FROM INDIVIDUAL INTENSITY VOXEL DATA TO INTER-INDIVIDUAL PROBABILISTIC ATLASES OF BIOLOGICAL OBJECTS BY AN INTERLEAVED REGISTRATION-SEGMENTATION APPROACH

Felix Bollenbeck[1], Diana Weier[2], Wolfram Schoor[1] and Udo Seiffert[1]

[1]*Fraunhofer Institute for Factory Operation and Automation (IFF), Sandtorstr. 22, D-39106 Magdeburg, Germany*

[2]*Leibniz Institute of Plant Genetics and Crop Plant Research, Corrensstr. 3, D-06466 Gatersleben, Germany*

Keywords: Registration, Segmentation, 3-D Averaging Atlases, Plant Models.

Abstract: In this paper we describe an automated processing of plant serial section data for high-resolution 3-D models of internal structures. The processing pipeline includes standardization and registration of large image stacks as well as multiple tissue recognition by a joint registration-segmentation approach. By integrating segmented data from multiple individuals in a common reference, a statistical three-dimensional description is used to represent the inherent biodiversity amongst specimen. Inter-individual 3-D models are a novelty in the context of plant microscopy, and along with meaningful visualisation they deliver new insights into growth and development as well as provide a framework for the integration of functional data.

1 INTRODUCTION

The importance and the value of three-dimensional computer models of tissues or organs can undoubtedly be taken for granted. These models often serve as anatomical atlases facilitating the integration of heterogeneous experimental information, such as functional or gene-expression data, with spatial or even spatio-temporal reference.

The inherently existing inter-individual diversity leads to a certain divergence between the model and an arbitrary natural individual. We are working towards statistically valid models by means of barley grains, based on a multitude of digital 3-D models from histological serial section data.

The advantages in resolution of serial-section data for digital models on a micrometer scale generally come with high costs in 3-D reconstruction (registration) and labelling (segmentation), since the object of interest is essentially destroyed for digitization, delivering several thousands of separate and unlabeled raw images. Existing works for 3-D model generation from microscopic serial section imaging employ interactive techniques (Gubatz et al., 2007) as well as automated processing (Dercksen et al., 2008) utilizing supervised classification schemes for tissue labelling. These approaches have in common that models are generated from data from a single individual, ignoring inter-individual variances in histology and morphology, whereby validity and predictive power for

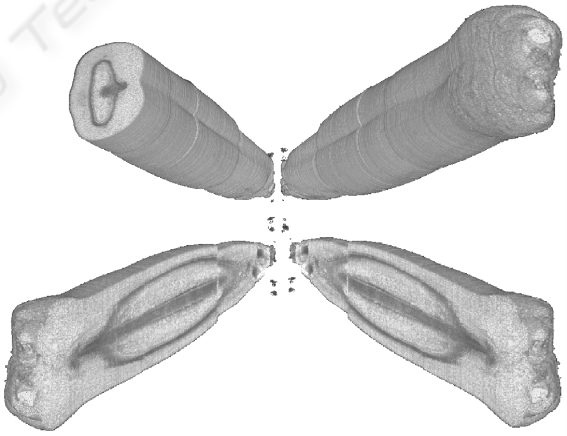


Figure 1: Perspectives of a digital barley grain. A stack of 2,217 individual section images (approx. 4 GB) which was standardized and registered to recompose an intact grain object visualized in a volume rendering, displaying the reconstructed histology in virtual lateral section.

Bollenbeck F., Weier D., Schoor W. and Seiffert U. (2009).
FROM INDIVIDUAL INTENSITY VOXEL DATA TO INTER-INDIVIDUAL PROBABILISTIC ATLASES OF BIOLOGICAL OBJECTS BY AN INTERLEAVED REGISTRATION-SEGMENTATION APPROACH.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 125-128
DOI: 10.5220/0001767001250128

new individuals is diminished. Probabilistic models, reflecting the diversity within biological phenotypes are highly feasible as reference atlases, apart from the possibility to quantify phenotypic variance itself.
Since high-throughput processing of section images is a prerequisite towards sufficiently large data-sets of anatomical data, we have developed a processing pipeline for an automated reconstruction of serially sectioned objects. By utilizing a joint segmentation–registration algorithm, we address the identification and labelling of relevant biological structures. In ongoing research we propose an initial modelling based on labelled voxel-data from a multitude of individuals for 3-D modelling comprising statistical descriptions of anatomical diversity and subsequent 3-D visualization.

2 INTERLEAVED REGISTRATION-SEGMENTATION OF SECTION DATA

The processing of image data towards inter-individual models is carried out from individual section images, to individual image stacks to statistical 3-D diversity models, where our proposed pipeline is based on interleaved segmentation and registration steps for iterative construction of inter-individual models.

Image Acquisition. For the imaging of functional units and tissues in microscopic plant organs, specimen of barley grains at distinct developmental time points are prepared for serial sectioning by embedding in a polymer and a contrasting agent. Embedded material is serially sectioned with a microtome into $3\mu m$ thick slices and digitized with a conventional light microscope ($1.83 \times 1.83 \mu m/px$), yielding roughly 2,000 images per grain, which are stored as *12-bit* grayscale image, since contrasting produces little color information.

Raw Image Segmentation and Registration. Manual handling and dust particles on the microscope slides produce high frequency noise as well as larger disturbances in section images. To remove these disturbances for the subsequent processing, the region-of-interest (*ROI*) is segmented, masked, and embedded in a uniform background. We employ a multi-resolution strategy using a variant of the *Level-Sets* approach (Caselles et al., 1997) for active contours segmentation suggested in (Li et al., 2005).
This initial segmentation of the section object is a preliminary for using the well-established *Principal Axis Transform* (*PAT*) (Alpert et al., 1990) for uniform image moments, since section slices appear in arbitrary orientation and positioning caused by manual preparation. Employing a *PAT* thereby serves as an initialization for the subsequent registration of the whole image stack (see (Schmitt et al., 2006) for a comprehensive study), allowing to re-establish three-dimensional coherence of sectioned objects, which is lost during sectioning.

Stack Registration and Tissue Recognition. By registering the full image stack, i.e. finding an optimal superposition over all images in the stack, the sectioned object is reconstructed. While finding an optimal affine transform, maximizing the correspondence of *all* stack images *at once* using numerical optimization schemes is computationally too complex, a pairwise sequential alignment of images is error-prone by propagating possible misalignments through the stack.
We use a spatially extended intensity-based image-to-image metric of a w_i, $i = 1,\ldots,N$ weighted sum of *SSD* values[1]

$$D(R \circ \varphi) := \sum_{i=2}^{M} \sum_{j=i-N}^{i-1} w_{i-j} \int_{\Omega} (R_j(x) \circ \varphi_j - R_i(x) \circ \varphi)^2 \, dx \stackrel{!}{=} min \quad (1)$$

within a local neighborhood N of all M slices (and respective transforms φ_j of images R_j for positive j) for more robust stack registration. The whole stack volume is finally resampled on an isotropic grid.
A uniform alignment of corresponding structures in the image stack (see fig. 1) in turn allows to relate a single section to a reference section using free-form deformations, which is exploited for the segmentation into prevailing tissues. In the segmentation step relevant biological structures within an image $I : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}^+$ are recognized and assigned a unique label $S : I \subset \Omega \mapsto \{1,\ldots,M\}$ for M tissues or classes. The classification of image grid points into multiple classes is a crucial step in the modelling pipeline. Here the raw intensity data is abstracted towards the rationale of the modelling process itself, where labeled voxel-data is the basis for quantification and surface-based modelling of internal structures. An automatic segmentation of sections is characterized by several requirements:

- A multitude of tissues must be recognized
- Images lack clearly defined edges and structures

[1]Here the computational cheap *Sum Of Squared Differences* (*SSD*) is used, because of histogram equalization in the preprocessing steps.

- The identification of tissue types needs expert knowledge

This necessitates the use of algorithms incorporating *a priori* information for robust multiclass-segmentation, where solely intensity-based techniques are clearly unfeasible.
In the context of section imaging in (Bollenbeck and Seiffert, 2008) we have suggested the segmentation into multiple classes by intensity driven registration and deformation of reference segmentations, performing equally accurate with image-feature based supervised classifiers like *support vector machines* and *multi layer perceptrons* in experiments on histological plant data, while being less computationally costly.
Employing the well known free-form deformable registration formulation

$$\mathbf{J}(u) := D(R,T;u) + \alpha s(u) \stackrel{!}{=} \min. \qquad (2)$$

of images $R,T : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{N}^+$ an *a priori* reference segmentation $S : R \subset \Omega \mapsto \{1,\ldots,M\}$ of R is adapted to segment T driven by an intensity based image-metric D subject to a regularized deformation u.
By using this method we classify voxels contained in the image stack to the respective tissue or material based on a small set of expert created reference segmentations.
These tissue mappings can be individually visualized by iso-surface renderings as in (Gubatz et al., 2007) and (Dercksen et al., 2008), where for valid inter-individual models *multiple* individual stacks are the basis for a statistical three-dimensional description.

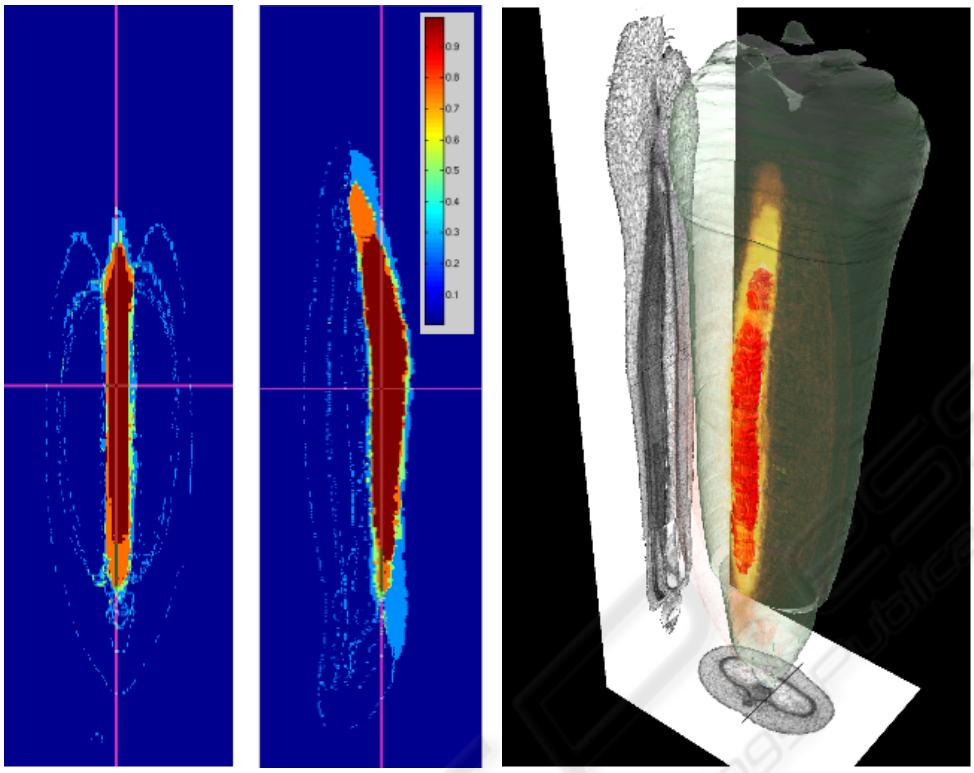
Statistical 3-D Models of Barley Grains. Whereas in the context of histological models, works so far have neglected biodiversity amongst specimen, the novelty of our approach is to provide a meaningful description of diversity amongst multiple specimen in one single model, allowing to quantify common themes and structures for further analysis and to provide a meaningful framework for the integration of functional data acquired from other individuals.
The quantification of inter-individual variability requires the mapping of data into a common reference frame allowing the estimation of ubiquitous tissues and regions of varying tissue composition.
While data sets generally vary in their physical extension, the task is to capture the composition of internal structures in terms of estimating a probability to each spatial coordinate for prevailing tissues based on segmented volume data, rather than constructing averaging surfaces or deformation models.
To obtain a transformation invariant to the actual tissue mapping, data sets are registered into a common coordinate frame by standardizing first image moments of the mass-centered intensities of the respective grayscale volumes of sectioned images (Alpert et al., 1990; Schmitt et al., 2006).
Instead of using registration approaches maximizing the correspondence of individual label- or grayscaled volumes directly with affine mappings, spline, or free-form deformations, a registration based only on individual image statistics can be considered *un-biased* in terms of leaving the inter-individual variances unaffected.
For each gridpoint $\vec{x} \in \Omega \subset \mathbb{R}^3$ and tissue M we estimate a probability for $\vec{x}$ belonging to tissue M empirically by

$$p_{\vec{x},M} := \frac{1}{|S|} \sum_{i=1}^{|S|} \delta(S_i(\vec{x}), M) \qquad (3)$$

from segmented data sets S_i.
Thereby a closed description of the spatial distribution of tissues and materials amongst specimen terms of a mapping $P : \Omega \mapsto \mathbb{R}^M$ is obtained.
While the spatial distribution of tissue probabilities is not based on assumptions on underlying distribution as with statistical deformation models, it can directly be related to underlying histological information (intensity volumes) as indicated in fig. 2.

3 RESULTS

For this work four individual grains at the same developmental stage were sectioned and digitized as described, yielding 2,128 to 2,736 slice images, each of size 1200×1600 pixels (approx. 30 GB image data).
Providing the basis for further modelling, intact individual grains were reconstructed from section images as proposed. Fig. 1 shows a volume rendering of a registered individual grain. By virtual lateral sections, revealing reconstituted histological features, such as cell walls etc., the performance of the registration approach in processing large image stacks could be initially validated, while detailed assessment is feasible, yet beyond the scope of this paper. Employing the described registration-segmentation algorithm, intensity stacks were segmented into their respective tissues based on a set of reference data defined by an expert.
For inter-individual description we registered and joined the segmented data in a common reference, yielding a volume of probabilities for each tissue (see fig. 2(a)). Using the probabilistic modelling, we addressed the biological questions (1.) how are specific tissues and relevant materials varying and (2.) what are ubiquitous themes amongst individuals. Thus, for an insightful 3-D visualization of probabilistic models



(a) (b)

Figure 2: Visualization of the inter-individual statistical atlas: 2(a) Two orthogonal length-sections through a volume of position-specific probabilities for the *nucellar projection*. 2(b) 3-D Rendering of a statistical model for the *nucellar projection*: Red volume represents tissue-material ubiquitous to all individuals., the opacity of the yellow volume rendering is proportional to the probability for the tissue amongst all individuals (projectional view and outer hull from a single individual).

or atlases, we are using a combination of two methods:

1. Volume rendering for the spatial distribution of tissue probability values
2. Surface rendering for ubiquitous regions, i.e. $p_{\vec{x},M} = 1$

Fig. 2 shows a combined rendering for the *nucellar projection*, which plays an important role in early grain development. A projectional view of a virtual lateral and section slice and transparent outer hull of an individual grain is displayed for better intuition. Visual analysis revealed that connected regions exist even for volumetrically small tissues such as the *vascular bundle* and *transfer cells*, with small variability amongst specimen, suggesting a determinant role in grain development.

4 DISCUSSION AND OUTLOOK

In this paper we present an interleaved registration and segmentation approach for automated 3-D model generation by integrating data from multiple individuals. By this inter-individual processing, high-resolution statistical models of internal structures are constructed towards high quality phenotyping of microscopic plant organs.

This significant extension of existing works on 3-D histological models provides the basis for systematic quantification of phenotypical properties, which is considered an urging topic in current plant biology. The presented modelling and quantification of inter-individual diversity further allows a reliable integration of functional data into a spatial atlas, whereas the description of diversity itself might also lead to new indications of functional interrelationships. Insightful visualization helps to identify common themes in morphology of developing seeds, especially structures related to storage compound aggregation.

Currently efforts for statistical models resolved on a

timeline are underway: For such digital *morphogenesis* population–averaging models are a preliminary.

ACKNOWLEDGEMENTS

This work is supported by BMBF grant 0313821. The authors want to thank U. Siebert and B. Zeike (IPK Gatersleben) for the material preparation, R. Pielot, and W. Weschke (IPK Gatersleben) for fruitful discussions and useful comments.

REFERENCES

Alpert, N. M., Bradshaw, J. F., Kennedy, D., and Correia, J. A. (1990). The principal axes transformation - a method for image registration. *The Journal of Nuclear Medicine*, 31:1717–1722.

Bollenbeck, F. and Seiffert, U. (2008). Fast registration-based automatic segmentation of serial section images for high-resolution 3-d plant seed modeling. *Biomedical Imaging: IEEE 5th Int. Symp.*, pages 352–355.

Caselles, V., Catte, F., Coll, T., and Dibos, F. (1997). A geometric model for active contours in image processing. *Int. J. Comput. Vision*, 22:158–175.

Dercksen, V. J., Brüß, C., Stalling, D., Gubatz, S., Seiffert, U., and Hege, H.-C. (2008). Towards automatic generation of 3D models of biological objects based on serial sections. In Linsen, L., Hagen, H., and Hamann, B., editors, *Visualization in Medicine and Life Sciences*, pages 3–25. Springer.

Gubatz, S., Dercksen, V., Brüß, C., Weschke, W., and Wobus, U. (2007). Analysis of barley (Hordeum vulgare) grain development using three-dimensional digital models. *The Plant Journal(12)*, 52(4):779–790.

Li, C., Xu, C., Gui, C., and Fox, M. (2005). Levelset set evolution without re-initialization: A new variational formulation. In *Proc. of the CVPR*.

Schmitt, O., Modersitzki, J., Heldmann, S., Wirtz, S., and Fischer, B. (2006). Image Registration of Sectioned Brains. *Int. J. Comput. Vision*, 73(1):5–39.