parameters given by equation (1) to the original image from Figure 1a.
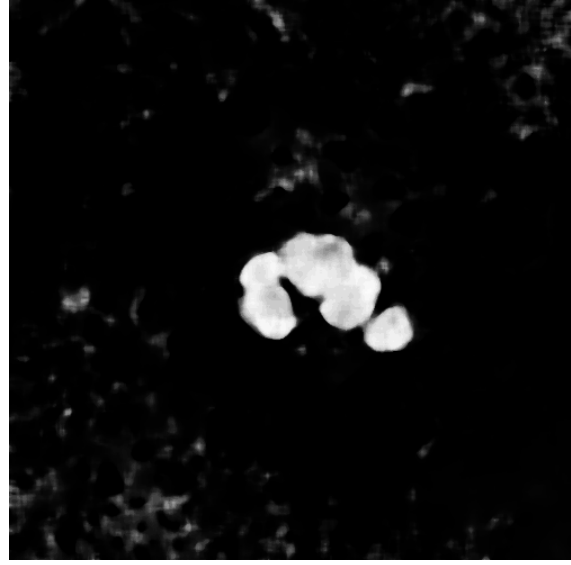
(b) The result of applying a median filter (over a range-5 neighbourhood) to the image from Figure 4a.

(c) A simple thresholding of the image from Figure 4b, in which all pixels with a greyscale value greater than 0.5 were selected.

(d) A morphological opening using a disk matrix of radius 2 as the structuring element, applied to the image in Figure 4c.

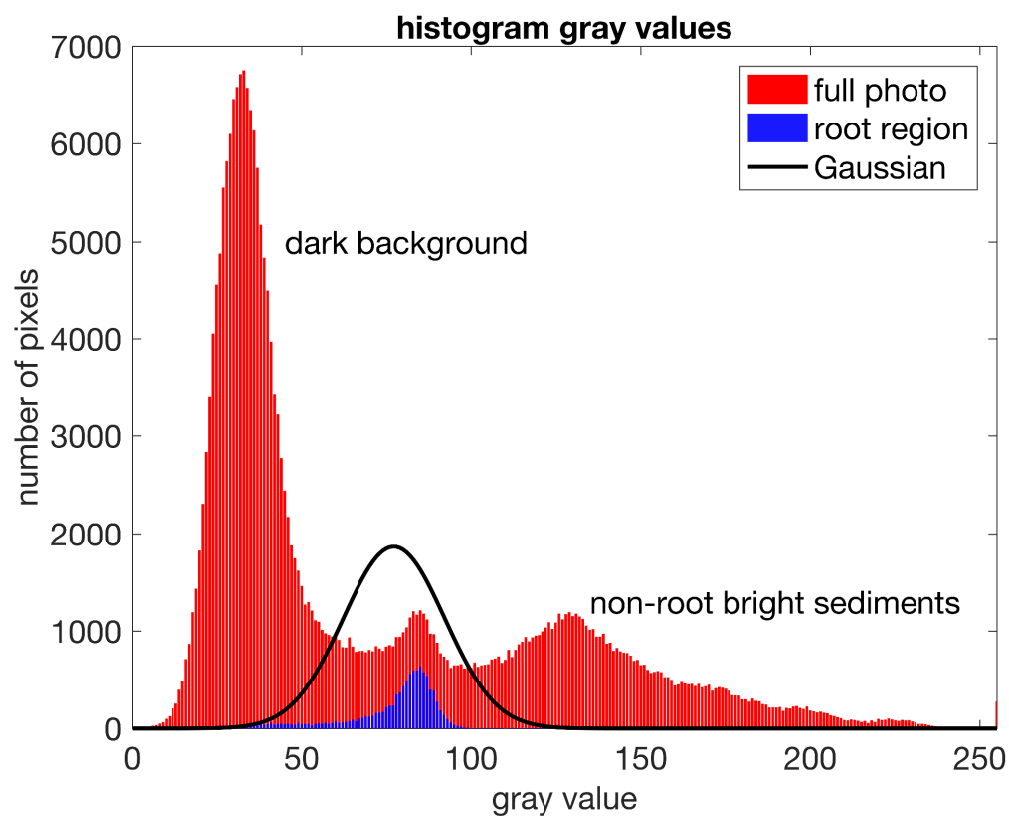Figure 4: Automatic segmentation of the first slice, steps 1.5 to 1.8.

Figure 5: A histogram of the number of occurances of each greyscale level in the top slice over the whole image (red), just the root region (blue) and the applied pseudo-Gaussian mapping function (black). Note that since this was an 8-bit image, the greyscale intensities are expressed as ranging between 0 (black) and 255 (white).
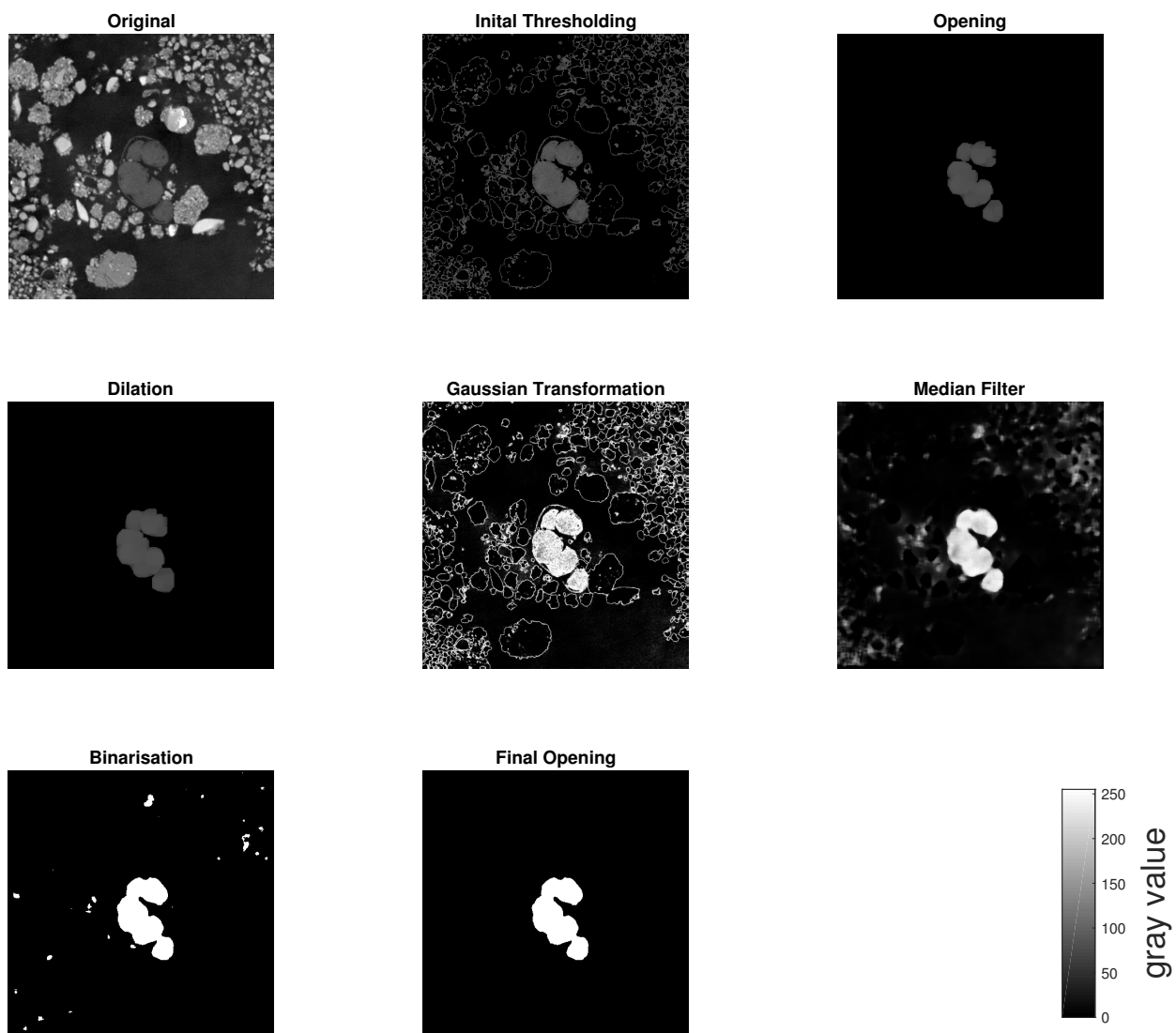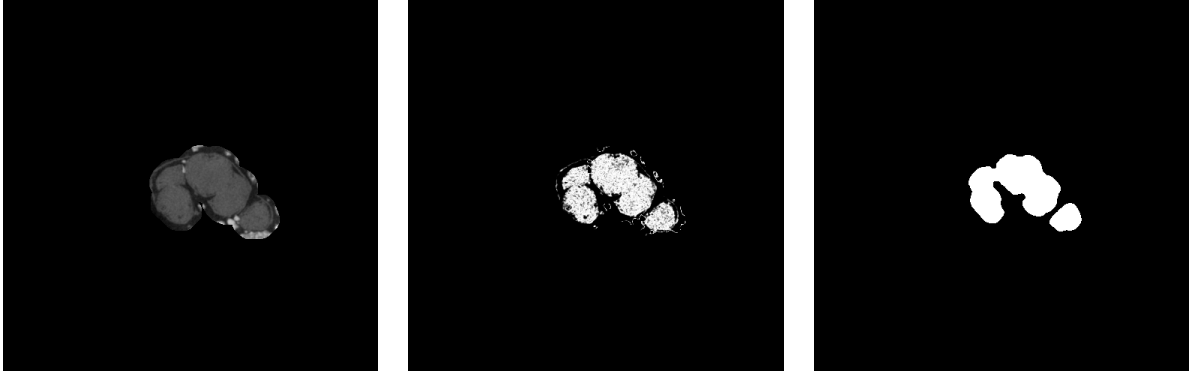
Figure 6: A summary of the image processing steps taken to segment the plant in the top slice.

(a) The result of dilating the image in Figure 4d using a radius 10 disk matrix as a structuring element, then image multiplying the resulting binary image by the next slice down.

(b) The result of applying the pixel greyscale remapping given by equation (2) to the image in Figure 7a using the mean and standard deviation calculated from the previous slice.

(c) The result of applying a median filter to Figure 7b over a range-5 neighbourhood, binarising using a threshold value of 0.4, then applying a morphological opening using a disk matrix of radius 2 as the structuring element.

Figure 7: Segmentation of slice 2, steps 2 and 3.

image multiplied with the next slice to determine a region of interest within which the roots are expected to lie. This method assumes that between one slice and the next, the roots will not have moved laterally by more than 10 pixels. The image resulting from this step is shown in Figure 7a.

Step 3: Remap the pixel greyscale levels

This step is the same as substep 5 from step 1, above, whereby the greyscale levels of the image are modified using the function given in equation (2); however, this time, the mean and standard deviation to input into the function are calculated from the previous slice. The result of applying this step to the image in Figure 7a is given in Figure 7b. We found that if the standard deviation calculated from a given slice gets too high, then this remapping erroneously maps soil particles to high greyscale levels and the algorithm therefore wrongly identifies some soil particles as being part of the root system. To overcome this difficulty, we capped the maximum allowed standard deviation at 0.025; therefore, if the calculated standard deviation of pixel greyscale values from slice i was higher than 0.025, then the value of σ used in equation (2) to transform slice $i + 1$ would be set to 0.025.

Figure 8: A three-dimensional visualisation of the three segmented root system.

Step 4: Apply a median filter, binarise, then open

This step is the same as substeps 6-8 from step 1, above. A median filter was applied over a range-5 neighbourhood, the image was binarised (this time the value 0.4 was used as the threshold greyscale level value), then a morphological opening was carried out using a disk matrix of radius 2 as the structuring element. The resulting image is shown in Figure 7c.

Step 5: Apply steps 2-4 to every subsequent slice

Finally, we repeat steps 2-4 for every other slice, each time using the immediately preceding slice as the input to step 2. The algorithm was applied to three different data sets (each representing the same plant at different stages of growth). The resulting three-dimensional volume representations of these root systems are shown in Figure 8.

Optional Step: Cleaning 3d detected roots

After the iterative application of the steps above the root structure can be detected but potentially also small fragments that, for example, represent organic material in the soil (see left panel in Fig. 9). Such small fragments have a much smaller volume, defined as the number of voxels, than the detected roots. Therefore we can apply an algorithm that detects and measures all connected components (`bwconncomp` in Matlab) and delete components below a certain size. This threshold should be chosen manually, in the example in Fig. 9 we deleted components with less than 120 voxels. If it is chosen to low to much organic material is still present and if it chosen to high parts of the root that are seemingly disconnected might be deleted, too.

Discussion

The algorithm contains a number of parameters that were selected either arbitrarily or via trial and error. These values worked well for the three data sets they were tested on during
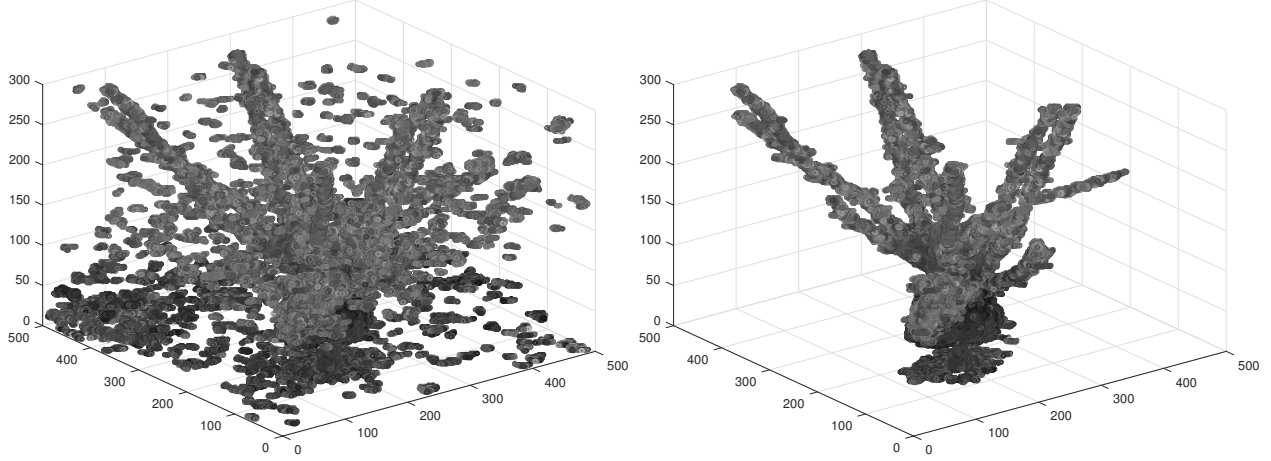
Figure 9: The cleaning of the detected root structure. Left: Uncleaned detected root structure. Right: Cleaned root structure after removal of small components.

the study group, but could be optimised to further improve the results. The relevant parameters are listed and discussed in Table 1. Using the values listed in the table, the algorithm presented is able to automatically segment a root system once the first slice has been segmented. It sometimes fails to detect smaller roots as the median filter and morphological opening steps result in smaller objects being removed. One possible extension would be to allow the parameters listed in the table to depend on depth and to reduce the size of the median filter neighbourhood and morphological opening structuring element matrix as the root diameters get smaller deeper into the soil. Further work is required to investigate the best way to do this.

One limitation of this methodology is that, if for any reason the algorithm loses track of a given root (for example if the binarisation threshold parameter was set too low to select a root after the median filter has been applied to a particular slice), it has no way of picking it back up unless it moves within 10 pixels of another root that is being tracked. To attempt to overcome this difficulty, we investigated removing step 2 from the algorithm. This results in the algorithm looking for roots in a slice globally based on the greyscale values of the roots in the previous slice. This approach allows us to pick up roots after they have been lost for a few slices, but also leads to some soil particles being selected (see Figure 10 and compare with the first image from Figure 8). One option would be to use this three-dimensional volume, remove the smaller soil particles by applying a three-dimensional morphological opening, then manually remove the larger soil particles (and those connected to the roots). The remaining root system could then be interpolated to fill in the gaps where the roots had been missed.

Besides the presented algorithm we also tested other approaches that were less successful. One attempt was to *analyse* the image slice-wise but rather as one three-dimensional array in a single step. Although this approach might seem more powerful on first sight, because it takes more information into account in a single step it was limited by mainly two factors:

11

Parameter	Step	Value	Discussion
Radius of disk matrix used for dilation	2	10	This parameter could be increased or decreased depending upon how much lateral movement is expected from slice to slice
Cap on the standard deviation input to equation (2)	3	0.025	We found that a relatively small value worked best for the data sets tested, but a different value may work better for other data sets.
Range neighbourhood of the median filter	4	5	This parameter should be varied depending upon the width of the soil grain edges that are highlighted in step 3.
Binarisation parameter	4	0.4	The algorithm is not very sensitive to this parameter, but optimsing it could potentially improve the results.
Radius of disk matrix used for opening	4	2	Should be kept as small as possible whilst still removing noise

Table 1: List of parameters used in the algorithm and the the values selected for the example data set.
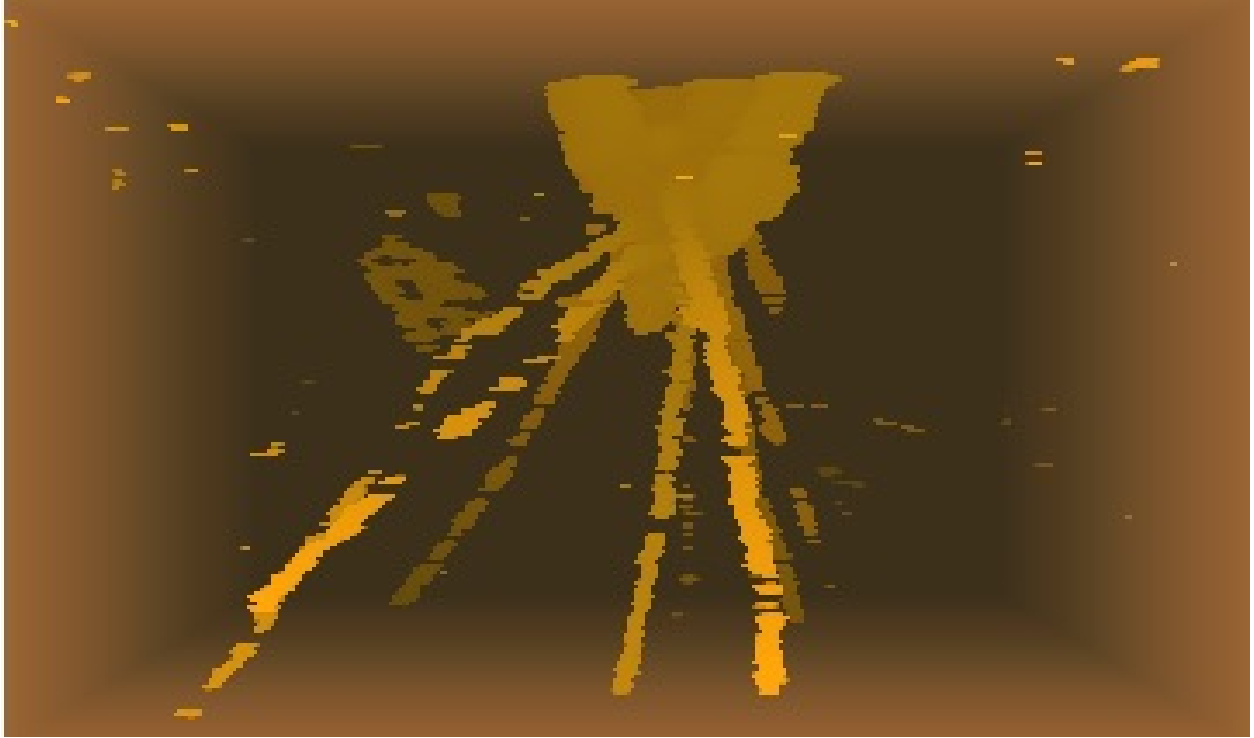
Figure 10: A three-dimensional visualisation of the first root system when step 2 is removed from the algorithm.

Firstly, the analysis of a large three-dimensional array is computationally more demanding as one that only computes on the information from two slices at a point in time. Therefore the performance of a slice-wise algorithm is better and the memory demand much lower. Note however, that out algorithm is not straightforward parallelisable since at each step the information from the slice before is needed. Secondly, in the practical work of image analysis a lot of fine-tuning of parameters under the inspections of the output is necessary. This procedure is much easier if the analysed slices can be inspected individually and not in one three-dimensional array. This could be overcome by a more advanced image-analysis pipeline but was not possible during the short time frame of the project.

Another algorithm was implemented in Matlab that allowed the user to click on a root within a 2D image slice. The algorithm then used the mean and standard deviation in the region of this initial selection to segment 40 slices below and above the image that was used as the initial selection. This worked well for small sections of individual roots and could be used as a method to fill the breaks within roots that can be seen in Figure 10.

All the codes developed during the workshop can be found at `https://github.com/gustavdelius/root_segmentation`.

Future Work

The algorithm will be validated by comparison with manual segmentation of the same scans. It can then be applied to the remaining scans in the original data set. This will enable the analysis of a statistically relevant sample size. The algorithm could also be validated for other plants, that may have morphologically different root structures. This will overcome the major bottleneck to investigating roots in soil.

References

Saskia D. Keesstra, Johan Bouma, Jakob Wallinga, Pablo Tittonell, Pete Smith, Artemi Cerdà, Luca Montanarella, John N. Quinton, Yakov Pachepsky, Wim H. van der Putten, Richard D. Bardgett, Simon Moolenaar, Gerben Mol, Boris Jansen, , and Louise O. Fresco. The significance of soils and soil science towards realization of the united nations sustainable development goals. *SOIL*, 2:111–128, 2016.

Samuel D. Keyes, Keith R. Daly, Neil J. Gostling, Davey L. Jones, Peter Talboys, Bernd R. Pinzer, Richard Boardman, Ian Sinclair, Alan Marchant, and Tiina Roose. High resolution synchrotron imaging of wheat root hairs growing in soil and image based modelling of phosphate uptake. *New Phytologist*, 198:1023–1029, 2013.

P. Barré and P. D. Hallett. Rheological stabilization of wet soils by model root and fungal exudates depends on clay mineralogy. *European Journal of Soil Science*, 60:525–538, 2009.

Sharif Ahmed, Trudy Naugler Klassen, Samuel Keyes, Michael Daly, David L. Jones, Mark N. Mavrogordato, Ian Sinclair, and Tiina Roose. Imaging the interaction of roots and phosphate fertiliser granules using 4d x-ray tomography. *Plant and Soil*, 401:125–134, 2015. doi: 10.1007/s11104-015-2425-5.