

A Visual Analytics Approach to Hypocotyl/Root Transition Detection in *Arabidopsis Thaliana*

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Julian Strohmayer

Matrikelnummer 01426125

an der Fakultät für Informatik
der Technischen Universität Wien

Betreuung: Viktor Vad, MSc.

Wien, 26. Dezember 2017

Julian Strohmayer

Viktor Vad

A Visual Analytics Approach to Hypocotyl/Root Transition Detection in *Arabidopsis Thaliana*

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Julian Strohmayer

Registration Number 01426125

to the Faculty of Informatics

at the TU Wien

Advisor: Viktor Vad, MSc.

Vienna, 26th December, 2017

Julian Strohmayer

Viktor Vad

Erklärung zur Verfassung der Arbeit

Julian Strohmayer
Tautenhayngasse 19/6, 1150 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 26. Dezember 2017

Julian Strohmayer

Danksagung

Ich bedanke mich herzlich bei meinem Betreuer MSc. Viktor Vad, für die Gelegenheit mit ihm an diesem faszinierenden Projekt arbeiten zu dürfen. Die ausgezeichnete Hilfestellung, während der Entwicklungsphase des PlateViewer Frameworks, erlaubte es mir die festgelegten Ziele rasch zu erreichen. Abgesehen davon möchte ich mich für die durchwegs angenehme Art der Betreuung bedanken.

Acknowledgements

I want to thank my supervisor MSc. Viktor Vad, for giving me the opportunity to collaborate with him on such an fascinating project. The excellent guidance he provided, throughout the developmental phase of the PlateViewer framework, allowed me to achieve our goals at a very fast pace. Besides that, his pleasant nature really made me enjoy working on this thesis.

Kurzfassung

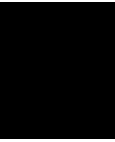
Die Durchführung von Phenotyping Experimenten im Bereich der Pflanzenzucht erfordert üblicherweise die Auswertung von hinreichend großen Datensätzen, wodurch händische Auswertung keine Option mehr ist. Zur Lösung dieses Problems wurden automatische Phenotyping Pipelines entwickelt, welche den manuellen Methoden hinsichtlich ihres Durchsatzes weit überlegen sind. Dies wird durch den Einsatz von automatischen Segmentierungs- und Detektionsmethoden erreichen, wodurch auf menschliches Eingreifen weitgehend verzichtet werden kann. Diesbezüglich existieren jedoch nicht für jede Art von Aufgabe ausreichend gute Lösungen. Ein solches, bisher ungelöstes Problem, ist die Bestimmung jenes Punktes auf der Pflanzenachse, an dem das Hypokotyl in die Wurzel übergeht. In dieser Arbeit wird ein neuer Ansatz zur Bestimmung dieses Punktes präsentiert, welcher Statistical Break Point Analysis zur Auswertung mehrerer Eigenschaften eines Keimlings verwendet. Zur Bestimmung geeigneter Eigenschaften wurde ein Visual Analytics Framework entwickelt, welches die Analyse von visuellen Daten individueller Arabidopsis Thaliana Keimlinge erlaubt. Die dafür verwendeten Daten entstammen einer automatischen Phenotyping Pipeline.

Abstract

Plant root phenotyping can be a tedious process if done manually, since it typically requires large data sets to be processed. The solution to this problem are automatic phenotyping pipelines, which allow significantly higher throughput than manual methods, by eliminating the need for human intervention. These pipelines rely on the robustness of automatic segmentation and detection methods for various plant characteristics. Due to numerous confounding factors, the detection of the hypocotyl/root transition point is still an unsolved task. In this thesis a novel approach to this problem, utilizing Statistical Break Point Analysis based on custom plant features, is presented. The approach has been developed using a visual analytics framework called PlateViewer, which was especially built for this task. The framework is able to analyze individual *Arabidopsis Thaliana* seedlings, taken from agar plate scan images, produced by an automatic phenotyping pipeline.

Contents

Kurzfassung	xi
Abstract	xiii
Contents	xv
1 Introduction	1
1.1 Phenotyping	1
1.2 Motivation	2
1.3 Related Work	3
1.4 Data Set	4
2 Methodology	7
2.1 Overview	7
2.2 Feature Selection	7
2.3 Statistical Break Point Analysis	12
3 Implementation	15
3.1 PlateViewer - A Visual Analytics Framework	15
4 Results	19
5 Discussion	25
List of Figures	27
List of Tables	29
List of Algorithms	29
Bibliography	31



Introduction

1.1 Phenotyping

The phenotype of an organism is a term used to refer to all of its observable characteristics, such as physiological and biochemical properties, morphology or behavior. A prime example for such a characteristic would be the human eye color, which is inherited and influenced by a multitude of different genes. This leads to the term genotype, referring to the genetic structure (DNA sequence) of an organism. Both were introduced by the geneticist Wilhelm Johannsen [Joh11] in the year 1911. As sometimes falsely assumed the genotype is not the only factor responsible for a certain phenotype, it is more a product of both genotype and environmental factors and their interaction, since some genes only express their observable characteristics under certain environmental conditions. In the case of plants this phenomenon is easily observable. Take two plants with identical genetic make up for example, where one is grown under ideal conditions (nutrient rich soil, ideal temperature, enough water and sunlight) and the second plant is exposed to critical growing conditions. This will most likely result in plants with different phenotypes, expressed as differences in their outer appearance, or other measurable properties such as the yield.

The quantification of these phenotypic properties is called phenotyping, which includes a wide variety of methods used to acquire data. These methods can either be classified as destructive or non-destructive. Destructive methods as the name suggests require the sample to be destroyed or damaged to conduct the measurement, whereas non-destructive methods leave the sample intact. Since the second case is often preferred or in fact necessary, a multitude of non-destructive methods have been developed which usually require some form of imaging technique, such as multispectral imaging [SMW18][ZPS⁺16] or thermography [CSA⁺17]. This class of methods has profited largely from the progress in the field of computer vision over the past decade and has become essential for plant phenotyping. Especially when high throughput is of importance. Due to the

computational nature of these methods, it is in some cases possible to create semi- or even fully automatic phenotyping pipelines, which allow significantly higher throughput than non-automatic pipelines. The demand for automatic, high throughput phenotyping pipelines is growing since most tasks require a large set of samples to be processed in a short amount of time.

1.2 Motivation

The germination phase of plants is a strict procedure. Once a seed has started to germinate, the radicle (root of the plant embryo) emerges from the seed and proceeds to grow further downward into the soil. At the same time the hypocotyl (stem of the plant embryo) grows towards the opposite direction, pushing the cotyledons (leaves of the plant embryo) upwards, until they emerge from the soil. The vertical growth direction of both radicle and hypocotyl is of course dependent on gravitational force, due to the gravitophims of plants. For plant phenotyping, the process of germination is of interest and has to be observed. However, in nature everything happens under ground, hidden from the human eye. This raises the need for plant root phenotyping under controlled environmental conditions, which is carried out by placing multiple seeds on agar plates to observe their phenotype, throughout the life cycle of a plant, starting with the germination phase. Depending on the experiment, environmental conditions such as the light spectrum, temperature, gravity, nutrient supply or the presence of elements/chemicals can be simulated, in order to trigger and observe certain characteristics of a plant. State of the art approaches usually use some form of semi- or fully automatic acquisition setup to acquire visual data at different points in time, which is then used as basis for further measurements conducted by software solutions [SGS⁺14]. However these measurements are not sufficient in some cases and require human supervision, which is limiting the throughput of the entire phenotyping pipeline. This leads to an entire class of unsolved problems, one being the detection of the hypocotyl/root transition point during the germination phase, treated in this thesis. The detection of this point in automatic phenotyping pipelines is important, since it is needed to keep track of developmental stages, understand the cotyledon/hypocotyl/radicle-composition and to conduct further measurements based on the location of this point. This means the quality of the detection is correlated with the outcome of following studies. It is therefore of utmost importance that the detected point describes reality as close as possible. This however is a non trivial task, since the transition region can vary in color, shape, length or diameter. In addition to this, limited image resolution and challenging light conditions are common problems. This fosters the need for robust and automatable detection methods that can be used in pheotyping pipelines. In this thesis, methods from the field of visual analytics and statistical analysis are used, to further understand the observable characteristics of plants growing under controlled environmental conditions and how they can be used to detect the hypocotyl/root transition point.

1.3 Related Work

This thesis builds on a project carried out by the Gregor Mendel Institute of Molecular Biology, located in Vienna [SGS⁺14]. The goal of this project was the creation of a low-cost, high throughput phenotyping pipeline for plant root growth, which could be applied on a large scale. The proposed method can be separated in a hardware and a software stage. The hardware stage is a Multi-CCD Flatbed Scanner Cluster consisting of eight Epson Perfection V600 photo scanners working in parallel, which were used for image acquisition. Figure 1.1 (a) shows this setup. Each scanner was fitted with a custom support frame to hold up to two 12 x 12 cm agar plates at a fixed position, throughout the scanning procedure. To improve contrast, the plates were scanned in a dark room with all the scanner lids open. Scans were performed with a resolution of 1200 x 1200 dpi which corresponds to an image size of ~6000 x 6000 pixels. Images were stored in the RGB TIFF format with 8 bit color depth. Figure 1.1 (b) shows one of these images. The visual analytics approach presented in this thesis uses these images as input. Since user interaction is usually limiting the throughput of supervised phenotyping pipelines, the described setup was designed in a way such that it can be operated by a single person using a dedicated control unit, which allows the scan of multiple agar plates simultaneously. This increases the throughput of this pipeline stage dramatically and further improves the scalability of this method. The second stage of the proposed method is an automated image processing software called BRAT (Busch-lab Root Analysis Toolchain), which was implemented as plug-in for the platform independent, open source framework Fiji. This software operates on the before acquired agar plate images, which can either be processed individually or as time series, where multiple images of the same agar plate were taken over a certain period of time. If the software detects such a time series, all plate images within the series are aligned using SIFT (Scale Invariant Feature Transform) [G.04]. These aligned images are then cropped to remove the plate corners, which hold no valuable information. After that, a threshold based on the background color, is calculated to create a binary image, in which regions of interest (potential seedlings) are drawn as foreground. Regions of interest are identified using the Marching Squares Algorithm [LC87]. Finally non seedling objects are eliminated. Once the seedling objects have been detected, 16 features based on the length, width and the topology of each seedling are calculated. This is done for all images in the time series. A complete list of all features is given in [SGS⁺14]. BRAT produces a multitude of object measurements as output, however not all of them are relevant to the approach presented in this thesis. Besides the raw agar plate scans, produced in the first stage, only object measurements such as the plant width and the pixel coordinates of the plant axis are needed. For details see section 1.4.

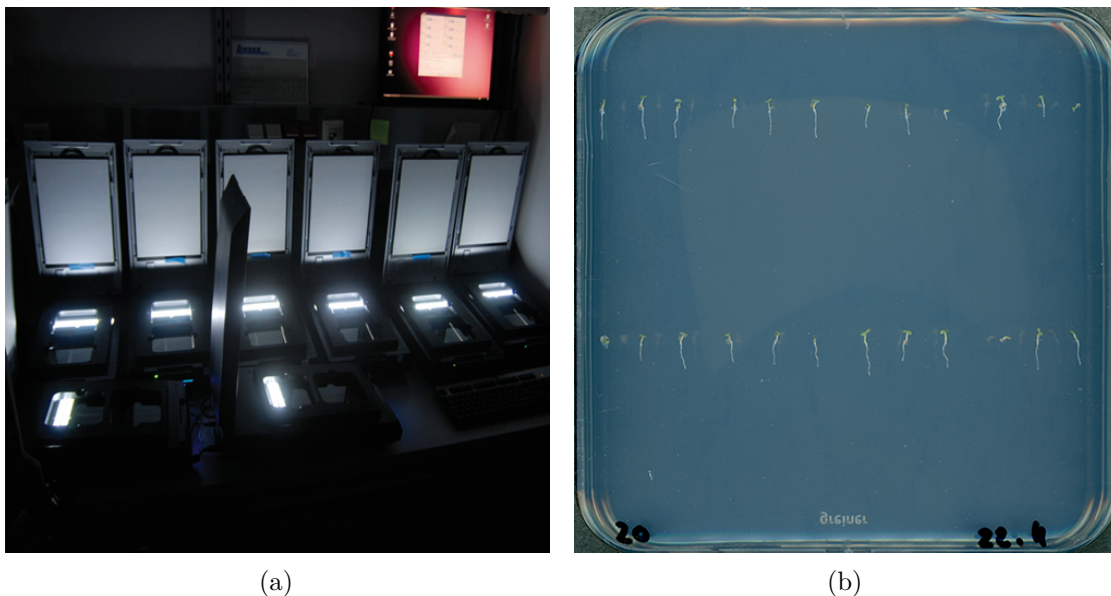


Figure 1.1: (a) GMI Multi-CCD Flatbed Scanner Cluster (b) Scan of agar plate with 24 Arabidopsis Thaliana seedlings selected from data set *A*, Source: [SGS⁺14]

1.4 Data Set

The detection capabilities, of the method presented in this thesis, were measured using data produced by the phenotyping pipeline in [SGS⁺14]. For this, two separate data sets were created, both having the same data structure. The general structure is as follows. There is an agar plate scan with Arabidopsis Thaliana seedlings in the TIFF image format, having an image size of $\sim 6000 \times 6000$ pixels and a file size of ~ 100 MB. This file has two corresponding text files which hold the ID of the plate scan in their name. An example of the naming scheme for these files is given by Table 1.1 (a)-(c). The first file contains the relevant measurements of all objects, which were detected on the plate scan, such as the object ID, the plant axis coordinates, the plant width and the pixel type. The entries in this file are 5 dimensional row vectors of the form [object ID, x , y , width, pixel type]. The second file holds a 2by3 affine transformation matrix, stored as 6 dimensional row vector, needed to transform the before mentioned plant axis coordinates in a way such that they match the plant on the plate scan. This is necessary due to a bug in the software, which was used to create these data sets.

Table 1.1: data structure and file naming scheme

(a) agar plate scan	RSMS_set6_day1_20111128_020.tif
(b) object coordinates	Object_Coordinates_RSMS_set6_day1_20111128_020.txt
(c) affine transformation	Affine_RSMS_set6_day1_20111128_020.txt

The larger of both data sets is a time series of a single agar plate, with 24 *Arabidopsis Thaliana* seedlings, scanned over a period of five days. This data set consists of five RGB TIFF images with an image size of 5612 x 6088 pixels and a file size of 97.7 MB, five object coordinate files and five affine transformation files. Figure 1.2 shows all plate scans of this data set and Figure 1.4 (a)-(e) shows the growth of a single seedling within this data set, over a series of five days. This data set can easily be distinguished from the second data set by the blue background color, opposed to the dark purple background color of the second data set. This data set will be referred to as data set *A* from now on.

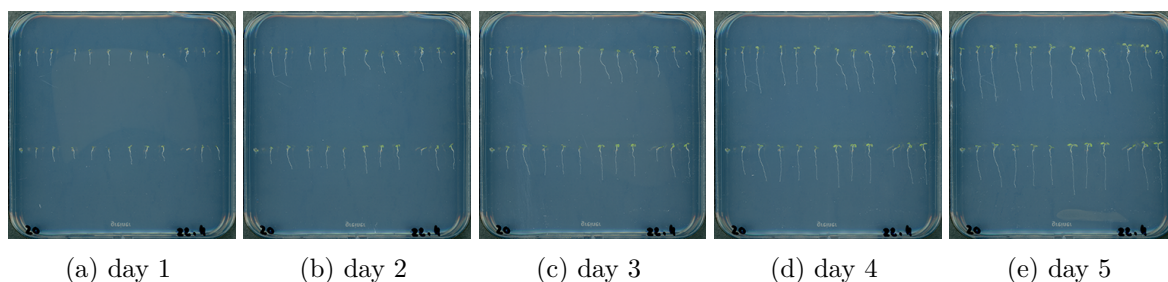


Figure 1.2: (a)-(e) Agar plate scans with *Arabidopsis Thaliana* seedlings over a series of five days taken from data set *A*

The second data set has two different agar plates with 24 *Arabidopsis Thaliana* seedlings, each scanned on the first day of the growth phase. The RGB TIFF images have an image size of 6608 x 6614 pixels and a file size of 125 MB. This data set also has two object coordinate files and two affine transformation files. As mentioned before, this data set has a dark purple background color, which might be due to a difference in the growing medium chosen for these particular plates. The difference is shown by Figure 1.3 and Figure 1.5. This data set will be referred to as data set *B* from now on.

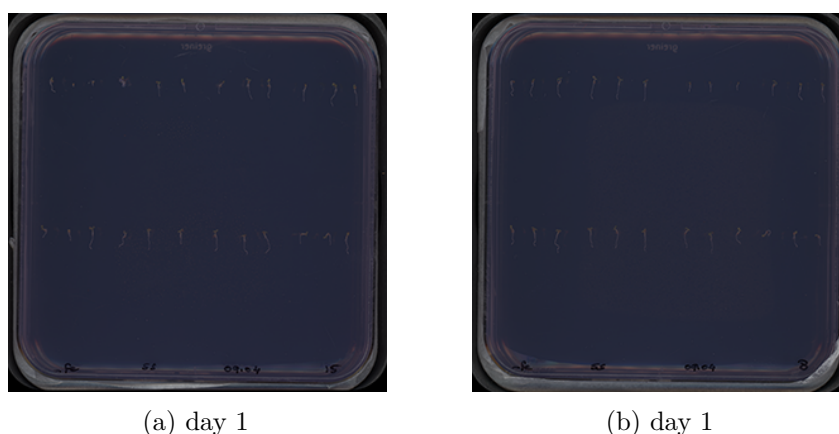


Figure 1.3: (a),(b) Two separate agar plate scans with *Arabidopsis Thaliana* seedlings at day one taken from data set *B*

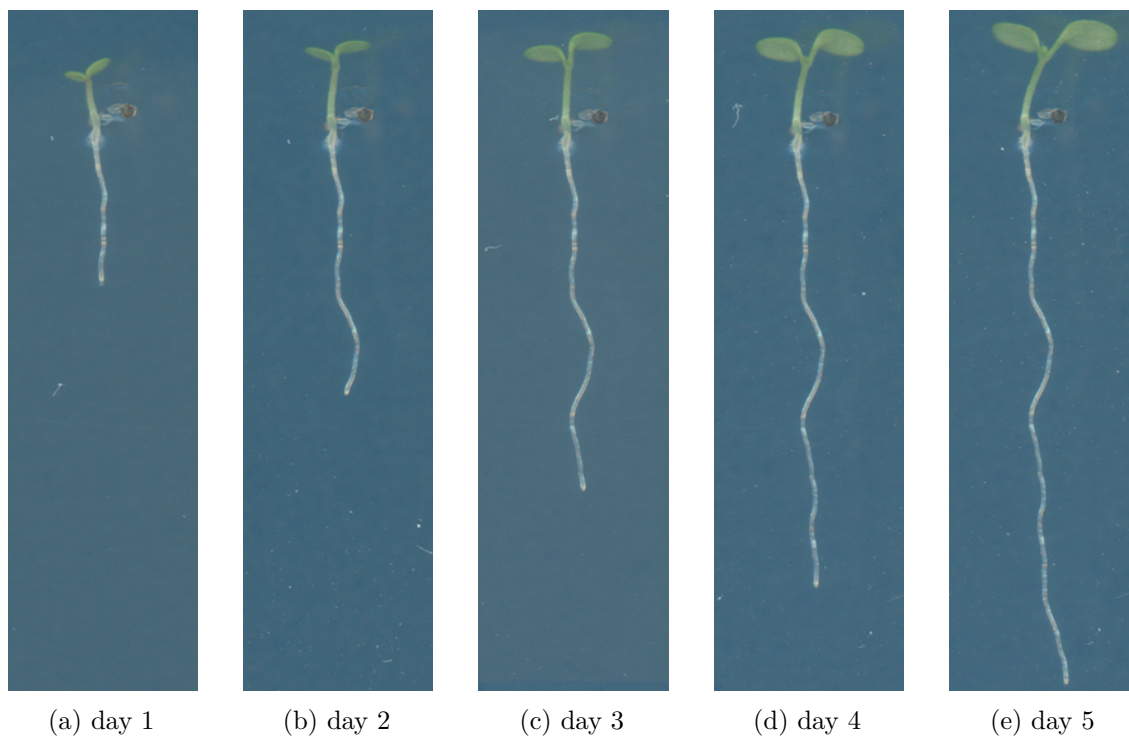


Figure 1.4: (a)-(e) Growth of Arabidopsis Thaliana seedling over a series of five days taken from data set *A*

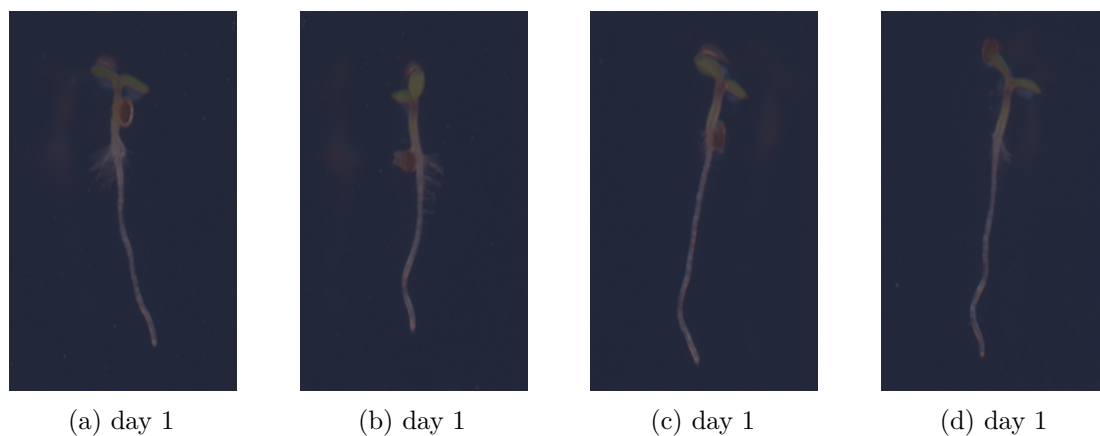


Figure 1.5: (a)-(c) Random Arabidopsis Thaliana seedlings at day one of the growth period taken from data set *B*

Methodology

2.1 Overview

Visual analysis of *Arabidopsis Thaliana* seedlings has hinted towards certain properties of the plant body, that may be used to create automatic and robust detection methods for the hypocotyl/root transition point. This thesis presents a visual analytics framework capable of measuring these properties, as well as the underlying correlation and present the result in an intuitive manner. For this task, agar plate scans with *Arabidopsis Thaliana* seedlings, taken from the data sets, are loaded into the framework where individual seedlings can be selected for detailed analysis. Once a seedling has been selected, automatic methods conduct measurements along the plant axis. These measurements include the CIELAB color values, the standard deviation of the horizontal intensity profile, the shape of the horizontal intensity profile (using DoG) and the width of the plant. The resulting set of feature vectors is then used to perform a statistical break point analysis, in order to find the point on the plant axis where the feature divergence is most prominent, which will be indicated by a global maximum. This point is assumed to be the hypocotyl/root transition. In the following section all features and the reasoning behind their selection is discussed in detail.

2.2 Feature Selection

2.2.1 CIELAB Color Coordinates

One prominent difference between the hypocotyl and the root is the greenness. This is due to the distribution of chlorophyll producing chloroplasts. The entire plant body shows a color gradient from green to white. Starting with the cotyledons having a bright green color, the greenness slightly decreases over the length of the hypocotyl. The hypocotyl/root transition region shows a rapid drop in greenness, which can be used for

detection purposes. The greenness of a plant is measured by sampling the RGB-values along the plant axis, on the plate scan image. The plant axis coordinates used, are a product of the segmentation of BRAT [SGS⁺14]. The extracted RGB-values are then converted into CIELAB color space coordinates, given by Figure 2.1 (b)-(d). Most notably in this figure is the divergence of the graphs (c) and (d) which are showing the green/red-difference and blue/yellow-difference along the plant axis. This point of divergence coincides exactly with the hypocotyl/root transition zone, which indicates a correlation between the change in greenness and the transition from hypocotyl to root.

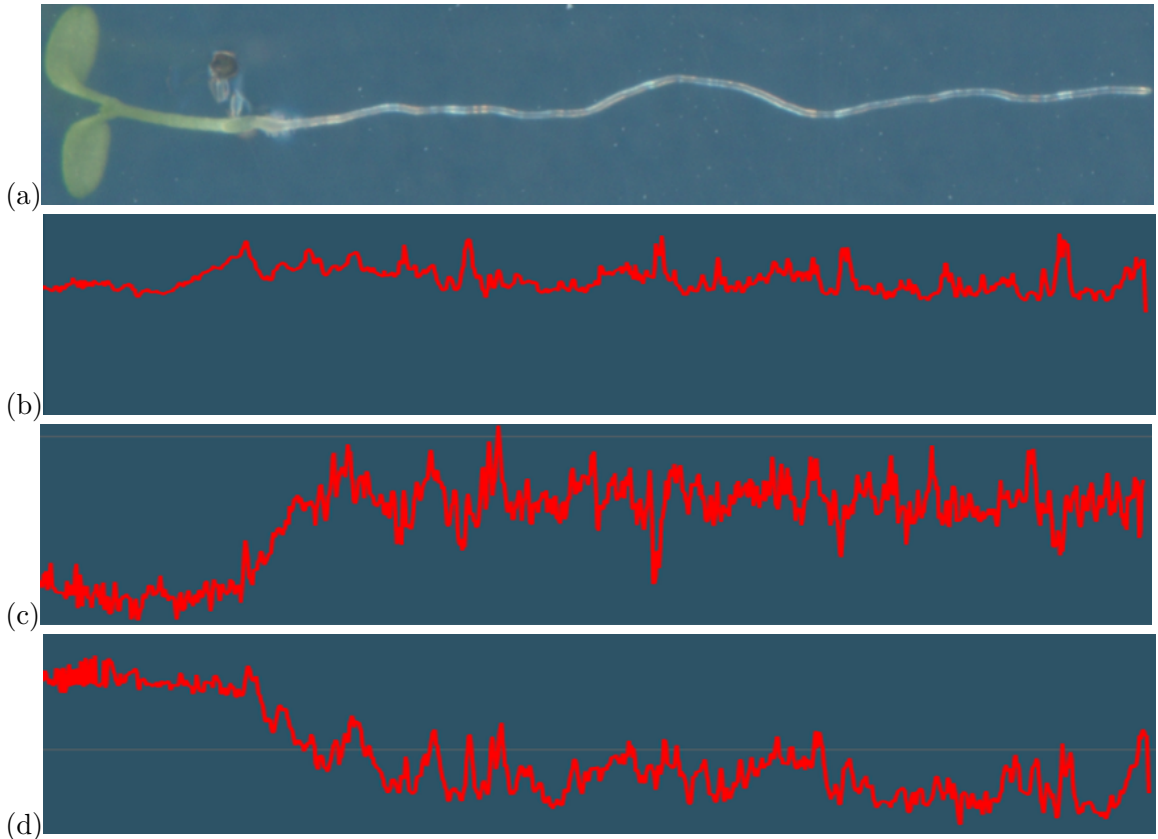


Figure 2.1: (a) Arabidopsis Thaliana seedling at day 4, (b) CIELAB L* (intensity) along the plant axis, (c) CIELAB a* (green-red) along the plant axis, (d) CIELAB b* (blue-yellow) along the plant axis

2.2.2 Plant Width

Seedlings show a visible difference in width (diameter) between the hypocotyl and the root, where the hypocotyl is usually twice the width of the root. Within the hypocotyl, the width remains almost constant until the hypocotyl/root transition zone is reached, where the width rapidly drops down to half of its original value. After the transition point, the width again remains almost constant, all the way to the tip of the root. This rapid change in width coincides with the hypocotyl/root transition point and can therefore be used for detection. The width values for each seedling on a plate scan were already included in the data set and could be used directly for analysis. Figure 2.2 shows this sudden change in width, at the hypocotyl/root transition.

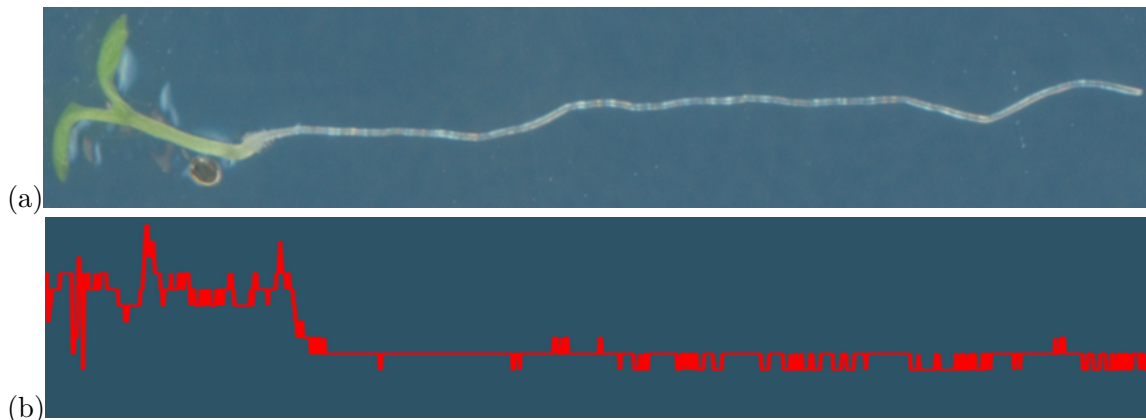


Figure 2.2: (a) Arabidopsis Thaliana seedling at day 4, (b) width (diameter) along the plant axis

2.2.3 Intensity Profile Shape (using DoG)

Figure 2.4 (a) shows a close up view of both the hypocotyl and the root. When compared, it is clearly visible that the hypocotyl has a more uniform intensity distribution across the horizontal axis, whereas the root shows a characteristic drop in intensity at the center. To detect the transition from one intensity profile to the other, a 1D DoG (Difference of Gaussian) filter kernel, mimicking the intensity profile of the root, was used. This was done by computing the horizontal intensity profile for each point on the plant axis, in the interval $[(x-w, y), (x+w, y)]$. Where x and y are the x and y coordinates of the point on the plant axis and w is the width (diameter) at this point. For each pixel in this interval the RGB-color values are sampled from the plate scan image and converted into the CIELAB color space, giving the horizontal intensity (CIELAB L^*) profile of a point on the plant axis. The intensity profile was then convolved with a 1D DoG kernel G , with a kernel size of w , created by $sG_1 - G_2$. Where G_1 and G_2 are both 1D Gaussian kernels with $\sigma_1 = 0.25$, $\sigma_2 = 0.1$ and scale factor $s = 0.8$. In order to mimic the intensity profile of the root as close as possible, the intensity profile of multiple seedlings was analyzed to find matching σ -values for the Gaussian kernels. The DoG kernel, as well as both

Gaussian kernels are given by Figure 2.4 (b). When convolved with the intensity profile, the DoG kernel produces a low response for points on the plant axis that lie within the hypocotyl and high responses for points on the plant axis that lie within the root. This is shown by Figure 2.3. Starting at the cotyledons the DoG kernel produces a low response. The response is then gradually increasing over the hypocotyl and is peaking at the point of the the hypocotyl/root transition. The entire root section is showing a high response with some drops, due to variations in intensity.

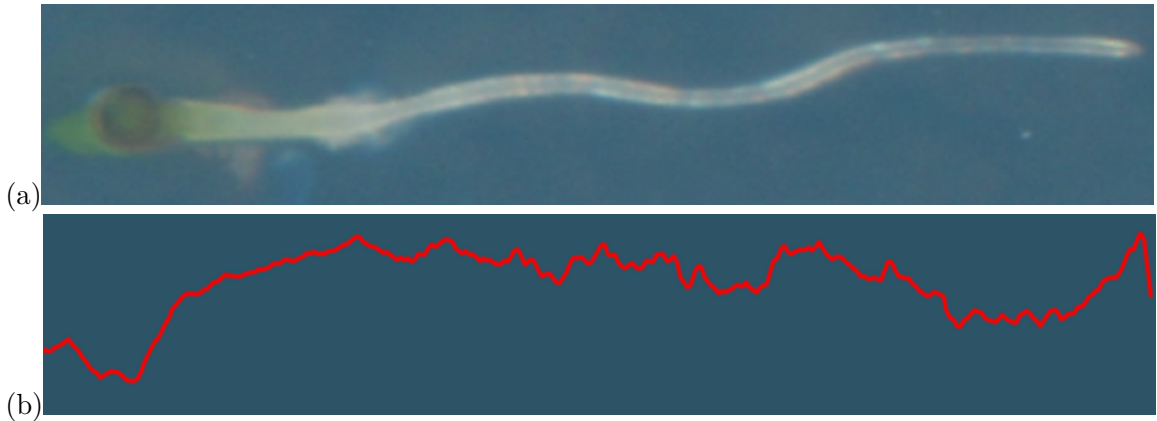


Figure 2.3: (a) Arabidopsis Thaliana seedling at day 1, (b) DoG filter response

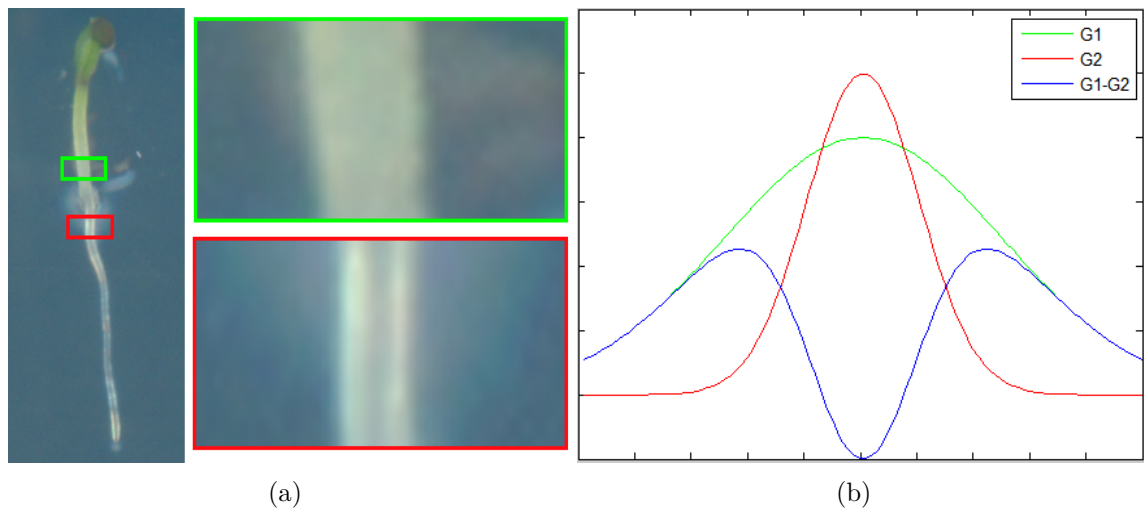


Figure 2.4: (a) Close up of hypocotyl (green) and root (red), (b) DoG kernel G (blue) Gaussian kernel G_1 with $\sigma_1 = 0.25$ and scale factor $s = 0.8$ (green) Gaussian kernel G_2 with $\sigma_2 = 0.1$ (red)

2.2.4 Intensity Profile Standard Deviation

The large variation within the intensity profile of the root, in contrast to the uniform intensity profile of the hypocotyl, suggests the use of a dispersion measurement to show the transition from hypocotyl to root. This was done using the standard deviation. For each point on the plant axis, the horizontal intensity profile for the interval $[(x - \lfloor \frac{w}{2} \rfloor, y), (x + \lfloor \frac{w}{2} \rfloor, y)]$ was computed, where x and y are the x and y coordinates of the point on the plant axis and w is the width at this point. For each pixel in this interval, the RGB-color values were sampled from the plate scan image and converted into the CIELAB color space. The resulting intensity values were then used to compute the standard deviation for the point on the plant axis. The resulting standard deviations, along the plant axis, for a random seedling, are given by Figure 2.5. It is clearly visible that the intensity values, within the intensity profiles of points on the cotyledons/hypocotyl, deviate much less than they are for points on the root. The transition from low to high deviation again coincides with the hypocotyl/root transition zone.

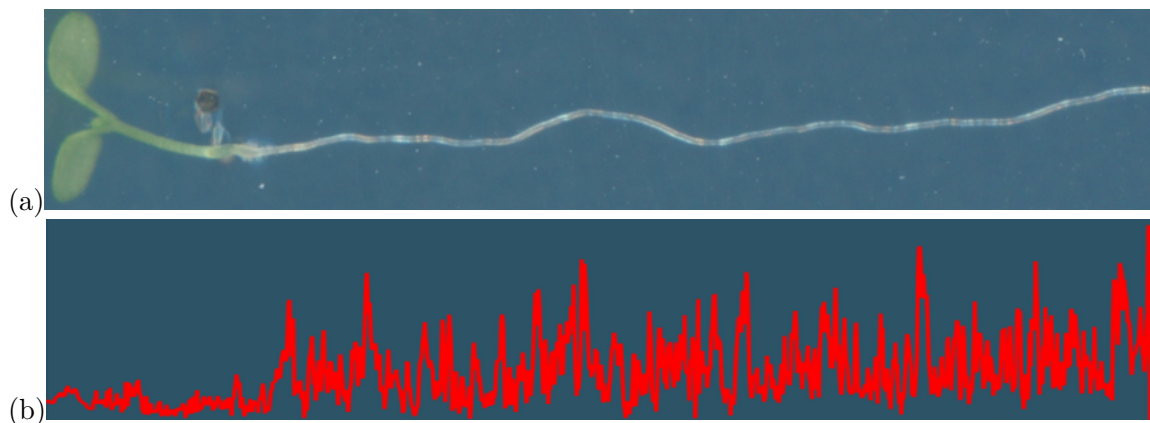


Figure 2.5: (a) Arabidopsis Thaliana seedling at day 5, (b) Standard Deviation within the intensity profile of points along the plant axis

2.3 Statistical Break Point Analysis

Each seedling with a plant axis of length n can be seen as multivariate data set of size n , where every point on the plant axis is represented by a 6-dimensional feature vector $y_k = (l, a, b, w, \sigma, r)$ with $k = 1, \dots, n$. Where l, a and b are the CIELAB color coordinates, w is the plant width (diameter), σ is the standard deviation within the intensity profile and r is the DoG response on the intensity profile. The hypocotyl/root transition is supposed to be the point on the plant axis where feature divergence is maximal. In order to find the break point in this multivariate data set, the method proposed in [AHHR09] was used. For each feature vector y_k , the the mean corrected version \tilde{y}_k is computed.

$$\tilde{y}_k = y_k - \frac{1}{n} \sum_{j=1}^n y_j, \quad k = 1, \dots, n \quad (2.1)$$

Then the mean corrected vectors are used to compute the extrapolated vectors Y_k , where $vech()$ is the symmetric matrix vectorizer operator, which stacks the columns of the lower triangular matrix, of a symmetric $d \times d$ matrix, to create a d' dimensional vector with a length of $d' = \frac{d(d+1)}{2}$.

$$Y_k = vech(\tilde{y}_k \tilde{y}_k^T), \quad k = 1, \dots, n \quad (2.2)$$

Now the vector quantities S_k , as well as the covariance matrix $\tilde{\Sigma}$ are computed. The vector quantities S_k are descriptors for the difference between the estimators $E[vech(\tilde{y}_k \tilde{y}_k^T)]$ of k observations and the estimator of all observations.

$$S_k = \frac{1}{\sqrt{n}} (\sum_{j=1}^k Y_j - k (\frac{1}{n} \sum_{j=1}^n Y_j)), \quad k = 1, \dots, n \quad (2.3)$$

$$\tilde{\Sigma} = \frac{1}{n} \sum_{j=1}^n Y_j Y_j^T \quad (2.4)$$

The final break indicator values λ_k , which will be plotted in the end, are given by $\lambda_k = S_k^T \tilde{\Sigma} S_k$. Let $\Lambda(p_k) = \lambda_k$ be the break indicator function that assigns a point p_k at position k on the plant axis its break indicator value λ_k . Now the hypocotyl/root transition point is supposed to be the point p on the plant axis with $\Lambda(p) = \max_{k=1, \dots, n} (\lambda_k)$. Additionally to this, the values $\bar{\lambda}_k$, which are based on the break indicator values, were also introduced. For a point p_k on the plant axis, $\bar{\lambda}_k$ is showing the mean break indicator value of the points p_{k+1}, \dots, p_n (p_n is the tip of the root).

$$\bar{\lambda}_k = \frac{1}{n-k} \sum_{j=k+1}^n \Lambda(p_j), \quad k = 1, \dots, n \quad (2.5)$$

Figure 2.6 shows the final result of the Break Point Analysis for a random seedling taken from data set *A*. Most notably is the prominent spike in graph (a) which is plotting the break indicator values λ_k along the plant axis. The maximum value coincides with the hypocotyl/root transition zone, which was the assumption. It can also be seen that the mean break indicator values $\bar{\lambda}_k$, given by graph (b), show small changes whenever a larger break occurs, even though the changes are very small in this particular case.

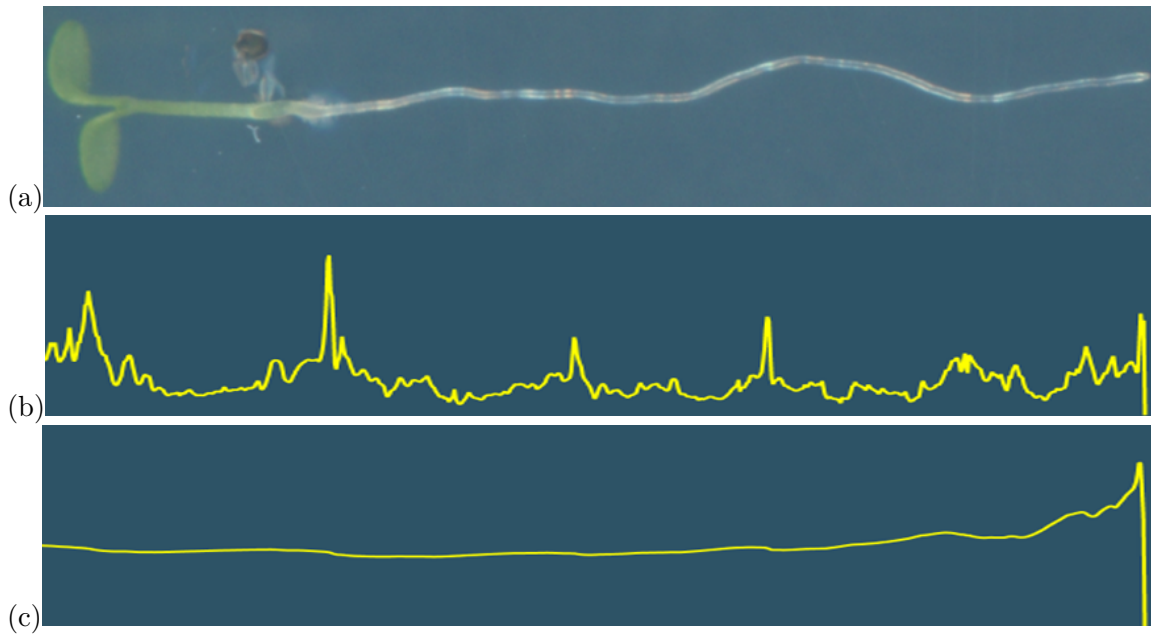


Figure 2.6: (a) Arabidopsis Thaliana seedling at day 3, (b) Break indicator values λ_k of the Break Point Analysis, (c) Mean break indicator values $\bar{\lambda}_k$

Implementation

3.1 PlateViewer - A Visual Analytics Framework

The data analysis in this thesis was carried out using a visual analytics framework called PlateViewer, which has been developed entirely for this task. The framework was implemented as stand-alone Java application. The code structure is based on the model-view-presenter architecture, allowing the decoupling of views from the underlying data model. This way, components can be replaced without much effort if the framework has to be altered, or expanded in the future. The graphical user interface layer was implemented in JavaFX and it uses methods from the field of visual analytics, such as multiple coordinated views to show the underlying data in different ways, which aims at supporting human pattern recognition. The combination of both automatic and manual analysis methods, provided by this framework, allows the user to find suitable feature combinations, needed to build automatic hypocotyl/root transition detection methods. The data analysis work flow in this application is comprised of two distinct stages.

The first stage requires a plate scan image, taken from one of the data sets, to be loaded into the application. At this point all corresponding data files have to be in the same directory as the plate scan image and have to have valid file names according to the file naming scheme, otherwise the loading process will fail. When all required files are in place, the application initiates the loading process of the plate scan image, which if successful, is indicated by the name of the plate scan image displayed in the graphical user interface. During the loading process the application also automatically generates a list of unique plant object ID's, extracted from the first column of the corresponding object coordinates file. Since seedlings are of primary interest, non-plant objects are ignored. The complete list of plant object ID's is made accessible in the graphical user interface by a drop down list, which allows the selection of individual seedlings from the plate scan image based on their ID. Once a plant object ID has been selected and the user decides to view the seedling, the application is reading all plant axis coordinates

3. IMPLEMENTATION

from the object coordinates file, having a matching plant object ID. Based on this list of x- and y-coordinates of the plate scan image, the minimum and maximum values are used to create a bounding box, which is again used to cut out a sub image from the plate scan image, only showing the seedling of interest. Two copies of this very same sub image are then aligned horizontally and combined to a single image. This is done because the entire plant axis, drawn in red, is superimposed on the left seedling, which could occlude regions of interest. By showing the seedling with the segmented plant axis next to the untouched seedling, no visual information is lost. Once the image is generated, it is displayed in the graphical user interface. The result is given by Figure 3.1. For the JavaFX 8 ImageView, used in the graphical user interface, custom zooming and panning operations haven been implemented, for even closer inspection of the seedling. Besides this main functionality, of showing individual seedlings, the application can also display the plate scan image as a whole. This concludes the first stage.

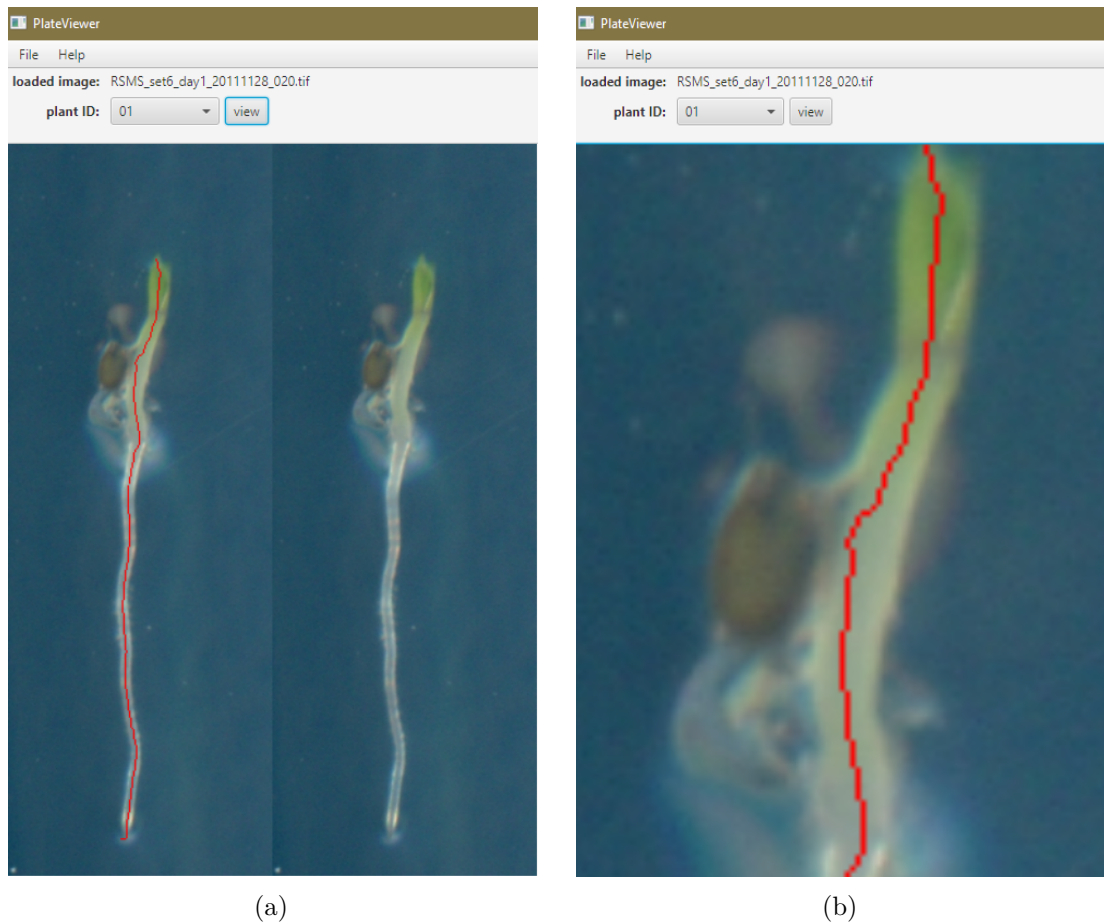


Figure 3.1: PlateViewer graphical user interface showing (a) the extracted sub image of the seedling with ID "01" and the superimposed plant axis drawn in red (b) the zoomed in view of the same seedling

In the second stage, in depth feature analysis is performed. For this, a seedling has to be selected as described in the previous stage and feature analysis has to be initiated by the user. When this is done a new window is created by the application and all feature values are automatically computed and the statistical break point analysis is performed. This way, multiple windows can be created, to allow analysis and comparison of multiple seedlings at the same time, by observing the windows side by side. The layout of the feature analysis window is basically split by a vertical line. On the left is the same image view as used in the previous stage, showing the seedling with the plant axis drawn in red, next to the seedling without the plant axis. The image view in this stage also shows a short, white horizontal marker line, which is intersecting with the plant axis. The marker line shows the hypocotyl/root transition point, estimated by the BRAT software (not Ground Truth). The estimated position is most of the time not correct and can therefore only be used as some kind of base line for a qualitative comparison, between the results produced by the method proposed in this thesis and the BRAT software. On the right side, multiple graphs (JavaFX 8 XYCharts) are stacked on top of each other. The data plotted in these charts is described by a headline on top of each chart. Besides that, there is also a visual difference between some of the charts. The difference being the color of the graph line. Most graphs have a red graph line, which is indicating that the plotted data are feature values, whereas the charts of the statistical break point analysis have yellow graph lines. This was done to highlight the importance of these graphs. In addition to this, the graphs of the statistical break point analysis are always at the top of the chart list. Since the analysis of a seedling is based on multiple different measurements along its plant axis, all graphs must share the same x-axis. The x-axis is showing integer values, corresponding to the pixels of the plant axis, which means of course the number of plant axis pixels are equal to the number of data points in each chart. This property has been used to implement custom marker lines (drawn in green), on top of the standard JavaFX 8 ImageView and JavaFX 8 XYCharts implementation, which link all views together. The graphs have a vertical marker line, whereas the image view has a horizontal marker line, due to the before mentioned relationship between the plant axis and the x-axis of the graphs. When a data point is selected in a graph, the marker lines in all other views are instantly moved to the position of the selected data point. This allows the comparison of local feature behavior, at a given plant axis position, possibly showing a correlation. In addition to this green marker line, which is present in all views, there is another type of marker line. The graph, plotting the cumulative mean of the statistical break point analysis, possesses a unique, light blue, horizontal marker line, which can be adjusted by a vertical slider at the right side of the graph. This marker line was intended to be used as a tool to find a threshold value, which coincides with the hypocotyl/root intersection. This could in theory be used to develop a automatic detection method. However this idea was not pursued any further, but the marker line was still left in the application, since it could be used in the future. The layout of the feature analysis window, as well as all described features are shown by shown by Figure 3.2 (some graphs are not displayed, due to space limitations).

3. IMPLEMENTATION

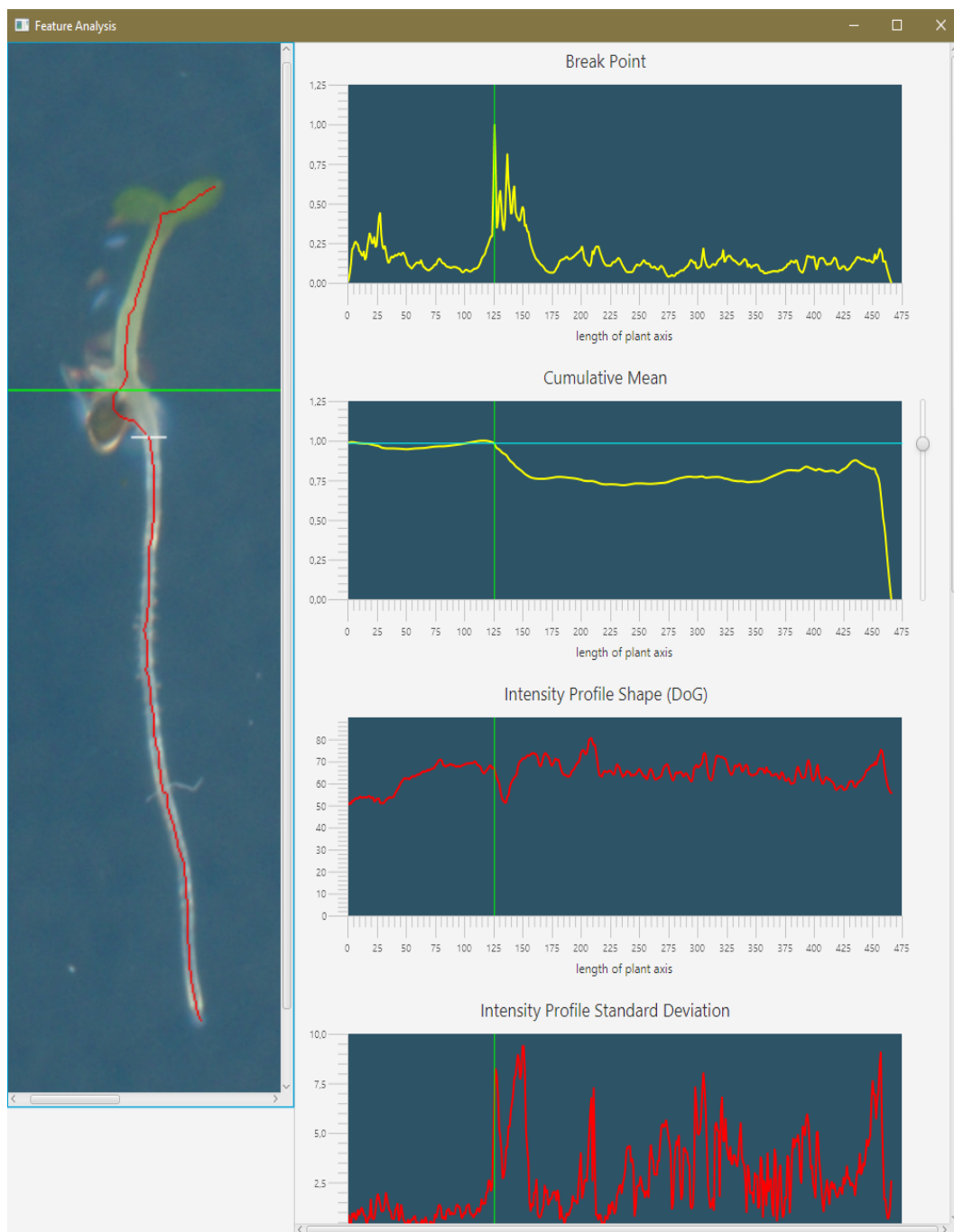


Figure 3.2: PlateViewer graphical user interface showing the results of the feature analysis stage (some graphs are not shown)

CHAPTER 4

Results

To evaluate the detection capabilities, regarding the hypocotyl/root transition point on the plant axis, of the proposed method, the PlateViewer framework was used. Data set *A* and data set *B* were used as input. The feature analysis function of the framework has been applied to each individual seedling within these data sets, in order to find the hypocotyl/root transition point, which is the point on the plant axis, corresponding to the maximum value of the statistical break point analysis. To determine the quality of the results, the estimated positions would ideally be compared to a Ground Truth labeling produced by a trained biologist. However both data sets do not contain any Ground Truth labeling and additional data was not available during the time of evaluation. In order to get reasonable measurements anyways, the results were compared to the hypocotyl/root transition point, estimated by the BRAT software. This measurement is shown by Figure 3.2 as the white, horizontal marker on the image of the seedling. To determine which method is superior in any given case, the distance between the "correct" hypocotyl/root transition point and the estimated positions was used as error metric. The method, producing a smaller error (distance) for a given seedling, is considered to be superior in that particular case. The superior method is denoted by "1" in the evaluation table and the inferior one by "0". The "correct" hypocotyl/root transition points, used for the metric, had to be estimated by a non expert, due to the previously mentioned reasons. The estimation may not be completely accurate in every case, however it is sufficient for the performed evaluation, since the relative difference between the produced error values is typically large enough to determine the clearly superior method. Additionally to this, seedlings that are considered as unusable for evaluation purposes are eliminated. This is done to eliminate incorrect measurements, that would otherwise only distort the evaluation result. A seedling is considered unusable under the following conditions: If the plant axis segmentation has failed, meaning that parts of the plant axis are completely missing, or the plant axis is severely misaligned. If the seedling is too small to see a difference between the hypocotyl and the root. If the seedling is severely malformed,

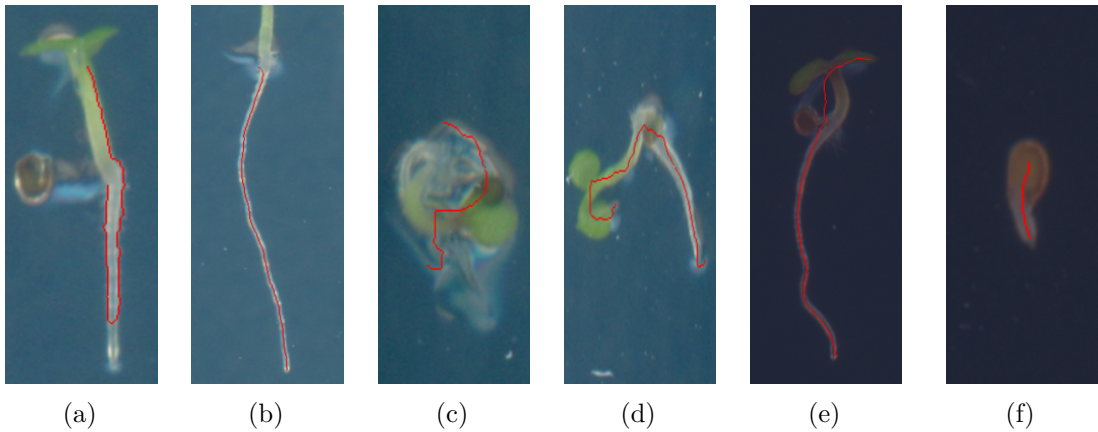


Figure 4.1: Examples of seedlings that are considered as unusable for the evaluation, (a) failed plant axis segmentation, (b) parts of the plant axis are missing, (c) severely malformed, (d) bad rotation, (e) severely misaligned plant axis, (f) seedling too small for evaluation

or rotated in a way such that the feature analysis would produce unusable results. A unusable seedling is denoted by "-" in the result tables. Examples of these conditions are given by Figure 4.1. Both data sets combined, had a total of 35 seedlings, meeting at least one of the conditions above. The reduced data set had total of 132 seedlings, suitable for evaluation purposes. This data set was evaluated twice. The first evaluation was carried out using all available features for the statistical break point analysis, in order to determine whether or not all features could be used for automatic detection methods. The results of each plate scan image, for this first evaluation, are given by the Tables 4.1-4.7. The overall result is given by Table 4.8, clearly showing that the BRAT software, with a success rate of 59.09 %, is superior to the PlateViewer detection method, only having a success rate of 40.91 %, when all features are used for the statistical break point analysis. Upon closer inspection of the cases, where the statistical break point analysis delivers worse results than the BRAT software, have revealed why this is the case. It turns out that most of the failed detection attempts are due to the growth of the cotyledons, after the second day. Once the cotyledons reach a certain size and begin to split apart, the object measurements produced by the BRAT software (which was used to create the data sets) are incorrect. The plant axis segmentation usually fails, producing a horizontal, or even a spiral shaped plant axis segmentation. This is especially problematic, since a failed plant axis segmentation near the cotyledons will result in large spikes in the plant width, which will destroy the result of the statistical break point analysis. Besides this, the results have shown that the intensity profile standard deviation feature is not suitable to detect the hypocotyl/root transition point, because large intensity fluctuations in the root will create too much noise in the data. Large intensity spikes near the tip of the root sometimes cause a maximum shift at a position where the hypocotyl/root transition point cannot be.

Table 4.1: Evaluation results of data set *A*, day 1 (all features)

Plant ID	1	2	3	4	5	6	7	8	9	10	11	12	15	16	17	18	19	20	21	22	23	24	25	26	sum
PLateViewer	1	1	1	0	1	0	0	0	-	1	-	-	-	1	1	1	0	0	0	1	0	-	-	0	9
BRAT	0	0	0	1	0	1	1	1	-	0	-	-	-	0	0	0	1	1	1	0	1	-	-	1	9

Table 4.2: Evaluation results of data set *A*, day 2 (all features)

Plant ID	1	2	3	4	5	6	7	8	9	10	11	13	14	17	18	19	20	21	22	23	24	25	26	27	28	sum
PlateViewer	1	1	1	1	-	1	1	1	0	0	0	1	-	-	0	0	1	0	0	0	1	0	-	0	1	11
BRAT	0	0	0	0	-	0	0	0	1	1	1	0	-	-	1	1	0	1	1	1	0	1	0	1	0	10

Table 4.3: Evaluation results of data set *A*, day 3 (all features)

Plant ID	1	2	3	5	6	7	8	9	10	11	13	14	18	19	20	21	22	23	24	25	26	27	28	29	sum
PlateViewer	0	0	1	0	1	0	0	0	0	1	0	-	-	0	0	0	0	1	0	-	0	-	1	0	5
BRAT	1	1	0	1	0	1	1	1	1	0	1	-	-	1	1	1	1	0	1	-	1	-	0	1	15

Table 4.4: Evaluation results of data set *A*, day 4 (all features)

Plant ID	1	2	3	5	6	7	8	9	10	11	13	14	18	19	20	21	22	23	24	25	26	27	28	29	sum
PlateViewer	0	0	1	-	1	1	0	0	0	1	0	-	-	0	1	1	0	0	1	1	0	-	0	1	9
BRAT	1	1	0	-	0	0	1	1	1	0	1	-	-	1	0	0	1	1	0	0	1	-	1	0	11

Table 4.5: Evaluation results of data set *A*, day 5 (all features)

Plant ID	1	2	3	5	6	7	8	9	10	12	13	18	19	20	21	22	23	24	25	26	27	28	29	sum
PlateViewer	1	0	1	0	1	0	0	0	1	1	-	-	1	0	0	0	0	0	0	0	-	0	0	6
BRAT	0	1	0	1	0	1	1	1	0	0	-	-	0	1	1	1	1	1	1	1	-	1	1	14

Table 4.6: Evaluation results of data set *B*, plate 1, day 1 (all features)

Plant ID	4	6	8	9	10	11	12	13	14	15	16	21	23	24	25	26	28	30	31	32	33	34	36	sum
PlateViewer	0	-	-	1	0	0	-	0	0	-	0	1	1	0	1	1	0	-	-	0	-	1	0	6
BRAT	1	-	-	0	1	1	-	1	1	-	1	0	0	1	0	0	1	-	-	1	-	0	1	10

Table 4.7: Evaluation results of data set *B*, plate 2, day 1 (all features)

Plant ID	3	4	5	6	7	8	10	11	12	13	14	15	22	23	24	25	27	28	29	31	32	33	34	35	sum
PlateViewer	1	-	-	-	0	1	0	0	1	0	0	0	1	-	1	1	0	0	1	1	-	-	-	0	8
BRAT	0	-	-	-	1	0	1	1	0	1	1	1	0	-	0	0	1	1	0	0	-	-	-	1	9

Table 4.8: Final result of the first evaluation (all features)

Plate	A1	A2	A3	A4	A5	B1.1	B1.2	sum	%
PlateViewer	9	11	5	9	6	6	8	54	40.91
BRAT	9	10	15	11	14	10	10	78	59.09

What the first evaluation has shown is that the cotyledons, as well as the standard deviation as a feature, do interfere with the detection capabilities of the statistical break point analysis. In order to resolve this issue, the standard deviation was excluded from the feature vector for the second evaluation. Additionally to this, the problem with the failed segmentation at the cotyledons was also resolved. For this the plant axis interval, used for the analysis, was limited to exclude problematic areas, that would otherwise destroy the detection of the hypocotyl/root transition point. Such areas are the cotyledons, as well as the tip of the root. These areas represent the start (cotyledons) and the end (root tip) of the plant axis, meaning the hypocotyl/root transition point must lie somewhere in between, but never within those areas. Knowing this, it is relatively easy to limit the searched area by ignoring a certain percentage of plant axis coordinates at the start and at the end of the plant axis. These percentages do not necessarily have to be equal. However, finding the perfect ratio is not trivial. If too many plant axis coordinates are excluded from the analysis, the hypocotyl/root transition point may be located within the excluded area and will therefore never be found by the statistical break point analysis. In contrast to this, if the percentage is too small, the cotyledons may again be interfering with the detection. Finding the optimal amount of plant axis coordinates to ignore is critical for the start of the plant axis (cotyledons), since the hypocotyl/root transition point is typically very close. At the end of the plant axis, the amount of ignored plant axis coordinates can be higher, because the hypocotyl/root transition point is usually located in the first half of the plant axis. The relative distance between the hypocotyl/root transition point and the tip of the root becomes even larger, the older a seedling grows. For the second evaluation, 15 % of plant axis coordinates at the start and 20 % at the end of the plant axis were excluded. However picking constant values is not ideal. During the growth of a seedling the distance ratios between start/end of the plant axis and the hypocotyl/root transition point change. For young seedlings (day 1) the ratio is typically close to equal, whereas for older seedlings (day 3+) the distance between the hypocotyl/root transition point and the tip of the root becomes very large. The relative distance between the start of the plant axis and the hypocotyl/root transition point also changes, but not as drastically. Instead of picking constant values, this knowledge could be used to choose dynamic percentage values, based on the age of the seedling, in order to find an optimal search interval. But this is subject to future work. After applying the mentioned changes to the feature analysis function of the PlateViewer framework, the statistical break point analysis delivered impressive results, when compared to the previous results. The evaluation results of all 132 seedlings, are given by the Tables 4.9-4.15. The final result of the second evaluation is given by Table 4.16, which is showing drastic improvements. The success rate of the PlateViewer method increased to 65.15 %, which is an improvement of 24.24 %, compared to the previous evaluation result. This clearly shows that the exclusion of the cotyledons is indeed necessary.

Table 4.9: Evaluation results of data set A , day 1 (without cotyledons and σ)

Plant ID	1	2	3	4	5	6	7	8	9	10	11	12	15	16	17	18	19	20	21	22	23	24	25	26	sum	
PLateViewer	0	1	1	0	1	0	0	1	-	0	-	-	-	1	1	0	1	1	1	1	1	1	-	-	1	12
BRAT	1	0	0	1	0	1	1	0	-	1	-	-	-	0	0	1	0	0	0	0	0	0	-	-	0	6

Table 4.10: Evaluation results of data set A , day 2 (without cotyledons and σ)

Plant ID	1	2	3	4	5	6	7	8	9	10	11	13	14	17	18	19	20	21	22	23	24	25	26	27	28	sum	
PlateViewer	1	1	1	1	-	1	1	1	0	0	1	0	-	-	1	1	1	1	1	0	0	1	1	-	0	0	14
BRAT	0	0	0	0	-	0	0	0	1	1	0	1	-	-	0	0	0	0	1	1	0	0	-	1	1	7	

Table 4.11: Evaluation results of data set A , day 3 (without cotyledons and σ)

Plant ID	1	2	3	5	6	7	8	9	10	11	13	14	18	19	20	21	22	23	24	25	26	27	28	29	sum	
PlateViewer	0	1	1	0	1	1	0	0	1	0	0	-	-	1	1	1	1	1	1	1	-	0	-	0	1	12
BRAT	1	0	0	1	0	0	1	1	0	1	1	-	-	0	0	0	0	0	0	0	-	1	-	1	0	8

Table 4.12: Evaluation results of data set A , day 4 (without cotyledons and σ)

Plant ID	1	2	3	5	6	7	8	9	10	11	13	14	18	19	20	21	22	23	24	25	26	27	28	29	sum	
PlateViewer	0	0	1	-	1	1	1	1	1	1	0	-	-	1	1	1	1	1	1	1	1	1	-	0	0	15
BRAT	1	1	-	-	0	0	0	0	0	0	1	-	-	0	0	0	0	0	0	0	0	0	-	1	1	5

Table 4.13: Evaluation results of data set A , day 5 (without cotyledons and σ)

Plant ID	1	2	3	5	6	7	8	9	10	12	13	18	19	20	21	22	23	24	25	26	27	28	29	sum	
PlateViewer	0	0	1	0	1	1	1	1	1	1	-	-	1	1	1	1	1	1	1	1	1	-	0	1	16
BRAT	1	1	0	1	0	0	0	0	0	0	-	-	0	0	0	0	0	0	0	0	0	-	1	0	4

Table 4.14: Evaluation results of data set B , plate 1, day 1 (without cotyledons and σ)

Plant ID	4	6	8	9	10	11	12	13	14	15	16	21	23	24	25	26	28	30	31	32	33	34	36	sum
PlateViewer	1	-	-	1	0	0	-	0	0	-	1	0	1	0	0	1	1	-	-	1	-	1	1	9
BRAT	0	-	-	0	1	1	-	1	1	-	0	1	0	1	1	0	0	-	-	0	-	0	0	7

Table 4.15: Evaluation results of data set B , plate 2, day 1 (without cotyledons and σ)

Plant ID	3	4	5	6	7	8	10	11	12	13	14	15	22	23	24	25	27	28	29	31	32	33	34	35	sum
PlateViewer	1	-	-	-	0	0	1	0	1	0	0	0	1	-	1	1	0	0	1	1	-	-	-	0	8
BRAT	0	-	-	-	1	1	0	1	0	1	1	1	0	-	0	0	1	1	0	0	-	-	-	1	9

Table 4.16: Final result of the second evaluation (without cotyledons and σ)

Plate	A1	A2	A3	A4	A5	B1.1	B1.2	sum	%
PlateViewer	12	14	12	15	16	9	8	86	65.15
BRAT	6	7	8	5	4	7	9	46	34.85

Discussion

It has been shown that the selected features, combined with the statistical break point analysis, can be used to detect the hypocotyl/root transition point. The proposed method even outperformed the BRAT software, in that regard. However, more testing has to be done. To verify the gathered results and to further improve the proposed method, larger data sets, including Ground Truth labeling are necessary. The segmentation quality within the used data sets should also be higher, since most measurements are dependent on correct segmentation. This fact was unveiled by the first evaluation, which showed the impact of incorrect segmentation on the detection. This shows the need for a robust cotyledon segmentation method, taking the separation of the cotyledons into account. Availability of correct segmentation for both cotyledons, would allow new approaches to the problem. As done in the second evaluation, the cotyledons can be excluded from the statistical break point analysis, which will be trivial to do, once correct segmentation is available. In this case, the constant exclusion percentages should be replaced by dynamic values, based on the age of a seedling. This will deliver better results, because constant values lead, in some cases, to the exclusion of plant axis areas that contain the hypocotyl/root transition point. This error is growth related and caused by the length ratio change of the hypocotyl and the root, during the growth of a seedling. Evaluation of seedlings in a time series, regarding their hypocotyl/root length ratio, could deliver usable exclusion percentages. The alternative approach would be, to use the entire plant axis, including the cotyledons. In this case the proposed method would have to be modified, in order to handle a forked plant axis. In addition to that, the (now correct) width values of the cotyledons would have to be handled in a way, such that the large drop in width, at the cotyledon/hypocotyl transition zone, is not affecting the detection of the hypocotyl/root transition point. This shows that the problem of hypocotyl/root transition point detection is not completely solved yet. Regardless of the chosen approach, many confounding factors are still present, which will have to be eliminated, in order to progress further. However this will be subject to future work.

List of Figures

1.1	(a) GMI Multi-CCD Flatbed Scanner Cluster (b) Scan of agar plate with 24 Arabidopsis Thaliana seedlings selected from data set A , Source: [SGS ⁺ 14]	4
1.2	(a)-(e) Agar plate scans with Arabidopsis Thaliana seedlings over a series of five days taken from data set A	5
1.3	(a),(b) Two separate agar plate scans with Arabidopsis Thaliana seedlings at day one taken from data set B	5
1.4	(a)-(e) Growth of Arabidopsis Thaliana seedling over a series of five days taken from data set A	6
1.5	(a)-(c) Random Arabidopsis Thaliana seedlings at day one of the growth period taken from data set B	6
2.1	(a) Arabidopsis Thaliana seedling at day 4, (b) CIELAB L* (intensity) along the plant axis, (c) CIELAB a* (green-red) along the plant axis, (d) CIELAB b* (blue-yellow) along the plant axis	8
2.2	(a) Arabidopsis Thaliana seedling at day 4, (b) width (diamter) along the plant axis	9
2.3	(a) Arabidopsis Thaliana seedling at day 1, (b) DoG filter response	10
2.4	(a) Close up of hypocotyl (green) and root (red), (b) DoG kernel G (blue) Gaussian kernel G_1 with $\sigma_1 = 0.25$ and scale factor $s = 0.8$ (green) Gaussian kernel G_2 with $\sigma_2 = 0.1$ (red)	10
2.5	(a) Arabidopsis Thaliana seedling at day 5, (b) Standard Deviation within the intensity profile of points along the plant axis	11
2.6	(a) Arabidopsis Thaliana seedling at day 3, (b) Break indicator values λ_k of the Break Point Analysis, (c) Mean break indicator values $\bar{\lambda}_k$	13
3.1	PlateViewer graphical user interface showing (a) the extracted sub image of the seedling with ID "01" and the superimposed plant axis drawn in red (b) the zoomed in view of the same seedling	16
3.2	PlateViewer graphical user interface showing the results of the feature analysis stage (some graphs are not shown)	18
		27

4.1	Examples of seedlings that are considered as unusable for the evaluation, (a) failed plant axis segmentation, (b) parts of the plant axis are missing, (c) severely malformed, (d) bad rotation, (e) severely misaligned plant axis, (f) seedling too small for evaluation	20
-----	--	----

List of Tables

1.1	data structure and file naming scheme	4
4.1	Evaluation results of data set A , day 1 (all features)	21
4.2	Evaluation results of data set A , day 2 (all features)	21
4.3	Evaluation results of data set A , day 3 (all features)	21
4.4	Evaluation results of data set A , day 4 (all features)	21
4.5	Evaluation results of data set A , day 5 (all features)	21
4.6	Evaluation results of data set B , plate 1, day 1 (all features)	21
4.7	Evaluation results of data set B , plate 2, day 1 (all features)	21
4.8	Final result of the first evaluation (all features)	21
4.9	Evaluation results of data set A , day 1 (without cotyledons and σ)	23
4.10	Evaluation results of data set A , day 2 (without cotyledons and σ)	23
4.11	Evaluation results of data set A , day 3 (without cotyledons and σ)	23
4.12	Evaluation results of data set A , day 4 (without cotyledons and σ)	23
4.13	Evaluation results of data set A , day 5 (without cotyledons and σ)	23
4.14	Evaluation results of data set B , plate 1, day 1 (without cotyledons and σ)	23
4.15	Evaluation results of data set B , plate 2, day 1 (without cotyledons and σ)	23
4.16	Final result of the second evaluation (without cotyledons and σ)	23

Bibliography

- [AHHR09] A. Aue, S. Hörmann, L. Horváth, and M. Reimherr. Break detection in the covariance structure of multivariate time series models. *ArXiv e-prints*, 2009.
- [CSA⁺17] A.C.W. Craparo, K. Steppe, P.J.A. Van Asten, P. Läderach, L.T.P. Jassogne, and S.W. Grab. Application of thermography for monitoring stomatal conductance of *coffea arabica* under different shading systems. *Science of The Total Environment*, 609(Supplement C):755 – 763, 2017.
- [G.04] Lowe, David G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.
- [Joh11] W. Johannsen. The genotype conception of heredity. *The American Naturalist*, 45(531):129–159, 1911.
- [LC87] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *SIGGRAPH Comput. Graph.*, 21(4):163–169, 1987.
- [SGS⁺14] Radka Slovak, Christian Göschl, Xiaoxue Su, Koji Shimotani, Takashi Shiina, and Wolfgang Busch. A scalable open-source pipeline for large-scale root phenotyping of *arabidopsis*. *The Plant Cell*, 26:2390–2403, 2014.
- [SMW18] Kate Sendin, Marena Manley, and Paul J. Williams. Classification of white maize defects with multispectral imaging. *Food Chemistry*, 243(Supplement C):311 – 318, 2018.
- [ZPS⁺16] Jianfeng Zhou, Mark J. Pavek, Seth C. Shelton, Zachary J. Holden, and Sindhuja Sankaran. Aerial multispectral imaging for crop hail damage assessment in potato. *Computers and Electronics in Agriculture*, 127(Supplement C):406 – 412, 2016.