Machine Vision and Applications (2016) 27:647–661
DOI 10.1007/s00138-015-0720-z

SPECIAL ISSUE PAPER

A framework for the extraction of quantitative traits from 2D images of mature *Arabidopsis thaliana*

Marco Augustin[1,5] · **Yll Haxhimusa**[2,3,5] · **Wolfgang Busch**[4] · **Walter G. Kropatsch**[5]

Received: 20 April 2015 / Revised: 30 July 2015 / Accepted: 4 September 2015 / Published online: 16 October 2015

Abstract In this work, we propose an image-based phenotyping framework for the determination of quantitative traits from mature *Arabidopsis thaliana* plants. Two-dimensional (2D) images taken from the dried and flattened plants are analyzed regarding their geometry as well as their branching topology. The realistic branching architecture is hereby reconstructed from a single 2D image using a tracing approach with a semi-circular search window. The centerline segments of the tracing procedure are subsequently merged and labeled based on a hierarchical approach combining continuity properties with geometrical and topological information determined during tracing. This paper covers a detailed description of the proposed plant phenotyping pipeline from the image acquisition process until the extraction of the quantitative traits. The framework is evaluated using a set of 106 images and compared to a manual phenotyping approach as well as a semi-automatic image-based approach. The most relevant results of this evaluation are presented.

✉ Marco Augustin
marco.augustin@meduniwien.ac.at

✉ Yll Haxhimusa
yll.haxhimusa@meduniwien.ac.at

1 Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria

2 Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Vienna, Austria

3 Faculty of Electrical and Computer Engineering, University of Prishtina, Prishtina, Kosovo

4 Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, Vienna, Austria

5 Pattern Recognition and Image Processing Group, Vienna University of Technology, Vienna, Austria

Keywords Image-based phenotyping · Geometrical traits · Topological traits · Tracing · Hierarchical reconstruction · Network of curvilinear structures

1 Introduction

In this section, we present the necessity for image-based plant phenotyping (Sect. 1.1), highlight relevant and inspiring state-of-the-art approaches (Sect. 1.2) and define the problem which is addressed in this work (Sect. 1.3).

1.1 Importance of plant phenotyping

Understanding the functional relation between the genotype, environmental conditions and the resulting appearance of a multicellular organism is crucial for different disciplines in life science. For instance, in medicine, the likelihood for certain pathologies when being carrier of a specific genetic variation is of interest. In animal biology, the outward manifestation (appearance, behavior) of different animal models is in scope of basic research enabling conclusions in a variety of fields [9].

In plant biology, the functional analysis of genes is often motivated by an increasing demand of nutrition as well as in ongoing climatic changes and its associated impacts on the natural environment of organisms [47]. Linking the genetic data of organisms with properties describing their appearance, so-called traits, is the goal of researches from multiple disciplines. While sequencing the genome of multicellular organisms can nowadays be solved effectively, the current bottleneck in large-scale functional genetic studies is the manual determination of traits [11].

Advances in digital imaging technologies offer nowadays the opportunity for high-throughput imaging [42] from dif-

ferent plants or plant components at different scales ranging from microscopic scale to larger scales (e.g., on the field) [21,39,46]. High-throughput imaging techniques combined with computer vision and image analysis methods can help to overcome the current bottleneck in large-scale genetic studies by providing frameworks/image analysis pipelines for an efficient extraction of traits from different organisms [41].

1.2 Recent developments in image-based plant phenotyping

The digital image-based phenotyping of plants is a rather new application in the field of computer vision and image analysis. The automated determination of quantitative root traits describing the root system architecture (RSA) was one of the earliest studied problems in plant phenotyping. Hence, the majority of current approaches deal with the determination of root plant traits of different plants like *Arabidopsis thaliana*, maize or rice [23,41].

Traits of interest regarding the RSA are for example the length of the main root, its curvature characteristics or the root's tip angle. These traits are observed over time (e.g., on a daily basis) during the early stages of growing and are combined to gain knowledge regarding the root growth rate [5,15,32,40]. A high temporal analysis of the root growth behavior is presented in [43] where the ability of root to react on gravity changes (root gravitropic experiments) is investigated.

While the previous research extracts traits from plants observed at different timestamps (days, minutes), other works analyze the architecture only at a specific moment. Besides the geometrical traits, also traits regarding the branching network (e.g., network depth) are extracted in [2]. The reconstruction of a 3D root shape from multiple 2D images for subsequent analysis is a topic of the approach presented in [51].

A further major application in image-based plant phenotyping is the extraction of traits describing the appearance of plant's leafs and rosette development. Images are acquired mainly from a camera system mounted on a robotic arm from the top view. After leaf segmentation, labeling and additional post-processing, quantitative traits like the leaf area or the leaf area distribution can be extracted [3,28–30,34]. To study the leaf growth behavior, images are acquired over a specific period of time.

Similar traits are also in focus for whole shoots (plant parts which are growing above ground) of crop plants like barley, wheat or pea. In [20,35], a 3D model of the plant is reconstructed from RGB images of multiple views and the traits like width, height or volume are calculated. An approach for the determination of the traits concerning the whole shoot of pea plant is presented in [18]. Three optical projections are used to calculate, e.g., the area of green components. The traits are extracted over a period of 21 days.

Moreover, image-based phenotyping can be used to characterize plant disease phenotypes [31].

As mentioned, there exists already a good amount and variety of approaches for phenotyping plants based on different image modalities and applications. Nevertheless, there exists only a limited number of methods concerning the analysis of the plant's shoot [1,12]. Especially when it comes to non-leaf traits. Furthermore, there is a lack of methods for the detailed (topological) analysis of networks of curvilinear structures like root or stem systems in the field of image-based phenotyping.

1.3 Aim of the work

This work deals with image-based phenotyping of the mature shoot system of plants. By the extraction of quantitative traits in the mature state of development, the final architecture of a plant can be analyzed. In this work, we present a framework which analyzes 2D images of mature *A. thaliana* plants.

A. thaliana is a small, flowering plant and a popular plant model for basic research which is widespread around the world [19]. The plants in this work were growing in a natural environment where the surrounding conditions were very well known. For the purpose of manual analysis, the plants were harvested and shipped to a lab where the subsequent phenotyping took place. For an adequate shipping and storing, the mature plants were pressed. In the mature state, *A. thaliana* mainly consists out of stems and siliques with some remaining dried flowers and leafs. A sample image is shown in Fig. 1.

Due to logistics and the single-shot 2D image acquisition, parts of the plant appear as overlapping regions on the images. Hence, in a segmented image, branching points would occur without having a physical relationship in nature. The problem addressed in this work is therefore defined as the reconstruction of the *realistic* plant architecture from single-shot 2D images of plants which appear as a network of curvilinear structures.

Prior applications of computer vision were already dealing with similar problems, e.g., when analyzing blood vessel networks, biological neural networks or aerial imaging [14,27]. Analyzing these structures is mainly carried out as two-step approach: In a first step, the object is segmented before parts of the network are grouped and labeled in a second step. Various methods were developed for the task of segmentation or skeleton extraction of such networks [13,14,25,27,33,48–50]. The amount of approaches which deal with the problem of automated labeling and grouping of trees and sub-trees is more limited [6,17,22,26,36,45].

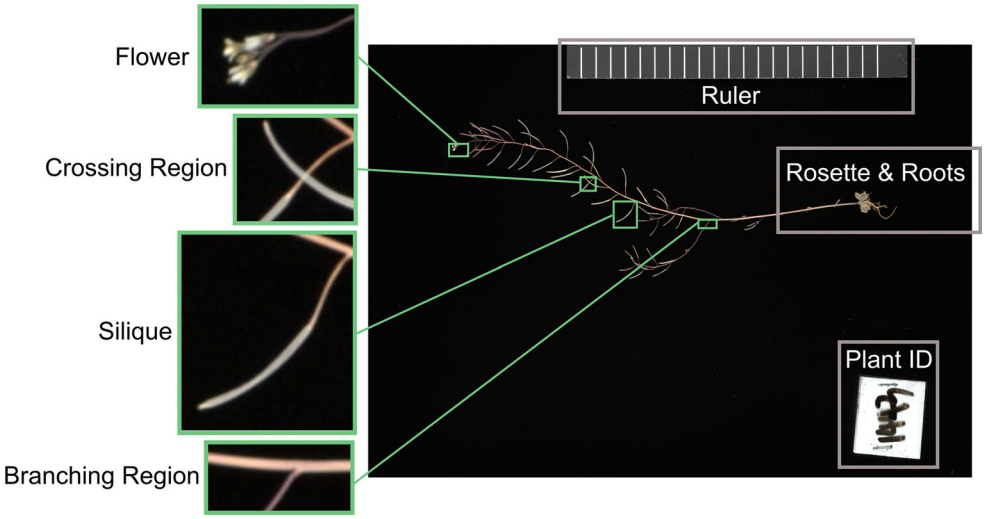


Fig. 1 Two-dimensional, top view, single-shot image of mature *A. thaliana*. In the mature state, the plant mainly consists out of stems and siliques. An example for a branching region as well as for a crossing region is emphasized. The image corresponds to approximately 45 × 30 cm in real world

The grade of complexity of the problem addressed in this work is defined by the number of critical points as well as imaging properties. Critical points are branching regions having a physical connection in nature, termination points and especially crossing (overlapping) regions. After reconstructing the plant's architecture, quantitative traits regarding the geometry of the plant (e.g., length) but also concerning the network topology can be determined.

In our work, we focused on the feasibility of high-throughput analysis for mature plants, which should also minimize the time spend for preparation of the plants before the images are acquired, thus we did not require to fix or/and to untangle the plants. We have also abounded "traditional" skeletonization-based methods for plant phenotyping [1,12] after a preliminary study. A lot of branch and silique overlapping occurs thus skeleton-based methods cannot correctly detect overlapped branches [1]. By untangling the plants, these problems can be overcome, which means also more time needed to prepare the plants before image capture. Tracing algorithms are alternative methods in the analysis of networks of curvilinear structures. Tracing is the process of following a structure, mostly curvilinear or tubular, from a defined starting point until a termination point. The advantage of tracing algorithms is the simultaneous segmentation and determination of local properties along (and only along) the structure. In this work, we propose the use of an adapted tracing algorithm with a semi-circular search window [8,48] to extract centerline segments of the network in a first step. In a subsequent step, the geometrical and topological properties of these segments are used to hierarchically reconstruct the plant's architecture based on continuity principles. The hierarchical reconstruction allows to deal with overlapping of branches and siliques in adequate manner.

Along with the core task of reconstructing the plant's architecture, a framework is developed for an effective phenotyping of the plant which includes pre-processing steps such as an automated spatial calibration but also generating appropriate resulting images for a quick quality control.

Within this work, we show the ability of the proposed framework to accurately extract geometrical traits in an semi-automated fashion. An automated detailed topological analysis is limited to the complexity of the network as well as by the image properties. To the best of our knowledge, this framework is the first approach dealing with the geometrical and topological analysis of the mature stem structure of *A. thaliana* plants using tracing approaches.

This work is an extension to our previous work presented in [4]. We provide in this paper a more detailed description of the methodology, especially regarding the required pre-processing steps. In Sect. 3.2, we present recent updates to our tracing approach for a reduction of user interaction along the steps of the proposed pipeline. Furthermore, we extended the evaluation section by a second phenotyping approach which was carried out by one of our collaborators.

The paper is structured as follows. In Sect. 2 an overview of the image acquisition setup as well as a more detailed description regarding the plants is given. In Sect. 3, we introduce the framework which was developed and describe the methodology. The paper is concluded with experiments (Sect. 4) and a conclusion (Sect. 5).

2 Imaging the mature *A. thaliana* shoot system

In this section, we will discuss the acquisition setup for imaging mature *A. thaliana* (Sect. 2.1) and the proposed model of the mature *A. thaliana* plant (Sect. 2.2).

2.1 Image acquisition setup

Imaging the mature state of plants which are growing in different regions of the world is a challenging task in terms of logistics. A trade-off between time consumption, cost-effectiveness as well as resulting limitations for high-throughput phenotyping has to be achieved. In this work, the images were taken as part of a project entitled *The molecular basis of local adaption in A. thaliana* led by Benjamin Brachi (Bergelson Lab, University of Chicago, US). The main intention of acquiring the images was a subsequent manual phenotyping. The use of computer vision and image analysis methods for digital phenotyping was not the main focus as no suitable framework was present at that moment.

The images of the plants were acquired indoors using a consumer digital camera (Nikon D700, digital single-lens reflex; focal length: 60 mm; f-number: $f/13$; exposure time: 1/125 s) with 12.1 megapixels (4284×2844 pixels) and stored as TIFF (tagged image file format) in RGB color mode with 8-bit color depth. The camera is fixed and two additional Bowens flash lights were used to maintain equal lighting conditions. The pressed plants were placed on a black velvet board to establish a high contrast between object and background. A velcro scale as well as a velcro wipeable number tag were added to the velvet for later analysis. Each plant is imaged once (single shot) and is completely visible, i.e., the shoot system, the rosette and the roots. A sample image from the dataset is shown in Fig. 1.

The benefits of this setup are the relatively low costs and the suitability for large-scale imaging. Due to logistical reasons (different harvesting locations), the plants were dried and pressed for shipping. Hence, some traits concerning the architecture of the plant cannot be reliably extracted, e.g., the original branching angles. Furthermore, the occurrence of crossing regions in the plant itself as well as in the 2D images leads to an increase of the complexity of the reconstruction process.

A 3D image acquisition setup would not automatically yield to a more promising analysis of the plants as they were already manipulated before the images are acquired. A complete and clear reconstruction of the mature plants would only be possible with on-field imaging technologies, having the drawback of high costs.

2.2 Model of the mature *A. thaliana* shoot system

The mature *A. thaliana* shoot system mainly consists out of stems and siliques. This appearance can be modeled as a network of curvilinear structures similar to blood vessel networks or biological neural networks. Hence, to simplify the task of analyzing the mature plant's shoot system, some constraints regarding the plant itself as well as regarding the constant image acquisition setup are defined.

The plants' stems and siliques are defined as piecewise linear (curvilinear) structures and their medial axis is represented by connected line segments. The stems and branches appear as bright structures on a dark background, their crosswise intensity profile approximates a Gaussian profile and the intensity changes along the stems and siliques are smooth. The stems and branches' diameters decrease coming closer to the plant's end points and do not change abruptly. The siliques show a stronger variation in diameter compared to regular stems.

The region of the rosette is defined as the origin of all main stems and the plant forms a tree-like structure for which side branches are building further sub-trees. The rosette is only of importance for finding the origin of the shoot system, but is not relevant in terms of phenotyping the shoot system. The same holds for the remaining roots which are sometimes present on the images.

The topological information of the plant is characterized by the physical connection between different parts of the plant (e.g., siliques, leafs, flowers, branches). Each of these parts can further be described by its geometrical characteristics such as size, shape, spatial location or orientation. The term *"architecture"* of a plant comprises the topological as well as the geometrical information and describes the different components of the plant regarding to space and time [16]. Characteristics describing the architecture of a plant are denoted in the following either as geometrical or topological traits. For, e.g., the number of siliques (on different stems) is considered as a topological trait, whereas the individual silique's length is considered as a geometrical trait. Regarding topological characteristics, the architecture of the plant can range from simple plants (low number of critical points, e.g., branching regions) to complex plants (high number of critical points).

3 Mature plant phenotyping framework: components and methodology

In this section, we propose a phenotyping pipeline for the purpose of the extraction of quantitative traits from 2D images of mature *A. thaliana* plants. The framework fulfills the tasks of pre-processing the images, extracting centerline segments of the plant's shoot system and reconstructing the realistic stem

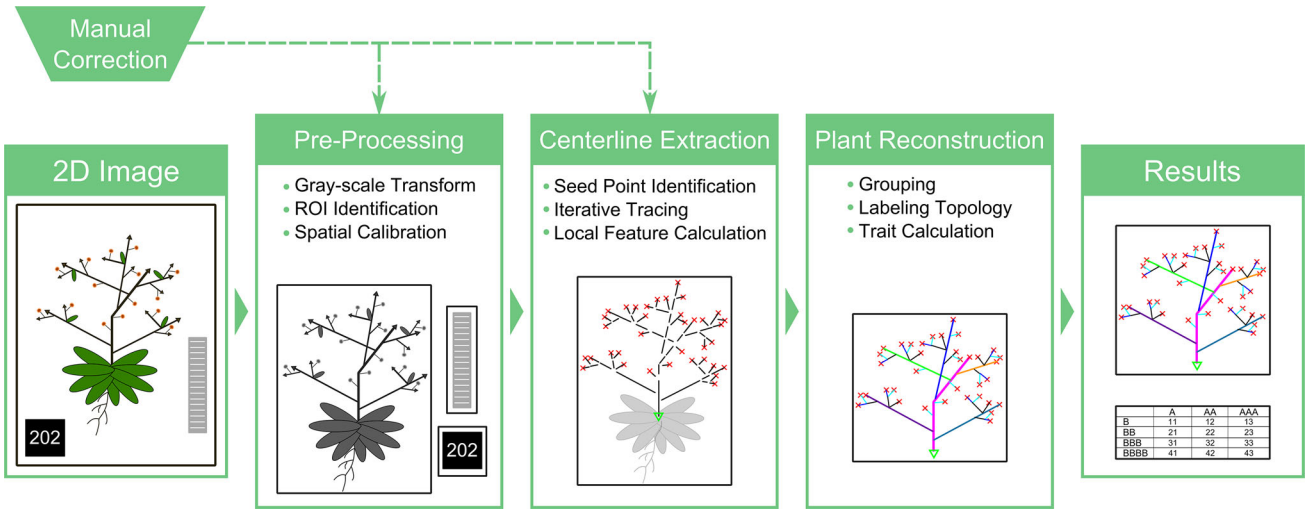


Fig. 2 Overview of the proposed mature plant phenotyping framework

architecture before the quantitative traits are calculated. An overview of the proposed pipeline is show in Fig. 2. This section is divided according to the three major steps of this pipeline.

3.1 Pre-processing

In this subsection, the basic steps of the image pre-processing are described.

3.1.1 Gray-scale transformation

Due to drying of the plants, the color information provided by the RGB images was unevenly in one image and over the whole dataset. Hence, the RGB images were transformed to 8-bit gray-scale images by extracting the R-channel of the RGB color space as the color information could not be used as an additional contrast for parts of the plant.

Another reasonable gray-scale transformation was achieved using the L-channel after transforming the RGB image into the L*a*b* color space. Certainly this transformation has to be considered as computationally more expensive. Nevertheless, in this work, a trade-off between the achieved contrast and the computational time was made using the R-channel over the L-channel.

3.1.2 Automated region-of-interest (ROI) identification

There are three ROIs identified in each image: the plant ID sign, the scale and the plant itself. The automated identification of these regions in each image $\mathbf{I}$ is achieved by the following procedure:

1. Down-sampling of $\mathbf{I}$ to $\mathbf{I}'$ by 1/8 of its initial size (3rd level of a Gaussian image pyramid).
2. Transformation to a binary image $\mathbf{B}'$ by calculating the maximum of the bit-planes 5–8 from $\mathbf{I}'$ at each position $\langle x, y \rangle$

$$\mathbf{B}'(x, y) = \max_{\text{bit}=5\ldots8} \mathbf{I}'_{\text{bit}}(x, y). \tag{1}$$

   With the chosen subset of bit-planes, this operation is equivalent to a gray-level threshold at level 15.
3. For each connected foreground region in $\mathbf{B}'$, the perimeter, the centroid and the bounding box are determined. The connectivity was chosen to be 8.
4. After filtering out small regions (perimeter is less than 5 % of the maximum), the remaining regions are sorted by their centroids' horizontal position of the plants. The arrangement of the ROIs in each image was constant during acquisition, hence the ROIs could be assigned to the regions in $\mathbf{B}'$.
5. The bounding box of each region in $\mathbf{B}'$ was rescaled to the initial size of $\mathbf{I}$ and the ROIs were extracted.

3.1.3 Automated spatial calibration

To get a relation between real-world metric units, e.g., mm, and pixel values a spatial calibration is needed. In this work, the spatial calibration is achieved by the use of the scale which is present on the black velvet board during acquisition.

The cropped image of the scale bar $\mathbf{I}_{\text{SB}}$ is extracted by the process described in automated ROI identification above. Mainly three colors are present in $\mathbf{I}_{\text{SB}}$: white (ruler units), mid-gray (ruler background) and black (velvet board). With use of these constraints, the process for the automated spatial calibration can be described as follows:

1. Binarization of $\mathbf{I}_{\mathrm{SB}}$ by a fixed threshold t which was set to 175.

$$\mathbf{B}_{\mathrm{SB}}(x, y) = \begin{cases} 1 & \text{if } \mathbf{B}_{\mathrm{SB}}(x, y) > t \\ 0 & \text{else} \end{cases}. \tag{2}$$

2. To calculate the distance between the scale graduations, remaining unwanted objects in $\mathbf{B}_{\mathrm{SB}}$, such as numbers have to be removed. This is done by morphological reconstruction operation ρ:

$$\mathbf{B}'_{\mathrm{SB}} = \rho_{\mathrm{mask}}(\mathrm{marker}). \tag{3}$$

   The binary image $\mathbf{B}_{\mathrm{SB}}$ is used as a mask, the marker image is calculated by morphological eroding ($\ominus$) of $\mathbf{B}_{\mathrm{SB}}$ with a structuring element **SE**. A horizontal line with a width of 1/3 of $\mathbf{I}_{\mathrm{SB}}$ was used as the structuring element.
3. Only the scale graduations remain in $\mathbf{B}'_{\mathrm{SB}}$ by the previous operations. For each foreground region in $\mathbf{B}'_{\mathrm{SB}}$, the centroid is calculated in a first step before the distance between neighboring scale graduations is determined. The average Euclidean distance and the a prior-known real-world metric unit is subsequently used for the final calculation of the conversion factor.

The procedure of automated spatial calibration is done for each image which is analyzed by the framework.

3.2 Centerline extraction

The centerline of an object is an efficient representation to analyze it in terms of geometry as well as topology. Tracing or line following approaches extract the centerline of a "thin" object by iteratively traveling along the object using a certain search window. The algorithm is initialized at certain starting points, so-called seed points and terminates if specific (stopping) criteria are fulfilled. In this work, a tracing algorithm using a semi-circular search window, similar to the one in [8,13], is used for extracting the centerline of the plants' stems. Local features like tracing direction or radius are determined already during the centerline extraction process and can be used for the plant reconstruction process.

3.2.1 Seed point identification

The initial set of seed points to start the iterative tracing procedure can either be chosen manually or by an automated procedure. Automated procedures focus mainly on the identification of ridges/edges along scan lines (grid) [7,13] or in the neighborhood of the ROI [17].

In this work, we propose to use two sets of seed points. The first set of seed points $\mathbf{S}_{\mathrm{I}}$ is used for initializing the tracing procedure and for the analysis of the branching network of the plant. A second set $\mathbf{S}_{\mathrm{II}}$ is used to guarantee the completeness of the centerline extraction process.

Automated identification of $\mathbf{S}_{\mathrm{I}}$ The rosette of the plant is the region where the stems originate and therefore is suitable to start the reconstruction of the plant. Therefore, an initial seed point is wanted on every stem leaving the rosette (denoted as main stems). An automated method to identify ridge points on the main stems is proposed as follows and motivated by the subsequent tracing procedure itself:

1. The location of the rosette is expected in the lower part of the image. Hence, the search area is reduced to the lower third of the image. The image is down-scaled to one half of its original size.
2. The image is binarized as proposed in Sect. 3.1.2 by (1).
3. The rosette is identified by its morphological difference compared to stems. While the stems are modeled as piecewise linear objects, the rosette and its overlapping leafs are modeled by an elliptical or circular shape. The diameter of the rosette is expected to be higher than the width of the biggest stem occurring in the image. An approximate segmentation of the rosette is achieved using repeated morphological opening using a disk-shaped structuring element. During each repetition, the disk diameter is decreased by 1 pixel. The procedure is repeated until only one connected component is present in the opened image. The initial diameter of the disk is set to 20 pixels.
4. On basis of the centroid $\mathbf{p} = \langle p_x, p_y \rangle$ and the equivalent diameter d_{eqv}, a semi-circular neighborhood in the growing direction is considered to identify possible ridge points. Specifically, a semi-circular gray-value profile $\mathbf{G}(r, \mathbf{p})$ with radius r from point $\mathbf{p}$ in an image $\mathbf{I}$ is defined as a sequence of n equally spaced samples taken along the circumference [8]:

$$\mathbf{G}(r, \mathbf{p}) = \{c_i,\ i = 0, 1, \ldots, n-1\}, \tag{4}$$

$$c_i = \mathbf{I}\left(p_x + r \cdot \cos(i \cdot \delta\theta),\ p_y + r \cdot \sin(i \cdot \delta\theta)\right), \tag{5}$$

$$\delta\theta = \frac{\pi}{n-1}. \tag{6}$$

5. The profile $\mathbf{G}(r, \mathbf{p})$ is subsequently filtered by a Gaussian filter (size $= 13$, $\sigma = 1$). The profile is determined for different scales of r starting with d_{eqv} and incremented by 5 pixels. The sampling rate n is set to $d_{\mathrm{eqv}} \cdot \pi$. At each scale, the maxima along the profile are identified and if the number of maximums remains constant for 5 iterations the location of the first maximum is the initial seed point set $\mathbf{S}_I$. If this condition is not fulfilled for 25 iterations, the seed points have to be selected manually.

Automated identification of $\mathbf{S}_{\mathrm{II}}$ The second set of seed points is used in the case that a stem is missed during the tracing procedure which would result in an incomplete cen-

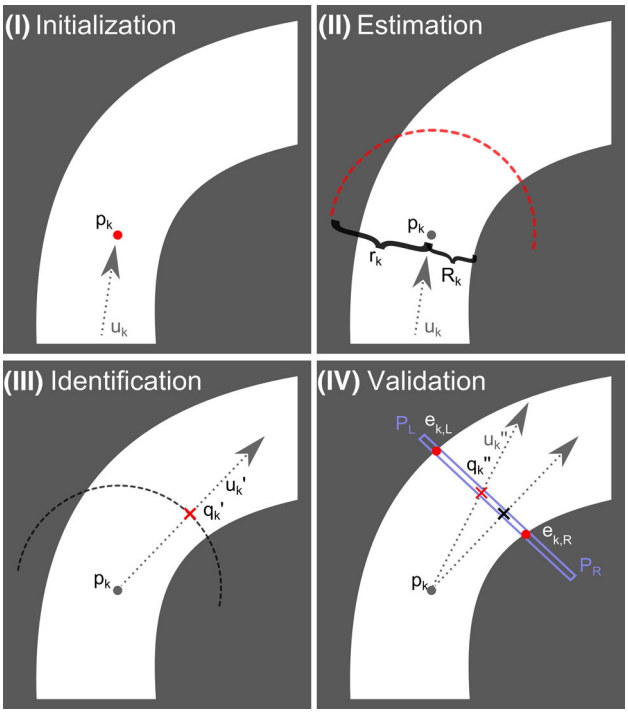


Fig. 3 The geometric determination and extraction of the local features at a centerline point

terline map of the plant. A quadratic grid is positioned in the center of the image, where the length of one square is set to approximately half of a silique length which were 75 pixels in this work [4]. Ridge points are detected along this grid. The set of valid ridge points is refined by not considering points below the lowest (y-position) seed point in $\mathbf{S}_\mathrm{I}$ and where the intensity value is below 20 % of the maximum intensity value. This approach helps us to overcome one of the limitations described in our previous work [4], which is the need for a manual selection of additional seed points in case the tracing procedure turns out incomplete.

3.2.2 Iterative tracing

The tracing procedure can be split in four phases: *Initialization*, *estimation*, *identification* and *validation* [8]. These phases are iteratively executed until a certain stopping criteria are fulfilled, discussed in the next section. A geometric description of these phases is shown in Fig. 3.

Initialization Each iteration k is initialized at a point $\mathbf{p}_k$. While in the beginning, the set of seed points $\mathbf{S}_I$ is used as initial points, additional points are determined during the tracing procedure when detecting neighboring centerline points or branching points.

Estimation The location of the next centerline point $\mathbf{p}_{k+1}$ is estimated by a dynamic semi-circular search window (see Fig. 4) which guarantees a constant look-ahead distance in

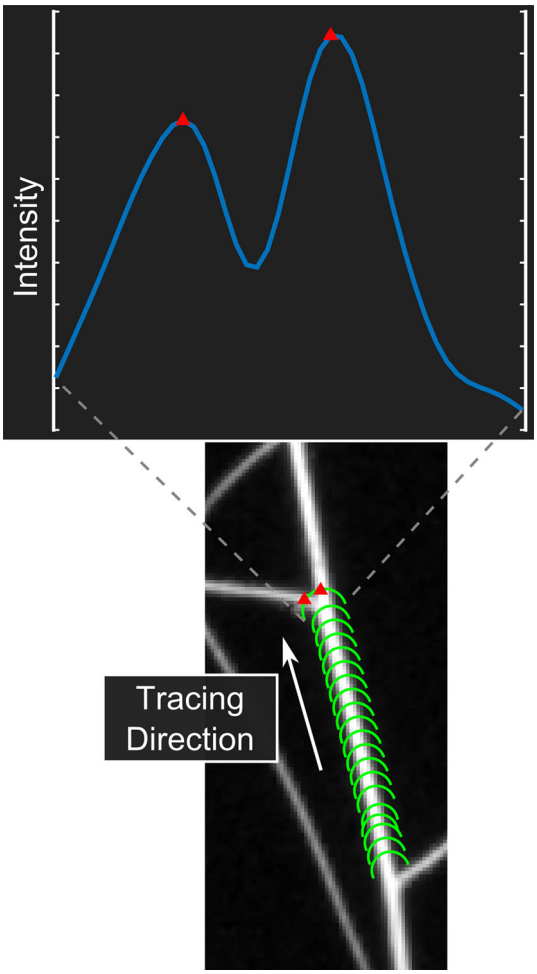


Fig. 4 A semi-circular template is propagating along a stem until a branching point is detected. The intensity profile along the semi-circular circumference of the search window is shown for the detection of the branching point. The *triangles* indicate the identified candidate points

all directions. The gray-value profile $\mathbf{G}\left(r_k, \mathbf{p}_k\right)$ along the circumference of a semi-circle with radius r_k at $\mathbf{p}_k$ in the image $\mathbf{I}$ is defined according to (4).

Since $\mathbf{I}$ is defined in the discrete space, the values of c_i are determined using the nearest pixel values. The gray values along the circumference are smoothed using a Gaussian filter (size $= 13$, $\sigma = 2$). The radius r_k must be defined big enough to cover different stem widths and must be small enough to not detect points along neighboring structures. Therefore r_k is adapted dynamically at each iteration by [8]:

$$r_k = \rho \cdot \left[\max\{R_k, R_{k+1}\}\right]. \tag{7}$$

R_k and R_{k+1} are the radius of the stem at the current and the next centerline point position. The constant factor ρ should be defined bigger than 1 to guarantee the coverage of the whole width of a stem. In this work, ρ was set to 2.

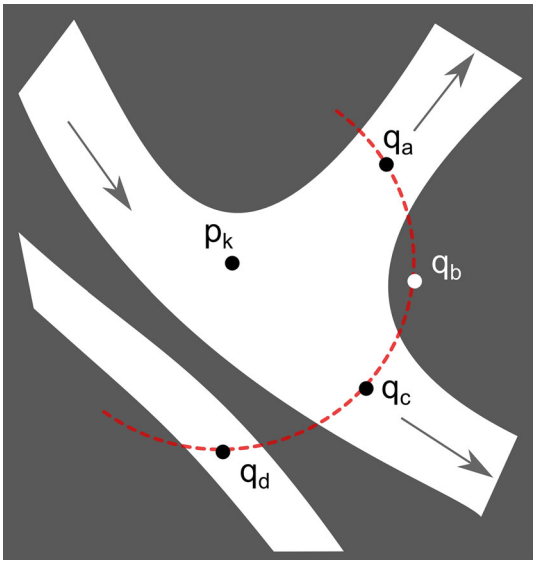


Fig. 5 Different types of candidate points can be identified: stem points ($\mathbf{q}_a$, $\mathbf{q}_c$), outlier points ($\mathbf{q}_d$) and non-stem points ($\mathbf{q}_b$)

Identification The stems appear as bright structures on a dark background and thus the intensity profile forms a Gaussian-like shape. To identify a centerline candidate point $\mathbf{q}_k$, the intensity profile $\mathbf{G}\left(r_k, \mathbf{p}_k\right)$ is probed for local maxima (see Fig. 4). Figure 5 shows an overview of different points which have to be considered and differentiated along the circumference of the search window: Stem points, outlier points and non-stem points.

Validation Each candidate point $\mathbf{q}_k$ is validated by determining its local features. The accuracy of the features depends on the tracing direction as well as on small intensity variations across the stems. To gain a higher accuracy regarding the location, the first estimate $\mathbf{q}_k^{'}$ is refined as follows (see Fig. 3 for illustration) [8,17,44]:

1. Calculate tracing direction for $\mathbf{q}_k^{'}$

$$\mathbf{u}_k^{'} = \frac{\mathbf{q}_k^{'} - \mathbf{p}_k}{\|\mathbf{q}_k^{'} - \mathbf{p}_k\|}. \tag{8}$$

2. Detect edge points ($e_{k,L}$ and $e_{k,R}$) along two linear intensity profiles $\mathbf{P}_\mathrm{L}$ and $\mathbf{P}_\mathrm{R}$ perpendicular to the direction $\mathbf{u}_k^{'}$ by finding the gradient's maxima. The length $\mathbf{P}_\mathrm{L}$ and $\mathbf{P}_\mathrm{R}$ is chosen to be the same value as r_k.
3. The location of $\mathbf{q}_k^{''}$ is calculated as the medial point between the edge points. The radius R_{k+1} is determined.
4. The final tracing direction $\mathbf{u}_k$ is refined by

$$\mathbf{u}_k = \frac{\mathbf{q}_k^{''} - \mathbf{p}_k}{\|\mathbf{q}_k^{''} - \mathbf{p}_k\|}. \tag{9}$$

Table 1 Geometrical and topological properties of a **STEL** are determined during extracting the centerline by iterative tracing

Geometrical properties	
$\mathbf{p}_k$	Location of current $\mathbf{STEL}_k$ in pixel units
$\mathbf{u}_k$	Tracing direction (unit vector) at $\mathbf{p}_k$
R_k	Radius of the stem at $\mathbf{p}_k$
$\mathbf{e}_{k,\mathrm{L}}$, $\mathbf{e}_{k,\mathrm{R}}$	Edge point positions
s_k	Normalized intensity level at $\mathbf{p}_k$
γ_k	Percent dynamic range from perpendicular intensity profile at $\mathbf{p}_k$
Topological properties	
ID	A unique ID for each STEL
Parent	ID of the previous STEL
Type	Regular (one child), branching/crossing (two or more children), root (no parent/seed point), termination (no child), outlier (no parent)

The final position of the validated point $\mathbf{p}_{k+1}$ can be influenced by a parameter α (=0.9 in this work) which regulates the step size according to:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha \cdot \mathbf{u}_k. \tag{10}$$

A centerline point, denoted as **STEL** (STem-ELement), comprises local geometrical as well as topological properties which are extracted during the tracing process. The properties are defined in Table 1.

3.2.3 Centerline representation and stopping criteria

To avoid tracing of already traced stems, a centerline representation is used. The centerline representation is a binary image with the size of the plant's image $\mathbf{I}$ (initialized with pixel values equal zero). After each iteration, the pixel values of the current stem segments are set to one. Two neighboring STELs are thereby connected using the Bresenham line drawing algorithm [10].

The centerline of the plant's stems is extracted by executing the previously described steps iteratively until at least one of the following criteria is fulfilled:

- Any of the pixels of the search window is outside the image range.
- No *valid* candidate points are identified.
- More than one valid candidate point is identified (branching/crossing).
- Connection between the current point $\mathbf{p}_k$ and one of the current candidate points $\mathbf{q}_k$ intersects the actual centerline image.

- Percent dynamic range falls below 10 %.

3.3 Plant reconstruction

The centerline extraction process leads to a set of unconnected centerline stem segments whose geometrical and partial topological properties are given by their STELs. The branching topology of *A. thaliana* is not known and therefore cannot be described as an a priori model. Furthermore, the local knowledge at critical regions during tracing is not sufficient for an appropriate reconstruction during the centerline extraction process. This demands an additional step which is the reconstruction of the plant out of the centerline segments from the tracing procedure. After reconstructing the *realistic* architecture out of the 2D information, the quantitative traits are calculated.

3.3.1 Grouping and labeling

In this work, we propose a hierarchical, graph-based reconstruction process. An overview of this process is shown in Fig. 6.

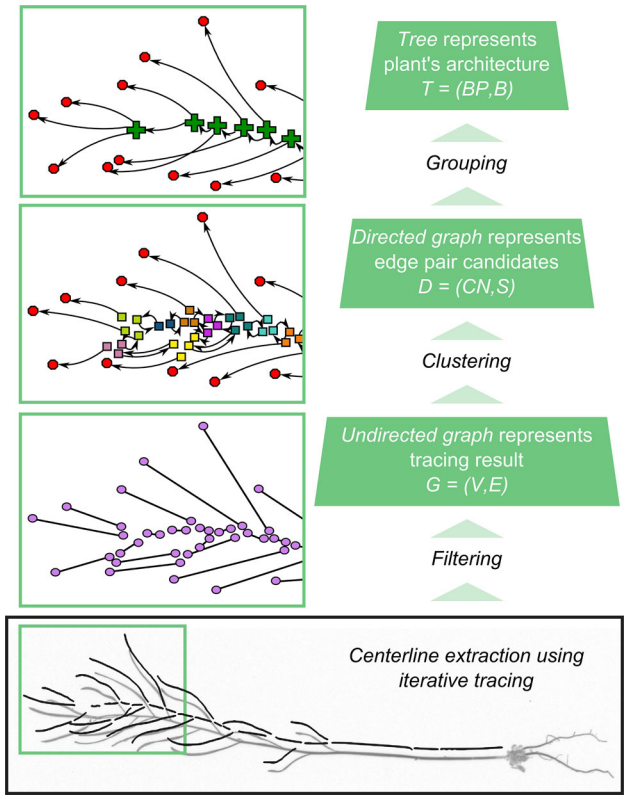


Fig. 6 Hierarchical reconstruction of the plants' architecture. The base layer consists of the centerline segments from the tracing procedure. The *top* layer represents the plant with use of branches and branching points. The root node is shown by a *triangle*, end nodes are represented by *circles* and branching points are represented by "+"-signs

Filtering Centerline segments which are shorter than 5 pixel are removed. The remaining centerline segments are represented as nodes $\mathbf{V}$ and possible connections between two segments as edges $\mathbf{E}$. The set of edges $\mathbf{E}$ and nodes $\mathbf{V}$ is represented as an undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$.

Clustering Assuming that each centerline segment $\mathbf{V}_i$ can be connected to any other centerline segment $\mathbf{V}_j$, we obtain a complete graph. As a segment at the very top of the image is unlikely to be connected to a segment at the very bottom we are looking for a representation with a reduced set of possible connections. At this level, each centerline segment is represented by an edge $\mathbf{S}_i$ and a set of possible connections between adjacent centerline segments as (cluster) node $\mathbf{CN}_i$. The set of possible connections is reduced by exploiting the topological information gained during tracing (STEL relations). Furthermore, we defined heuristic rules and an Euclidean distance to integrate unrelated segments which are present after tracing.

Grouping The classification of merging segments to a path is based on continuity principles as we can expect smooth variation in terms of shape and texture along the stems. Therefore, the cost term $c\left(\mathbf{S}_i, \mathbf{S}_j\right)$ is proposed based on the edge direction similarity $\theta\left(\mathbf{S}_i, \mathbf{S}_j\right)$. To strongly weight resulting tortuous paths, the tortuosity $\tau\left(\mathbf{S}_i, \mathbf{S}_j\right)$ of the resulting path is weighted with the exponential function and multiplied by the linkage distance d_l between $\mathbf{S}_i$ and $\mathbf{S}_j$. Hence, the cost function therefore is defined as

$$c\left(\mathbf{S}_i, \mathbf{S}_j\right) = \theta\left(\mathbf{S}_i, \mathbf{S}_j\right) + d_l \cdot e^{\tau\left(\mathbf{S}_i, \mathbf{S}_j\right)}. \tag{11}$$

The edge direction similarity $\theta\left(\mathbf{S}_i, \mathbf{S}_j\right)$ between two segments $\mathbf{S}_i$ and $\mathbf{S}_j$ with corresponding directions $\mathbf{u}_i$ and $\mathbf{u}_j$ is defined as in [45]

$$\theta\left(\mathbf{S}_i, \mathbf{S}_j\right) = \arccos\left(\frac{\mathbf{u}_i \cdot \mathbf{u}_j}{|\mathbf{u}_i| \cdot |\mathbf{u}_j|}\right). \tag{12}$$

The tortuosity of the resulting path is calculated as the ratio between the path length l_{ij} and the Euclidean distance d_{ij} between the resulting path's endpoints

$$\tau\left(\mathbf{S}_i, \mathbf{S}_j\right) = \frac{l_{12}}{d_{12}}. \tag{13}$$

The grouping process follows a trace-back principle which means that the tree is constructed from the exterior regions to the interior regions. Furthermore, the cluster nodes are visited from less complex to more complex nodes and the cost function is determined subsequently. The number of possible edge-pairs in a cluster is defined as the grade of complexity. A tree $\mathbf{T} = (\mathbf{BP}, \mathbf{B})$ containing branches $\mathbf{B}_i$ and branching points $\mathbf{BP}_i$ is iteratively build by minimizing the cost function in (11).

3.3.2 Geometrical and topological traits

After reconstructing the plant's realistic architecture, the quantitative traits are determined. Topological traits are used to quantify the "network topology" of the plant and describe the relation between different plant parts. Therefore, we define different branch types from which their occurrence is determined:

- N_{MS}: number of (main) stems (MS) which originate from the rosette area.
- N_{SIL}: number of terminating branches (siliques—SIL) containing the seeds and having a characteristic width variation.
- N_{L}: number of terminating branches (leaf—L) which are not classified as siliques.
- N_{SB}: number of side branches (SB) originating from a main stem or another side branch.

The geometrical traits are calculated for each branch $\mathbf{B}_i$:

- l_{B}: path length of the branch's centerline calculated by the number of odd N_{o} and even N_{e} Freeman (8-directional) chain codes along the centerline:

$$l_{\mathrm{B}} = N_{\mathrm{e}} + N_{\mathrm{o}} \cdot \sqrt{2}. \tag{14}$$

- d_{B}: the Euclidean distance between the branch's end points [26].

Silique detection The majority of the exterior parts of mature *A. thaliana* are siliques. Furthermore, siliques generally tend to have a higher variation in width and the lengths of siliques on one plant is more evenly distributed than the length of other leaf parts. For this reason, the path length, the average width and the standard deviation of the width of the exterior branches combined with outlier detection methods are used to differentiate between siliques and other leafs. If the number of leafs on a plant is greater or equal 20, the robust *Minimum Covariance Determinant* estimator is used (outlier detection based on all three features) [38]. Otherwise the *Median Absolute Deviation* is used based on the branches' path lengths [37].

4 Experiments and evaluation

To evaluate the performance of the proposed phenotyping framework for mature plants, a set of 106 images was used (see Sect. 2). The extracted traits by the proposed plant phenotyping framework (PPFW) are compared to a semi-automatic analysis using *Fiji Simple Neurite Tracer* [24] (GT Fiji) and a manual phenotyping approach (GT UK). The manual analysis of the plants was done by one of our collaborators in Norwich, UK before the framework was developed. One expert determined hereby traits like the number of siliques or length of the siliques by visual inspection. To get a more detailed analysis, especially regarding the geometrical traits, the second analysis GT Fiji was carried out. Furthermore, the framework is analyzed regarding the throughput and the amount of manual corrections which are needed along the phenotyping pipeline.

4.1 Throughput of the proposed plant phenotyping framework

In addition to the accuracy of the extracted traits using a phenotyping framework also their need for manual interventions vs. the throughput of the system is crucial when it comes to a laboratory use. A manual correction to rise the throughput rate along the analysis is provided for the pre-processing as well as the centerline extraction process. For the step of plant reconstruction, only a manual quality control by visual inspection is provided and the user decides whether he accepts the result of the reconstruction process or if he wants to exclude the results from the trait analysis. The framework was implemented in the MathWorks® Matlab® environment.[1]

4.1.1 Pre-processing

During the automated ROI identification as well as the automated spatial calibration, there were no manual corrections needed for all 106 images.

4.1.2 Centerline extraction

During the automated seed point identification for $\mathbf{S}_I$ (see Sect. 3.2.1), 7 clicks are needed to overcome false positives and negatives. The F-value without a manual correction was 97.01 %.

4.1.3 Plant reconstruction

The plant reconstruction process is mainly evaluated for the completeness of the reconstruction. In some cases, wrongly made local decisions in a cluster node lead to missing of some plant parts as the grouping process yields to multiple separated (sub)trees instead of one tree per main stem. Hence a quantitative trait extraction would not make sense in those cases. Therefore, the completeness of the reconstruction is evaluated by visual inspection of a resulting image overlaid on the original image (see Fig. 10). Incompleteness was noticed in 16 out of 106 images, which corresponds to a throughput rate of 84.91 %.

[1] http://www.mathworks.com/products/matlab/.

Table 2 Comparison regarding the branches' lengths (path length l_B) and number of branches (N_B)

Method	Trait	Branch type		
		MS	SB	SIL
GT Fiji	$l_B\ [Px]$	1971 (±737)	615 (±408)	179 (±24)
	N_B	102	148	3051
PPFW	$l_B\ [Px]$	2017 (±762)	236 (±188)	168 (±21)
	N_B	100	545	2701
GT UK	$l_B\ [Px]$	–	–	129 (±23)
	N_B	–	–	2912

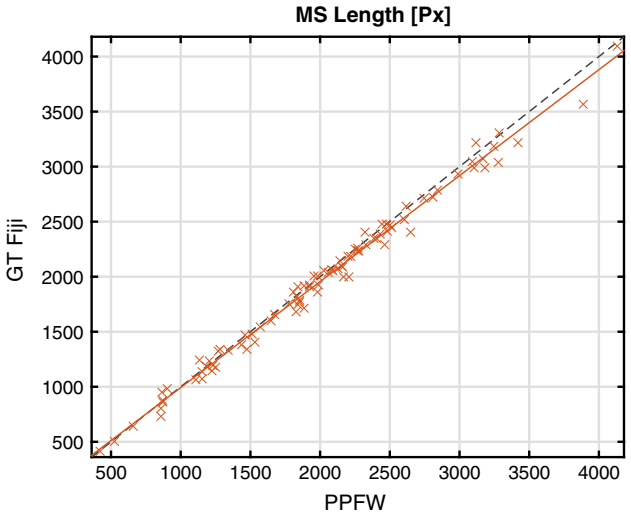


Fig. 7 Scatter plot comparing the GT Fjiji and the PPFW MS lengths (Pearson's $r = 0.99$ ($p < 0.01$))

4.2 Measurement of the plant size

The accuracy of the final traits is evaluated using the 90 images where the grouping process was judged positive. The lengths of certain branch types (MS, SB and SIL) are averaged per plant and compared to GT Fiji. Furthermore, the lengths of the siliques are also compared to manual phenotyping (abbreviated with GT UK) as this trait was already evaluated for the given dataset before the framework was developed.

4.2.1 Comparison with GT Fiji

The evaluation concerning the stems' lengths is summarized in Table 2. The average relative error between the results of PPFW and GT Fiji concerning the MS length is (3.64 ± 3.25) %. The average relative error concerning the siliques' lengths between these two measurements was determined to be (6.57 ± 4.71) %. Scatter plots showing the individual measurements for this comparisons are shown in Figs. 7 and 8.

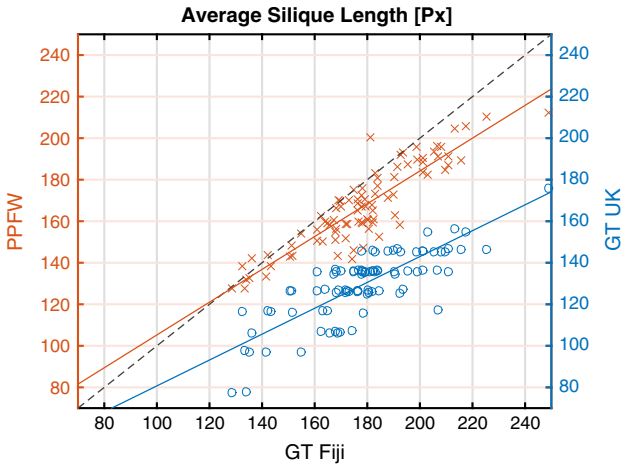


Fig. 8 Scatter plot comparing the GT Fiji, GT UK and the PPFW siliques' lengths

The evaluation regarding the side branches shows a much higher deviation between GT Fiji and PPFW (see Table 2). This is explained by extensive overlapping which are present in the very exterior regions (e.g., siliques) of the plants. Thereby the overlapping is wrongly resolved as one new (mostly short) side branch and terminating branches instead of only terminating branches. During the trait extraction, this leads to an overestimation of number of side branches and an underestimation of number of siliques (see Table 2).

Furthermore, this is also an explanation of the underestimation of the siliques' lengths which is noticeable in the scatter plot (Fig. 8) and the underestimation of the averaged side branches' lengths in Table 2.

4.2.2 Comparison with GT UK

During the manual phenotying study, the siliques' lengths were determined for each plant. The average silique length for all plants was hereby determined with (128.73 ± 22.54) pixels and all together 2912 siliques were counted (see Table 2). The scatter plot in Fig. 8 shows a comparison between the GT UK and the GT Fiji. One can notice an offset of the linear regression line which shows an underestimation of the siliques' lengths during the manual phenotyping. This effect can be explained by different definitions of silique length. While in the GT Fiji approach, the silique length was defined from the branching point to the tip, it was measured from the beginning of the carpel to the silique's tip in the GT UK approach. Apart from this it can be noticed that the measurement during the manual phenotyping approach has a lower resolution compared to the results of GT Fiji. The Pearson's correlation coefficient r between GT Fiji and

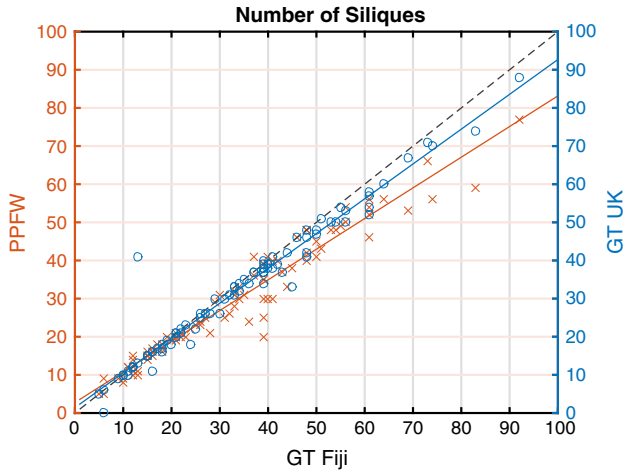


Fig. 9 The number of siliques are compared between GT Fiji, GT UK and PPFW

the PPFW for the siliques' lengths was determined with $r = 0.90\,(p < 0.01)$. Between the independent GT measurements, r was determined with $r = 0.82\,(p < 0.01)$. This shows a high inter-operator variability for the given dataset.

4.3 Number of siliques

To identify the number of siliques on a plant is of interest for biologists. As the siliques contain the seeds, their size and number are an important indicator in terms of reproduction of a plant. To evaluate the silique counting ability of PPFW, we determined the relative error compared to the GT Fiji and to GT UK. The mean relative error between GT Fiji and PPFW was determined (11.81 ± 10.14) % and (11.08 ± 11.16) % between GT UK and PPFW. The scatter plot for comparing the individual number of siliques for each plant is shown in Fig. 9. The determined number of siliques is shown in Table 2. It can be noticed that the number of siliques is underestimated by the proposed framework. One can also notice that the precision decreases with an increasing number of siliques on a plant. This let us conclude that an automated analysis of this trait is limited to plants with lower morphological complexity.

4.4 Images for quality control

To allow an effective judging of the quantitative trait extraction process, the resulting labeled skeleton and the critical points identified are overlaid on the (inverted) original images. Five exemplary results are shown in Fig. 10. Starting points are marked with a green triangle near the rosette, termination points with a red triangle and branching points with a magenta colored "+"-sign. Side branches are labeled in different colors while siliques are colored with a light blue and other leafs with a dark blue. The corresponding table of traits for the images in Fig. 10 is shown in Table 3.

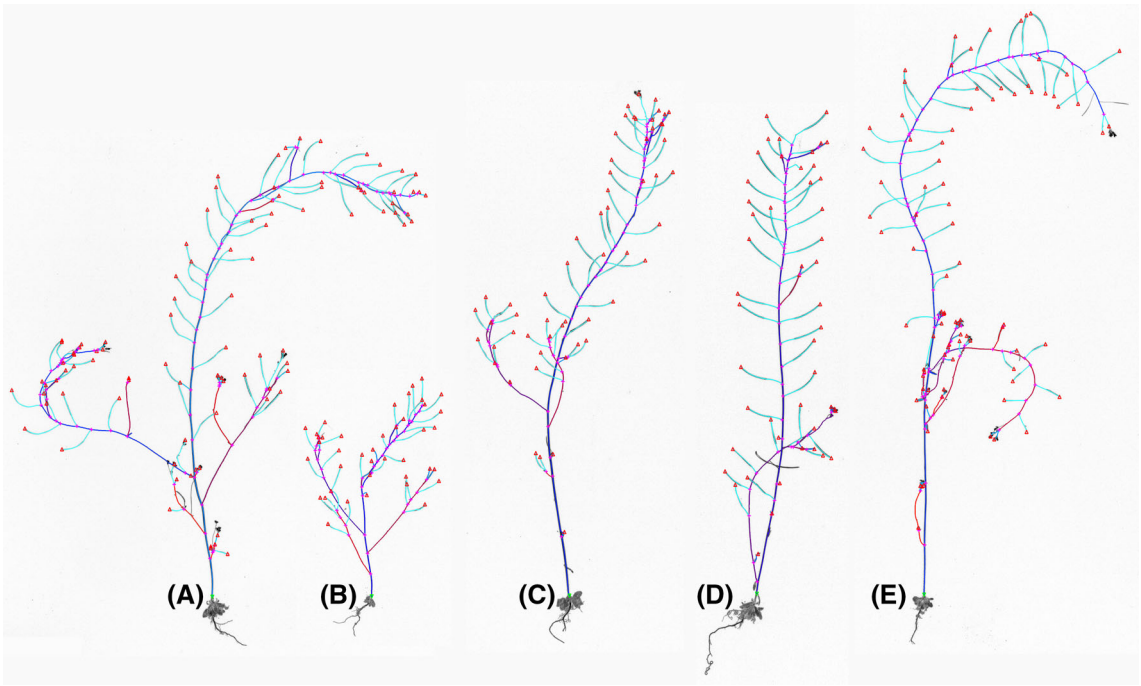


Fig. 10 Exemplary output for five images of the dataset which are processed with the PPFW. Table 3 summarizes the plants' traits determined by GT Fiji and PPFW

Table 3 Comparison of 5 exemplary images of the dataset between the GT Fiji and the results of PPFW

Plant	Method	N_{MS}	N_{SB}	N_{SIL}	$\bar{l}_{MS}$ $[Px]$	$\bar{l}_{SIL}$ $[Px]$
A	GT Fiji	1	4	64	3306	206
	PPFW	1	17	56	3282	196
B	GT Fiji	1	3	33	1185	143
	PPFW	1	7	28	1185	138
C	GT Fiji	1	3	48	2786	178
	PPFW	1	7	42	2844	176
D	GT Fiji	1	1	39	2520	217
	PPFW	1	5	35	2607	206
E	GT Fiji	1	1	48	4097	195
	PPFW	1	13	48	4130	187

5 Conclusion

A framework to analyze the architecture of mature *A. thaliana* plants based on 2D single-shot images is presented. Quantitative traits describing the geometry as well as the topology of the plant are extracted and compared to a manual phenotyping approach as well as a semi-automatic image-based phenotyping approach. With the evaluation comparing the extracted geometrical and topological traits, we showed the benefits of tracing algorithms for centerline extraction when applied to image-based plant phenotyping problems in which curvilinear/tubular structures are analyzed.

The accuracy of the reconstruction process from the real branching network out of the 2D images is limited by the quality of the images itself as well as by the morphological complexity of the plants, i.e., limited number of branching and crossing regions in a certain area. The evaluations showed the ability of the PPFW to quantify geometrical properties of mature plants with a high accuracy. The weakest link in the proposed image processing chain is the hierarchical grouping and labeling process for the extracted center line segments from the tracing procedure. We encountered difficulties to gain reliable topological properties if multiple overlaps occur very close to each other and the involved segments are short. In this case, a local decision in a cluster node can lead to an error propagation and a concluding unsatisfying reconstruction result. This problem should be addressed in future works.

The reliable extraction of topological traits would further increase the number of quantitative traits as they could be combined with geometrical traits. Such traits could for example be the bifurcation ratio or the inter-nodal distance between different branch types.

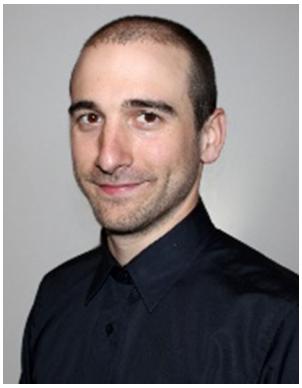
The image-based phenotyping of plants in the mature state can play an important role in large-scale phenotyping studies. Combined with approaches to phenotype plants in the early stages of growing this would enable plant phenotyping over the whole life-cycle.

Acknowledgments We want to thank Svante Holm (Mid Sweden University, SE) and Alison Anastasio (University of Chicago, US) for planting and harvesting the plants, Man Yu and Andrew Davis for taking the photos and Benjamin Brachi (Bergelson Lab, University of Chicago, US) for his valuable inputs and support along the stages of development.

References

1. Al-Tam, F., Adam, H., Anjos, A., Lorieux, M., Larmande, P., Ghesquiere, A., Jouannic, S., Shahbazkia, H.: P-TRAP: a panicle trait phenotyping tool. BMC Plant Biol. **13**(1), 122 (2013)
2. Armengaud, P., Zambaux, K., Hills, A., Sulpice, R., Pattison, R.J., Blatt, M.R., Amtmann, A.: EZ-Rhizo: integrated software for the fast and accurate measurement of root system architecture. Plant J. **57**(5), 945–956 (2009)
3. Arvidsson, S., Perez-Rodriguez, P., Mueller-Roeber, B.: A growth phenotyping pipeline for Arabidopsis thaliana integrating image analysis and rosette area modeling for robust quantification of genotype effects. New Phytol. **191**(3), 895–907 (2011)
4. Augustin, M., Haxhimusa, Y., Busch, W., Kropatsch, W.G.: Image-based phenotyping of the mature Arabidopsis shoot system. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) Computer Vision—ECCV 2014 Workshops. Lecture Notes in Computer Science, vol. 8928, pp. 231–246. Springer, Berlin (2015)
5. Basu, P., Pal, A., Lynch, J.P., Brown, K.M.: A novel image-analysis technique for kinematic study of growth and curvature. Plant Physiol. **145**(2), 305–316 (2007)
6. Benmansour, F., Fua, P., Türetken, E.: Automated reconstruction of tree structures using path classifiers and mixed integer programming. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 566–573 (2012)
7. Boroujeni, F.Z., Wirza, R., Rahmat, O., Mustapha, N., Affendey, L.S., Maskon, O.: Automatic selection of initial points for exploratory vessel tracing in fluoroscopic images. Def. Sci. J. **61**, 443–451 (2011)
8. Boroujeni, F.Z., Rahmat, O., Wirza, R., Mustapha, N., Affendey, L.S., Maskon, O.: Coronary artery center-line extraction using second order local features. Comput. Math. Methods Med. **2012** (2012). doi:10.1155/2012/940981
9. Brachi, B., Morris, G.P., Borevitz, J.O.: Genome-wide association studies in plants: the missing heritability is in the field. Genome Biol. **12**(10), 232 (2011)
10. Bresenham, J.E.: Algorithm for computer control of a digital plotter. IBM Syst. J. **4**(1), 25–30 (1965)
11. Cobb, J.N., DeClerck, G., Greenberg, A., Clark, R., McCouch, S.: Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. **126**(4), 867–887 (2013)
12. Crowell, S., Falcão, A.X., Shah, A., Wilson, Z., Greenberg, A.J., McCouch, S.R.: High-resolution inflorescence phenotyping using a novel image-analysis pipeline, panorama. Plant Physiol. **165**(2), 479–495 (2014)
13. Delibasis, K.K., Kechriniotis, A.I., Tsonos, C., Assimakis, N.: Automatic model-based tracing algorithm for vessel segmentation and diameter estimation. Comput. Methods Programs Biomed. **100**(2), 108–122 (2010)
14. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G., Barman, S.A.: Blood vessel segmentation

methodologies in retinal images—a survey. Comput. Methods Programs Biomed. **108**(1), 407–433 (2012)
15. French, A.P., Ubeda-Tomas, S., Holman, T., Bennett, M., Pridmore, T.: High-throughput quantification of root growth using a novel image-analysis tool. Plant Physiol. **150**(4), 1784–1795 (2009)
16. Godin, C., Costes, E., Sinoquet, H.: A method for describing plant architecture which integrates topology and geometry. Ann. Botany **84**(3), 343–357 (1999)
17. Huang, Y., Zhang, J., Huang, Y.: An automated computational framework for retinal vascular network labeling and branching order analysis. Microvascu. Res. **84**(2), 169–177 (2012)
18. Humplík, J.F., Lazár, D., Fürst, T., Husičková, A., Hýbl, M., Spíchal, L.: Automated integrative high-throughput phenotyping of plant shoots: a case study of the cold-tolerance of pea (Pisum sativum L.). Plant Methods **11**(1), 1–11 (2015)
19. The Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**(6814), 796–815 (2000)
20. Klodt, M., Cremers, D.: High-resolution plant shape measurements from multi-view stereo reconstruction. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) Computer Vision—ECCV 2014 Workshops. Lecture Notes in Computer Science, vol. 8928, pp. 174–184. Springer, Berlin (2015)
21. Li, L., Zhang, Q., Huang, D.: A review of imaging techniques for plant phenotyping. Sensors **14**(11), 20078–20111 (2014)
22. Lin, K.S., Tsai, C.L., Tsai, C.H., Sofka, M., Chen, S.J., Lin, W.Y.: Retinal vascular tree reconstruction with anatomical realism. IEEE Trans. Biomed. Eng. **59**(12), 3337–3347 (2012)
23. Lobet, G., Draye, X., Périlleux, C.: An online database for plant image analysis software tools. Plant Methods **9**(38), 1–7 (2013)
24. Longair, M.H., Baker, D.A., Armstrong, J.D.: Simple neurite tracer: Open source software for reconstruction, visualization and analysis of neuronal processes. Bioinformatics **27**(17), 2453–2454 (2011)
25. Martinez-Perez, M.E., Hughes, A.D., Stanton, A.V., Thom, S.A., Bharath, A.A., Parker, K.H.: Retinal blood vessel segmentation by means of scale-space analysis and region growing. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 90–97 (1999)
26. Martinez-Perez, M.E., Hughes, A.D., Stanton, A.V., Thom, S.A., Chapman, N., Bharath, A.A., Parker, K.H.: Retinal vascular tree morphology: a semi-automatic quantification. IEEE Trans. Biomed. Eng. **49**(8), 912–917 (2002)
27. Meijering, E.: Neuron tracing in perspective. Cytom. Part A **77A**(7), 693–704 (2010)
28. Minervini, M., Abdelsamea, M.M., Tsaftaris, S.A.: Image-based plant phenotyping with incremental learning and active contours. Ecol. Inform. **23**, 35–48 (2014)
29. Minervini, M., Giuffrida, M.V., Tsaftaris, S.: An interactive tool for semi-automated leaf annotation. In: Tsaftaris, S.A., Scharr, H., Pridmore, T. (eds.) Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP), pp. 6.1-6.13. BMVA Press (2015)
30. Müller-Linow, M., Pinto-Espinosa, F., Scharr, H., Rascher, U.: The leaf angle distribution of natural plant populations: assessing the canopy with a novel software tool. Plant Methods 11(1) (2015). doi:10.1186/s13007-015-0052-z
31. Mutka, A., Bart, R.: Image-based phenotyping of plant disease symptoms. Front. Plant Sci. **5**, 734 (2015). doi:10.3389/fpls.2014.00734
32. Naeem, A., French, A.P., Wells, D.M., Pridmore, T.: High-throughput feature counting and measurement of roots. Bioinformatics **27**(9), 1337–1338 (2011)
33. Nguyen, U.T.V., Bhuiyan, A., Park, L.A.F., Ramamohanarao, K.: An effective retinal blood vessel segmentation method using multi-scale line detection. Pattern Recogn. **46**(3), 703–715 (2013)
34. Pape, J.M., Klukas, C.: 3-D histogram-based segmentation and leaf detection for rosette plants. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) Computer Vision—ECCV 2014 Workshops. Lecture Notes in Computer Science, vol. 8928, pp. 61–74. Springer, Berlin (2015)
35. Pound, M.P., French, A.P., Murchie, E.H., Pridmore, T.P.: Automated recovery of three-dimensional models of plant shoots from multiple color images. Plant Physiol. **166**(4), 1688–1698 (2014)
36. Robben, D., Türetken, E., Sunaert, S.: Simultaneous segmentation and anatomical labeling of the cerebral vasculature. In: International Conference on Medical Image Computing and Computer Assisted Intervention (2014)
37. Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. Wiley, New York (1987)
38. Rousseeuw, P.J., Driessen, K.V.: A fast algorithm for the minimum covariance determinant estimator. Technometrics **41**(3), 212–223 (1999)
39. Rousseau, D., Chéné, Y., Belin, E., Semaan, G., Trigui, G., Boudehri, K., Franconi, F., Chapeau-Blondeau, F.: Multiscale imaging of plants: current approaches and challenges. Plant Methods **11**(1), 1–9 (2015)
40. Slovak, R., Göschl, C., Su, X., Shimotani, K., Shiina, T., Busch, W.: A scalable open-source pipeline for large-scale root phenotyping of Arabidopsis. Plant Cell **26**(6), 2390–2403 (2014)
41. Sozzani, R., Benfey, P.: High-throughput phenotyping of multicellular organisms: finding the link between genotype and phenotype. Genome Biol. **12**(3), 219–225 (2011)
42. Spalding, E.P., Miller, N.D.: Image analysis is driving a renaissance in growth measurement. Curr. Opin. Plant Biol. **16**(1), 100–104 (2013)
43. Subramanian, R., Spalding, E.P., Ferrier, N.J.: A high throughput robot system for machine vision based plant phenotype studies. Mach. Vis. Appl. **24**(3), 619–636 (2013)
44. Sun, Y.: Automated identification of vessel contours in coronary arteriograms by an adaptive tracking algorithm. IEEE Trans. Med. Imaging **8**(1), 78–88 (1989)
45. Türetken, E., Gonzalez, G., Blum, C., Fua, P.: Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors. Neuroinformatics **9**(2–3), 279–302 (2011)
46. Walter, A., Liebisch, F., Hund, A.: Plant phenotyping: from bean weighing to image analysis. Plant Methods 11(1) (2015). doi:10.1186/s13007-015-0056-8
47. Weigel, D.: Natural variation in arabidopsis: from molecular genetics to ecological genomics. Plant Physiol. **158**(1), 2–22 (2012)
48. Yin, Y., Adel, M., Bourennane, S.: Retinal vessel segmentation using a probabilistic tracking method. Pattern Recogn. **45**(4), 1235–1244 (2012)
49. Zana, F., Klein, J.C.: Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. Trans. Image Process. **10**(7), 1010–1019 (2001)
50. Zhang, Y., Zhou, X., Degterev, A., Lipinski, M., Adjeroh, D., Yuan, J., Wong, S.: A novel tracing algorithm for high throughput imaging screening of neuron-based assays. J. Neurosci. Methods **160**(1), 149–62 (2007)
51. Zheng, Y., Gu, S., Edelsbrunner, H., Tomasi, C., Benfey, P.: Detailed reconstruction of 3D plant root shape. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2026–2033 (2011)



Marco Augustin received his BSc and MSc in medical informatics from the Technical University of Vienna, Austria, in 2010 and 2014, respectively. Currently, he is a Ph.D. student at the Center for Medical Physics and Biomedical Engineering at the Medical University of Vienna, Austria. His research interests include optical imaging, image processing and pattern recognition particularly in life sciences.

Yll Haxhimusa received a Dipl.Ing. degree in Electrical and Computer Engineering from the University of Prishtina, Kosova. He obtained his Ph.D. degree in Computer Science from the Vienna University of Technology in 2006. In 2007, he became a postdoctoral associate at Purdue University (USA) in the department of psychological sciences. Dr. Haxhimusa's research interests span from computer vision, pattern recognition and machine learning, to human problem solving. His current interest is the usage of machine learning and pattern recognitions methods in personalized medicine. He is an IEEE and an IAPR member.

Wolfgang Busch received his diploma in biology from the University of Tübingen (Germany). He obtained his Ph.D. at the Max Planck Institute for Developmental Biology in Tübingen in 2008. In 2008, he became a postdoctoral associate at Duke University (USA) in the department of biology and the center of systems biology. Since 2011, he is a group leader at the Gregor Mendel Institute (Austria) where he seeks to comprehend how plant growth and development are quantitatively regulated at the molecular level. His achievements were recognized with multiple awards including the President's Medal of the Society of Experimental Biology.

Walter G. Kropatsch is full professor at Vienna University of Technology since 1990. He received his Dipl.Ing. (1975) and his Dr.techn. (1982) from the Technical University of Graz. In addition, he received the maître d'informatique from the University in Grenoble (1976). His interest in image pyramids goes back to 1984 when he spent one year at the Center for Automation Research of the University of Maryland on invitation of Prof. Azriel Rosenfeld. Since then he explored different variants of hierarchical data and processing structures. In collaboration with Peter Meer, Annick Montanvert and Jean-Michel Jolion, he extended the scope of pyramids to graphs and then to plane graphs for which the first proof of topology preservation could be shown. He edited and contributed to more than 30 books, authored 39 articles in peer reviewed journals and published more than 200 technical papers in conferences and workshops. His current interest are in the extension of the pyramidal concept towards dynamical hierarchical description of the 3D moving objects in their environment and its derivation from one or more 2D image sequences. In this high-dimensional context, topology receives more importance. He served the IAPR in many positions, was awarded IAPR fellowship in 1994 and was its president from 2004 to 2006. Together with J.-M. Jolion, he initiated the IAPR TC15 on graph-based representations in 1996.