

Joint Registration and Segmentation of Histological Volume Data by Diffusion-Based Label Adaption

Felix Bollenbeck and Udo Seiffert

Fraunhofer Institute for Factory Operation and Automation IFF, D-39004 Magdeburg, Germany
felix.bollenbeck@iff.fraunhofer.de

Abstract

Three-dimensional serial section imaging delivers high spatial resolution and histological detail, which facilitates analysis of differentiation and development by exact labelling of tissues and cells, unknown to other 3-D imaging modalities. We propose an algorithm for interleaved reconstruction and segmentation of tissues in serial section volumes by diffusion-based registration and adaption of two-dimensional reference labellings. Iterative refinement of the global image congruence and local deformation of labellings delivers an efficient algorithm for processing of large volume data-sets. The benefits of the approach are shown by means of reconstruction and segmentation of giga-voxel serial section volumes of plant specimen.

1 Introduction

The investigation of growth and development of biological specimen in their natural spatial extension delivers a high benefit, and has become widely available to researches within the past decades. Apart from direct 3-D imaging by tomographic or laser scanning methods, serial section imaging remains an important method to exhibit histological details in 3-D. Particularly for investigating growth and differentiation processes, precise identification of prevailing cell types and tissues is important [3].

This paper proposes a method for joint reconstruction and segmentation of tissue types in un-aligned stacks of light microscopic serial section images, in typical data sets comprising $10^9 - 10^{10}$ voxels at approximately isotropic resolution. Since the three-dimensional coherence of structures is lost by preparation for digitization, the required registration and segmentation of volumes are inter-dependent tasks: In order to reconstruct specimen in the dimension of the sectioning, slice images need to be segmented into unique structures, which are again to be registered and brought into correspondence, reconstituting the intact three-dimensional organs. Concise identification and delimitation of multiple tis-

sues in sections is difficult, even for histological experts. We therefore believe, that the input of expert-created reference data is essential for correct segmentation of tissues in such histological volumes.

As a result of the dense sectioning, **i)** tissue segmentations in aligned adjacent slices differ in local deformations, and **ii)** identified local structures deliver additional constraints for a global registration of the image stack. To exploit these observations, we propose a model for interleaved registration and segmentation of images by iterative refinement of global transformations and deformation of references to capture local features. With the proposed model, we aim at mutually improving both tasks for a robust scheme for 3-D tissue reconstruction from section data. We perform experiments showing the usefulness of the approach in reconstruction and segmentation of giga-voxel sized serial section data sets of cereal grain specimen.

2 Related Work

In the context of processing volume data from serial sectioning a comprehensive study of registration schemes in reconstructing rat and human brains is given in [6], without addressing the segmentation of tissues in reconstructed volumes. In recent works [2] segmentation of such data using neural networks and a high-dimensional set of non-linear and computationally expensive features is proposed. In [1] random forest classifiers with a rich feature set are used for segmentation of blockface section volumes, using topological grouping post-processing to reduce over-segmentation. Atlas-based approaches for segmentation as in [5], using complete 3-D references have been widely used for medical data like MRI volumes. Such approaches depend on complete anatomical knowledge, being consistent in structure with the sampled image intensities to some degree, which is unavailable for our kind of data.

3 Method

Our approach aims at addressing registration and segmentation of image stacks inter-dependently: Re-

construction requires alignment of structures in sections, segmentation by image driven local deformation of reference labellings depends on a sufficient global alignment. We propose to join both tasks: Iterative affine registration, using initial and deficient tissue labellings, and segmentation based on image intensities in transformed sections by diffusion-based deformation of reference segmentations.

Denote a reconstructed and segmented volume by

$$\begin{aligned} \mathcal{V}(\vec{x}) &= (\mathcal{I}(\vec{x}), \mathcal{S}(\vec{x})) \\ \mathcal{I}(\vec{x}) &: \mathbb{R}^3 \mapsto \mathbb{R} \\ \mathcal{S}(\vec{x}) &: \mathbb{R}^3 \mapsto \{1, \dots, M\} \end{aligned} \quad (1)$$

with intensities \mathcal{I} and class labels \mathcal{S} . A measure for the error of a computed volume \mathcal{V}^* is

$$\mathcal{E}(\mathcal{V}^*) = E_1(\mathcal{I}^*) + E_2(\mathcal{S}^*). \quad (2)$$

E_1 and E_2 usually represent an intensity-based image-to-image metric, and a voxel- or surface-based labeling error measure, for assessing registration \mathcal{I}^* and segmentation \mathcal{S}^* , respectively. In turn, assuming a correct segmentation, E_1 can be evaluated using \mathcal{S}^* , and E_2 is based on how local deformation in image contents \mathcal{I}^* relates to aligned structures.

Rigid Registration: Registration of section stacks is performed by finding the transform parameters, maximizing image similarity of pairs of consecutive images. Improved optimization can be achieved by iterating over different scales of a scale-space representation of an image $I(x; t) = G_t * I(x)$, as described in [7]. Although the overall alignment result can significantly degrade, since structures on smaller scales are ignored. By registering images based on (possibly incomplete) segmentation of structures, these ambiguities can be overcome, yet for an ultimate alignment, an exact segmentation is needed.

For a stack of k images, an optimal transform is assumed as a set of pairwise alignments of adjacent segmentations $S_1, S_2 : \Omega \subset \mathbb{R}^2 \mapsto \{1, \dots, M\}$

$$\begin{aligned} E_1(\mathcal{S}^*) &:= \sum_k \int_{\Omega} D(S_1(\mathbf{x}), S_2(\mathbf{x})) d\mathbf{x} \\ &= \sum_{i=1}^k D(S_i, S_{i+1}). \end{aligned} \quad (3)$$

with the metric D being the cardinality of the set of pixels that match exactly between the moving and fixed images D_C . For the transformation $\varphi : \mathbb{R}^2 \mapsto \mathbb{R}^2$ we allow translation and rotation and the minimization problem $D_C(S_1, S_2; \varphi) := D_C(S_1, S_2 \circ \varphi) = \min$ is solved by downhill optimization.

Diffusion-Based Deformation: For section images I_1, I_2 with a reference segmentation $S_1 : \Omega \mapsto \{1, \dots, M\}$ of I_1 , a transformation of S_1 to segment I_2

can be found by deformable registration of I_2 to match I_1 , i.e. the segmentation is driven by similarity of underlying *images*, instead of similarity to an *atlas*. Given reference and target intensity image I_1, I_2 the goal is to find a transformation $\phi(x) = x + u(x)$ by point displacements u to maximize the similarity between both images.

The problem of finding an optimal deformation u minimizing the sum of squared intensity differences D_{SSD} must be regularized w.r.t. u :

$$D(I_1, I_2; u) := D_{SSD}(I_1, I_2 \circ u) + \alpha s(u) = \min. \quad (4)$$

and

$$E_2(\mathcal{T}^*) := \sum_{i=1}^k D(I_i, I_{i+1}; u) \quad (5)$$

We use a regularizer s based on the diffusion equation regularizing the *gradient* of the displacement

$$s(u) := \frac{1}{2} \sum_{l=1}^d \int_{\Omega} \|\nabla_{u_l}\|^2 dx. \quad (6)$$

While constraining a smooth displacement field, the diffusion-registration problem can be solved in $\mathcal{O}(N)$ operations per iteration using *additive operator splitting* (AOS) solution schemes [8] to compute displacement fields and $S_2^* = S_1 \circ u$ to obtain a segmentation of I_2 .

Iterative Refinement: For mutually improving alignment and segmentation, we iterate both steps in an EM-like heuristic over scale representations of images: Registration is improved with more exact labelling of structures, and vice versa.

For a set of Gaussians g_t of width t , I_t is the image representation $I_t := I * g_t$ at scale t , i.e. all structures smaller \sqrt{t} are attenuated. A pair of unaligned scale-space images $I_{1;t}, I_{2;t}$, the reference segmentation S_1 is fixed, while $S_{2;t}^*$ is updated iteratively over decreasing values of t . An initial segmentation of $S_{2;t,max}$ is calculated by estimating the conditional intensity distributions from S_1 on the coarsest scale $I_{1,t,max}$, which is determined by the spatial image resolution.

Iterating over increasingly finer scales t , registration and segmentation are updated according to

$$\begin{aligned} \varphi_t &= \operatorname{argmin}_{\varphi} D(S_{1;t}, S_{2;t}^*; \varphi) \\ u_t &= \operatorname{argmin}_{\varphi} D(I_{1;t}, I_{2;t} \circ \varphi; u) \\ S_{2;t}^* &:= S_{1,t} \circ u_t \quad t = T, \dots, 1. \end{aligned} \quad (7)$$

The registration-step aligns both images based on currently available segmentation. Given the new alignment of the intensity images, the segmentation-step then updates S^* based on the transformed intensity images (see fig. 2) as summarized in algorithm 1 (**DIF**).

4 Evaluation

We have evaluated our method on histological serial section volumes of developing barley grains,

Algorithm 1 DIF()

- 1: compute scale-space representations $I_{1,2} * G_t, t = 1, \dots, t_{\max}$
- 2: **for** $t = t_{\max}$ **downto** 1 **do**
- 3: compute $\phi = \operatorname{argmin}_{\phi} D_C(S_{1;t}, S_{2;t}; \phi)$ (*Downhill*)
- 4: apply transform $I_{2;t} = I_{2;t} \circ \phi$
- 5: compute $u_t = \operatorname{argmin}_{\phi} D(I_{1;t}, I_{2;t} \circ \phi; u)$ (*AOS Scheme*)
- 6: apply deformation $S_{2;t-1}^* = S_{1,t} \circ u_t$
- 7: **end for**

which were acquired by biologists in the context of works towards three-dimensional developmental atlases. Depending on the developmental stage, each individual data volume comprises between $2 \cdot 10^9$ and $1 \cdot 10^{10}$ voxels.

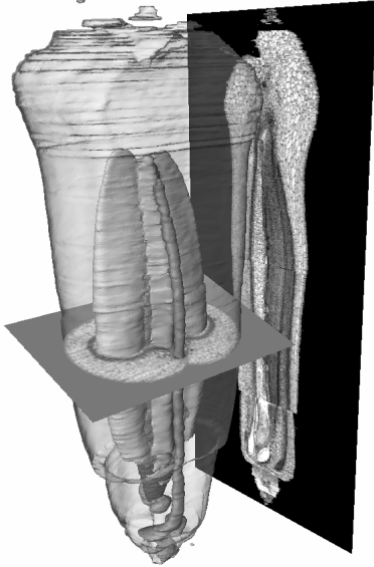


Figure 1: 3-D reconstruction and tissue segmentation from section data of developing barley grains: Surface rendering of segmented tissues and virtual histological sections ($1,600 \times 1,200 \times 1,954$ px).

Data sets: Image data was acquired from *Hordeum vulgare* specimen, bred under standardized conditions at three timepoints, yielding 1,058, 1,809 and 1,954 images respectively, at slice thickness of $3.0 \mu\text{m}$. Bright-field micrographs were digitized at a spatial resolution of $1.83 \times 1.83 \mu\text{m}$ per pixel to 1200×1600 px 12-bit grayscale images. Image backgrounds were removed using the method described in [4]. By experiments we aimed at assessing **i**) the segmentation performance of the proposed algorithm in terms of classification accuracy and **ii**) the ability of the approach for processing whole 3-D datasets for tissue segmentation and reconstruction.

Addressing **i**), a stack of 80 section images (i.e. 4.1% of the entire stack) were completely segmented by an

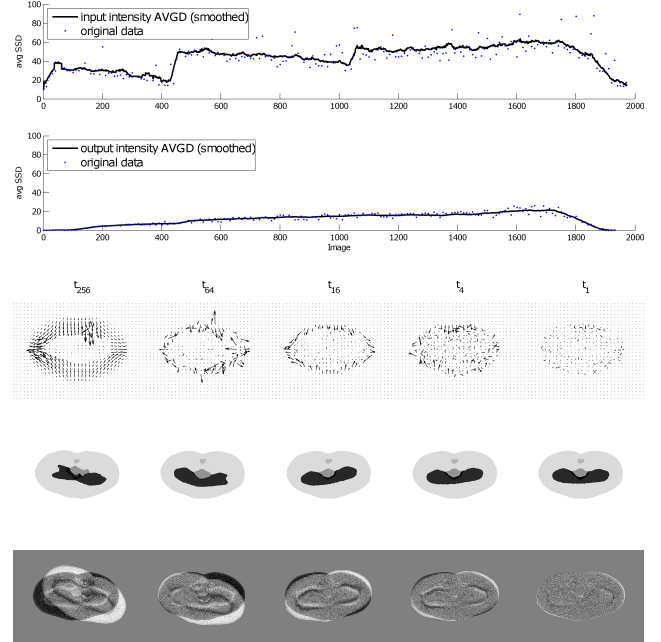


Figure 2: (Top down:) Registration results regarding similarity of adjacent stack images by average SSD of consecutive slices of the 7 DAF data set ($1,600 \times 1,200 \times 1,954$ px) *before* 1) and *after* 2) reconstruction. Iterative alignment and segmentation: 3) Displacement vector fields, 4) warped segmentations, and 5) corresponding rigid alignment over decreasing scales t_i .

expert biologist into six tissue types as ground truth data. We compared the segmentation accuracy of our method against feature-based classification, using a *one-vs-rest* RBF kernel SVM ensemble, and a 30 – 15 *logsig multi-layer perceptron* (MLP) for discrimination of 50-dimensional vectors of local image features (e.g. *Gabor-filters*, *gray-level co-occurrences*, and derived measures) into one of six classes.

Addressing **ii**) we applied the method to full datasets, where no manual ground truth was obtained due to data sizes.

Performance Measures: To measure the segmentation accuracy against ground truth class memberships, we used a confusion matrix and class-wise *accuracy* as an indicator regarding mis-classified pixels.

As an additional measure we used *distance to closest point* (*DCP*) as a weighting factor for accounting the spatial distribution of segmentation errors. For $S(x) = m$ and $S^*(x) = n$ being *true* and *predicted* class of $x \in \Omega$, the minimal distance

$$DCP(x) := \begin{cases} \min_{\hat{x}, S(\hat{x})=m} \|x - \hat{x}\|_{L_2} & \text{if } S(x) \neq S^*(x) \\ 0 & \text{if } S(x) = S^*(x) \end{cases} \quad (8)$$

to the *true* segment is aggregated for a weighted loss function.

Table 1: Class specific image segmentation accuracy for supervised classification (SVM, MLP) and proposed method (DIF) (6 classes with their fraction in the entire dataset)

Class	1 (42.18%)	2 (40.64%)	3 (1.21%)	4 (4.37%)	5 (1.22%)	6 (10.42%)	Total
DIF	0.99 ± 0.007	0.99 ± 0.004	0.83 ± 0.031	0.96 ± 0.017	0.73 ± 0.024	0.91 ± 0.095	0.97 ± 0.03
SVM	0.89 ± 0.035	0.89 ± 0.018	0.16 ± 0.343	0.85 ± 0.133	0.42 ± 0.19	0.83 ± 0.025	0.87 ± 0.116
MLP	0.87 ± 0.008	0.81 ± 0.027	0.17 ± 0.308	0.79 ± 0.311	0.15 ± 0.289	0.78 ± 0.262	0.82 ± 0.201

Table 2: Per-class average *distance-to-closest-point* (in px).

Class	1 (42.18%)	2 (40.64%)	3 (1.21%)	4 (4.37%)	5 (1.22%)	6 (10.42%)
DIF	1.46 ± 0.487	2.70 ± 0.189	8.17 ± 0.478	2.18 ± 0.858	2.23 ± 0.435	3.40 ± 1.062
SVM	18.63 ± 3.11	16.92 ± 1.12	48.17 ± 4.78	42.18 ± 2.58	42.23 ± 10.43	34.00 ± 10.62
MLP	18.17 ± 4.61	20.98 ± 2.01	65.15 ± 9.57	35.15 ± 12.28	58.63 ± 17.21	34.10 ± 11.0

4.1 Results

i) Segmentation accuracy on ground truth data:

For the ground-truth data features were extracted from available images. The data set was split into 10%/90% uniformly sampled test and training set ($\approx 1.5 \cdot 10^7 / \approx 1.5 \cdot 10^8$ samples) for SVM and MLP segmentation. 10% of equidistant segmented images were used for our method (DIF). The results are summarized in tables 1-2: Pixel classification by SVM and MLP achieve total accuracies of 87% and 82% respectively. Our method performs significantly better with 97% correct classification rate. Particularly for small classes the achieved accuracy is higher. While this problem might possibly be overcome by tuning training and -data, average DCP values show that mis-classified points are widely spread over the image domain. Here the constraints in method ensures that segmentation errors are locally bound.

ii) Reconstruction and segmentation of complete section volumes: To test the usefulness of our algorithm in reconstructing and segmenting large serial section data sets, we processed complete data sets of sectioned grains. For our results 5% of sections were segmented as references by an expert. Concerning stack reconstruction, metric values of consecutive stack slices are a fundamental indicator (fig. 2a), yet the quality of the alignment is ultimately determined by inspecting the reconstructed structures in 3-D. By virtual slices and renderings of segmented structures (fig. 1), biologists confirmed correct alignment of structures and tissue labeling within the accuracy of human perception in 3-D navigation.

5 Conclusion

We propose an algorithm for joint reconstruction and segmentation of serial section data, based on iterative rigid registration and diffusion-based deformation of reference segmentations. While we see the input of expert knowledge as being essential to any system recognizing multiple tissues in such data, interleaving registration and segmentation allows the use of expert knowledge for both reconstruction *and* labelling. The approach is limited when changes in image contents between slices are strong, however data acquisition protocols usually aim at isotropic resolutions where this should not be the case. Experiments show that the approach delivers high accuracies in segmentation, equal to feature-based supervised schemes, while avoiding computationally costly feature extraction. This makes the method particularly useful for processing very large high-resolution image volumes from serial histological sections. Recently, we could successfully use the proposed algorithm in different applications involving autoradiograph and confocal image data.

References

- [1] B. Andres, U. Kothe, M. Helmstaedter, W. Denk, and F. Hamprecht. Segmentation of SBFSEM volume data of neural tissue by hierarchical classification. In *Proc. 30th DAGM symp. on Pattern Recognition*, pages 142–152. Springer, 2008.
- [2] V. J. Dercksen, C. Brüß, D. Stalling, S. Gubatz, U. Seifert, and H.-C. Hege. Towards automatic generation of 3D models of biological objects based on serial sections. In L. Linsen, H. Hagen, and B. Hamann, editors, *Visualization in Medicine and Life Sciences*, pages 3–25. Springer-Verlag Berlin Heidelberg, 2008.
- [3] S. Gubatz, V. J. Dercksen, C. Brüß, W. Weschke, and U. Wobus. Analysis of barley (*Hordeum vulgare*) grain development using three-dimensional digital models. *The Plant Journal*, 6:779–790, 2007.
- [4] C. Gui and M. Fox. Level set evolution without re-initialization: a new variational formulation. In *IEEE Conf. on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, volume 1, pages 430–436, 2005.
- [5] T. Rohlfing, D. Russakoff, and C. Maurer. Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE Trans. on Med. Imaging*, 23(8):983–994, 2004.
- [6] O. Schmitt, J. Modersitzki, S. Heldmann, S. Wirtz, and B. Fischer. Image registration of sectioned brains. *Int. J. Comput. Vision*, 73(1):5–39, 2006.
- [7] P. Thevenaz, U. Ruttimann, and M. Unser. Iterative multi-scale registration without landmarks. In *Image Processing, 1995. Proc., Internl. Conf. on*, volume 3, 1995.
- [8] J. Weickert, B. Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998.