

Assignment Flow For Order-Constrained OCT Segmentation

Dmitrij Sitenko, Bastian Boll, and Christoph Schnörr

Image and Pattern Analysis Group (IPA) and Heidelberg Collaboratory for Image Processing (HCI), Heidelberg University, Germany

Abstract. At the present time Optical Coherence Tomography (OCT) is among the most commonly used non-invasive imaging methods for the acquisition of large volumetric scans of human retinal tissues and vasculature. Due to tissue-dependent speckle noise, the elaboration of automated segmentation models has become an important task in the field of medical image processing.

We propose a novel, purely data driven *geometric approach to order-constrained 3D OCT retinal cell layer segmentation* which takes as input data in any metric space. This makes it unbiased and therefore amenable for the detection of local anatomical changes of retinal tissue structure. To demonstrate robustness of the proposed approach we compare four different choices of features on a data set of manually annotated 3D OCT volumes of healthy human retina. The quality of computed segmentations is compared to the state of the art in terms of mean absolute error and Dice similarity coefficient.

1 Introduction

Overview. Optical Coherence Tomography (OCT) is a non-invasive imaging technique which measures the intensity response of back scattered light from millimeter penetration depth. We focus specifically on the application of OCT in ophthalmology for aquisition of high-resolution volume scans of the human retina. This provides information about retinal tissue structure in vivo to understand human eye functionalities. OCT devices record multiple two-dimensional B-scans in rapid succession and combine them into a single volume in a subsequent alignment step. Taking an OCT scan only takes minutes and can help detect symptoms of pathological conditions such as glaucoma, diabetes, multiple sclerosis or age-related macular degeneration.

The relative ease of data aquisition also enables to use multiple OCT volume scans of a single patient over time to track the progression of a pathology or quantify the success of therapeutic treatment. To better leverage the availability of raw OCT data in both clinical settings and empirical studies, much work has focused on automatic extraction of relevant information, in particular automatic cell layer segmentation, detection of fluid and reconstruction of vascular structures. The difficulty of these tasks lies in challenging signal-to-noise

ratio which is influenced by multiple factors including mechanical eye movement during registration and the presence of speckle.

Related Work. Effective segmentation of OCT volumes is a very active area of research. Several methods for segmenting human retina were proposed in [16], [12],[11] and [2] which rely on graphical models. To increase robustness, the retina segmentation approaches proposed in [23] and [9] employ shape priors using soft constraints. In [19] Rathke first introduced a parallelizable segmentation method based on probabilistic graphical models with global low-rank shape prior. Variational approaches given in [8,26] and [18] model retina layers by zero level sets with properly chosen functionals including soft constraints. Much recent work has focused on the use of deep learning to address the task of cell layer segmentation in a purely data driven way. Methods presented in [17,21] rely on the U-net architecture [20] which yields good predictive performance in settings with limited availability of training data. To enforce global order of cell layers along a spatial axis as well as additional regularization, local predictions have been tied together through graph-based methods [10] or through a second machine learning component [14]. However, if global context is already used in feature extraction, the risk of overfitting remains and unseen pathologies may result in unpredictable behavior.

Approach. Our segmentation approach is a smooth image labeling algorithm based on geometric numerical integration on an elementary statistical manifold. It can work with input data from any metric space, making it agnostic to the choice of feature extraction and suitable as plug-in replacement in diverse pipelines. In addition to respecting the natural order of cell layers, our segmentation process has a high amount of built-in parallelism such that modern graphics acceleration hardware can easily be leveraged. We evaluate the effectiveness of our novel approach for a selection of input features ranging from traditional covariance descriptors to convolutional neural networks.

Contribution. We propose a geometric assignment approach that extends the approach introduced by [4] to retinal layer segmentation with the following novel characteristics:

- (i) By leveraging a continuous characterization of layer ordering, our method is able to simultaneously perform local regularization and to incorporate the global topological ordering constraint in a *single smooth* labeling process. The segmentation is computed from a distance matrix containing pairwise distances between data for each voxel and prototypical data for each layer in some feature space. This highlights the ability to extract features from raw OCT data in a variety of different ways and to use the proposed segmentation as a plug-in replacement for other graph-based methods.
- (ii) Computationally fast and high-quality cell layer segmentations of OCT volumes are obtained by using only local features for each voxel. This is in contrast to competing deep learning approaches which commonly use information from an entire B-scan as input. In addition, the exclusive use of local features combats bias introduced through limited data availability

in training and enables the incorporation of three-dimensional information without compromising runtime scalability.

- (iii) The highly parallelizable segmentation approach with global voxel interaction enables robust cell layer segmentation of entire OCT volumes without using any prior knowledge other than local regularity and order of cell layers. In particular, no global shape prior is used as opposed to segmentation methods relying on graphical models like, e.g., [19]. Figure 1.1 shows a typical result obtained with our novel approach after segmenting healthy retina tissues with labels specified in (4.1).

Organization. Our paper is organized as follows. The assignment flow approach is briefly summarized in Section 2 and extended in Section 4 in order to take into account the order of layers as a global constraint. In Section 3, we consider the Riemannian manifold \mathcal{P}_d of positive definite matrices as a suitable feature space for local OCT data descriptors. The resulting features are subsequently compared to local features extracted by a convolutional network in Section 5. The evaluation of performance measures for OCT segmentation of our novel approach are reported in Section 5 and compared to the state-of-the-art method given in [16].

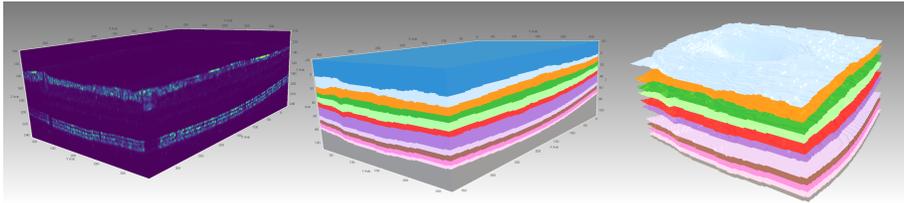


Fig. 1.1: From left to right: 3D OCT volume scan dimension $512 \times 512 \times 256$ of healthy human retina with ambiguous locations of layer boundaries. The resulting segmentation of 11 layers expressing the order preserving labeling of the proposed approach. Illustration of boundary surfaces between different segmented cell layers.

2 Assignment Flow

We summarize the assignment flow approach introduced by [4] and refer to the recent survey [22] for more background and a review of recent related work.

Assignment Manifold. Let $(\mathcal{F}, d_{\mathcal{F}})$ be a metric space and $\mathcal{F}_n = \{f_i \in \mathcal{F} : i \in \mathcal{I}\}$, $|\mathcal{I}| = n$ given data. Assume that a predefined set of prototypes $\mathcal{F}_* = \{f_j^* \in \mathcal{F} : j \in \mathcal{J}\}$, $|\mathcal{J}| = c$ is given. *Data labeling* denotes assignments $j \rightarrow i$, $f_j^* \rightarrow f_i$ to be determined in a spatially regularized fashion. The assignments at each pixel $i \in \mathcal{I}$ are encoded by *assignment vectors* $W_i = (W_{i1}, \dots, W_{ic})^{\top} \in \mathcal{S}$ in the relative interior $\mathcal{S} = \text{rint } \Delta_c$ of the probability simplex, that becomes a

Riemannian manifold (\mathcal{S}, g) endowed with the Fisher-Rao metric g from information geometry. The *assignment manifold* (\mathcal{W}, g) , $\mathcal{W} = \mathcal{S} \times \cdots \times \mathcal{S}$ ($n = |\mathcal{I}|$ factors) is the product manifold whose points encode the label assignments at all pixels.

Assignment Flow. Based on the given data and prototypes, the distance vector field $D_{\mathcal{F};i} = (d_{\mathcal{F}}(f_i, f_1^*), \dots, d_{\mathcal{F}}(f_i, f_c^*))^\top$, $i \in \mathcal{I}$ is well defined. This data representation is lifted to the assignment manifold by the *likelihood map* and the *likelihood vectors*, respectively,

$$L_i: \mathcal{S} \rightarrow \mathcal{S}, \quad L_i(W_i) = \frac{W_i e^{-\frac{1}{\rho} D_{\mathcal{F};i}}}{\langle W_i, e^{-\frac{1}{\rho} D_{\mathcal{F};i}} \rangle}, \quad i \in \mathcal{I}. \quad (2.1)$$

This map is based on the affine e-connection of information geometry and the scaling parameter $\rho > 0$ is used for normalizing the a-prior unknown scale of the components of $D_{\mathcal{F};i}$ that depends on the specific application at hand. The likelihood vectors are spatially regularized by the *similarity map* and the *similarity vectors*, respectively,

$$S_i: \mathcal{W} \rightarrow \mathcal{S}, \quad S_i(W) = \text{Exp}_{W_i} \left(\sum_{k \in \mathcal{N}_i} w_{ik} \text{Exp}_{W_i}^{-1}(L_k(W_k)) \right), \quad i \in \mathcal{I}, \quad (2.2)$$

where $\text{Exp}_p(v) = \frac{pe^{v/\rho}}{\langle p, e^{v/\rho} \rangle}$ is the exponential map corresponding to the e-connection and positive weights w_{ik} , $k \in \mathcal{N}_i$, that sum up to 1 on every patch around pixel i indexed by \mathcal{N}_i , determine the regularization properties.

The *assignment flow* is induced on the assignment manifold \mathcal{W} by the locally coupled system of nonlinear ODEs

$$\dot{W}_i = R_{W_i} S_i(W), \quad W_i(0) = \mathbb{1}_{\mathcal{S}}, \quad i \in \mathcal{I}, \quad (2.3)$$

where the map $R_p = \text{Diag}(p) - pp^\top$, $p \in \mathcal{S}$ turns the right-hand side into a tangent vector field and $\mathbb{1}_{\mathcal{W}} \in \mathcal{W}$ denotes the barycenter of the assignment manifold \mathcal{W} . The solution $W(t) \in \mathcal{W}$ is numerically computed by geometric integration [27] and determines a labeling $W(T)$ for sufficiently large T after a trivial rounding operation. Convergence and stability of the assignment flow have been studied by [28].

3 OCT Data Representation by Covariance Descriptors

In this section, we briefly sketch the basic geometric notation for representation of OCT data in terms of covariance descriptors $f_i \in \mathcal{F}_n$ [25] and identify the metric data space $(\mathcal{F}, d_{\mathcal{F}})$ underlying (2.1).

The Manifold \mathcal{P}_d . The Riemannian manifold (\mathcal{P}_d, g) of positive definite matrices of dimension $\frac{(d+1)(d)}{2}$ and the Riemannian metric are given by

$$\mathcal{P}_d = \{S \in \mathbb{R}^{d \times d}: S = S^\top, S \text{ is positive definite}\} \quad (3.1a)$$

$$g_S(U, V) = \text{tr}(S^{-1}US^{-1}V), \quad U, V \in T_S\mathcal{P}_d = \{S \in \mathbb{R}^{d \times d}: S^\top = S\}. \quad (3.1b)$$

The Riemannian distance is given by

$$d_{\mathcal{P}_d}(S, T) = \left(\sum_{i \in [d]} (\log \lambda_i(S, T))^2 \right)^{1/2}, \quad (3.2)$$

whereas the globally defined exponential map reads

$$\exp_S(U) = S^{\frac{1}{2}} \expm(S^{-\frac{1}{2}} U S^{-\frac{1}{2}}) S^{\frac{1}{2}}, \quad (3.3)$$

with $\expm(\cdot)$ denoting the matrix exponential. Given a smooth objective function $J: \mathcal{P}_d \rightarrow \mathbb{R}$, the Riemannian gradient is given by

$$\text{grad } J(S) = S(\partial J(S))S \in T_S \mathcal{P}_d, \quad (3.4)$$

where the symmetric matrix $\partial J(S)$ is the Euclidean gradient of J at S .

Region Covariance Descriptors. To apply the introduced geometric framework, we model each OCT volume by a mapping $I: \mathcal{D} \rightarrow \mathbb{R}_+$ where $\mathcal{D} \subset \mathbb{R}^3$ is an underlying spatial domain.

To each voxel $v \in \mathcal{D}$, we associate the local feature vector $f: \mathcal{D} \rightarrow \mathbb{R}^{10}$,

$$f(v) := (I(v), \nabla_x I(v), \nabla_y I(v), \nabla_z I(v), \sqrt{2} \nabla_{xy} I(v), \dots, \nabla_{zz} I(v))^{\top} \quad (3.5)$$

assembled from the raw intensity value $I(v)$ as well as first- and second-order responses of derivatives filters capturing information from larger scales following [13]. By introducing a suitable geometric graph spanning \mathcal{D} , we can associate a neighborhood $\mathcal{N}(i)$ of fixed size with each voxel $i \in \mathcal{I}$ as in (2.2). For each neighborhood, we define the regularized *region covariance descriptor* S_i as

$$S_i := \sum_{j \in \mathcal{N}(i)} \theta_{ij} (f_j - \bar{f}_i)(f_j - \bar{f}_i)^{\top} + \epsilon I, \quad \bar{f}_i = \sum_{k \in \mathcal{N}(i)} \theta_{ik} f_k, \quad (3.6)$$

as a weighted empirical covariance matrix with respect to feature vectors f_j . The small value $1 \gg \epsilon > 0$ acts as a regularization parameter enforcing positive definiteness of S_i . In the following, we use the shorthand notation $[n] = \{1, \dots, n\}$ for natural numbers n .

Computing Prototypical Covariance Descriptors. Given a set of covariance descriptors

$$\mathcal{S}_N = \{(S_1, \omega_1), \dots, (S_N, \omega_N)\} \subset \mathcal{P}_d \quad (3.7)$$

together with positive weights ω_i , we next focus on the solution of the problem

$$\bar{S} = \arg \min_{S \in \mathcal{P}_d} J(S; \mathcal{S}_N), \quad J(S; \mathcal{S}_N) = \sum_{i \in [N]} \omega_i d_{\mathcal{P}_d}^2(S, S_i), \quad (3.8)$$

with the distance $d_{\mathcal{P}_d}$ given by (3.2). From (3.3), we deduce

$$U = \exp_S^{-1} \circ \exp_S(U) = S^{\frac{1}{2}} \logm(S^{-\frac{1}{2}} \exp_S(U) S^{-\frac{1}{2}}) S^{\frac{1}{2}} \quad (3.9)$$

with the matrix logarithm $\text{logm} = \text{expm}^{-1}$ [15, Section 11]. The efficient mean retrieval of (3.8) regarding the evaluation of (3.2) requires a nontrivial matrix decomposition that has to be applied multiple times to every voxel (vertex) of a 3D gridgraph. This results in an overall quite expensive approach in particular for a large data set. Therefore we reduce the computational costs by relying on an approximation of the Riemannian mean by employing surrogate metrics and distances introduced below.

Log-Euclidean Distance and Means. A computationally cheap approach was proposed by [3] (among several other ones). Based on the operations

$$S_1 \odot S_2 = \text{expm}(\text{logm}(S_1) + \text{logm}(S_2)), \quad (3.10a)$$

$$\lambda \cdot S = \text{expm}(\lambda \text{logm}(S)), \quad (3.10b)$$

the set $(\mathcal{P}_s, \odot, \cdot)$ becomes isomorphic to the vector space where \odot plays the role of addition. Consequently, the mean of the data \mathcal{S}_N given by (3.7) is defined analogous to the arithmetic mean by

$$\bar{S} = \text{expm}\left(\sum_{i \in [N]} \omega_i \text{logm}(S_i)\right). \quad (3.11)$$

While computing the mean is considerably cheaper than integrating the flow induced by (3.4) with respect to objective (3.8), the geometry (curved structure) of the manifold \mathcal{P}_d is ignored. Therefore, in the next section, we additionally consider another approximation of the Riemannian mean that better respects the underlying geometry but can still be evaluated efficiently.

S-Divergence and Means. For an approximation of the objective function (3.8), we replace the Riemannian $d_g^2(p, q)$ distance by the *Stein divergence* proposed by Sra [24]

$$D_s(S_1, S_2) = \log \det\left(\frac{S_1 + S_2}{2}\right) - \frac{1}{2} \log \det(S_1 S_2), \quad S, S_1, S_2 \in \mathcal{P}_d, \quad (3.12)$$

and avoid involved generalized eigenvalue problem for evaluation of (3.2) by replacing (3.8) with

$$\bar{S} = \arg \min_{S \in \mathcal{P}_d} J_s(S; \mathcal{S}_N), \quad J_s(S; \mathcal{S}_N) = \sum_{i \in [N]} \omega_i D_s(S, S_i). \quad (3.13)$$

We refer to, e.g., [6,5] for a more complete exposition of divergence functions. The Riemannian gradient flow for this specific problem reads

$$\dot{S} = -\text{grad } J_s(S; \mathcal{S}_N) \stackrel{(3.4)}{=} -S \partial J(S; \mathcal{S}_N) S \quad (3.14a)$$

$$= -\frac{1}{2}(SR(S; \mathcal{S}_N)S - S), \quad R(S; \mathcal{S}_N) = \sum_{i \in [N]} \omega_i \left(\frac{S + S_i}{2}\right)^{-1}. \quad (3.14b)$$

Discretizing the flow using the geometric explicit Euler scheme with step size h yields,

$$S_{(t+1)} = \exp_{S_{(t)}} \left(-h \operatorname{grad} J_s(S_{(t)}; \mathcal{S}_N) \right) \quad (3.15a)$$

$$\stackrel{(3.3)}{=} S_{(t)}^{\frac{1}{2}} \operatorname{expm} \left(\frac{h}{2} \left(I - S_{(t)}^{\frac{1}{2}} R(S_{(t)}; \mathcal{S}_N) S_{(t)}^{\frac{1}{2}} \right) \right) S_{(t)}^{\frac{1}{2}}. \quad (3.15b)$$

Using as initial point $S_{(0)}$ the log-Euclidean mean (3.11) defines the following algorithm that we use for mean retrieval throughout the present paper.

Algorithm 1: Geometric Matrix Mean Based on the S -divergence.

Initialization, ϵ (termination threshold)

$t = 0$, $S_{(0)}$ solves (3.11)

$\epsilon_0 > \epsilon$ (any value ϵ_0)

while $\epsilon_t > \epsilon$ **do**

$$\left[\begin{array}{l} LL^\top = S_{(t)} \\ L_i L_i^\top = \frac{S_{(t)} + S_i}{2} \text{ for } i \in [N] \\ U = I - S_{(t)}^{\frac{1}{2}} \left(\sum_{i \in [N]} \omega_i (L_i L_i^\top)^{-1} \right) S_{(t)}^{\frac{1}{2}} \\ S_{(t+1)} = S_{(t)}^{\frac{1}{2}} \operatorname{expm} \left(\frac{h}{2} U \right) S_{(t)}^{\frac{1}{2}} \\ \epsilon_{t+1} := \|U\|_F, \quad t \leftarrow t + 1 \end{array} \right.$$

4 Ordered Layer Segmentation

In this section, we work out an extension of the assignment flow (Section 2) which is able to respect the order of cell layers as a global constraint while remaining in the same smooth geometric setting. In particular, existing schemes for numerical integration still apply to the novel variant.

4.1 Ordering Constraint

With regard to segmenting OCT data volumes, the order of cell layers is crucial prior knowledge. Figure 4.1 illustrates for a schematic OCT volume acquisition of 11 retina layers and 3 separating membranes (ILM,ELM,BM) and typical scan notations used throughout the paper. To incorporate this knowledge into the geometric setting of Section 2, we require a smooth notion of ordering which allows to compare two probability distributions. In the following, we assume prototypes $f_j^* \in \mathcal{F}$, $j \in [n]$ in some feature space \mathcal{F} to be indexed such that ascending label indices reflect the physiological order of cell layers.

Definition 1 (Ordered Assignment Vectors). *A pair of voxel assignments $(w_i, w_j) \in \mathcal{S}^2$, $i < j$ within a single A-scan is called ordered, if $w_j - w_i \in K =$*

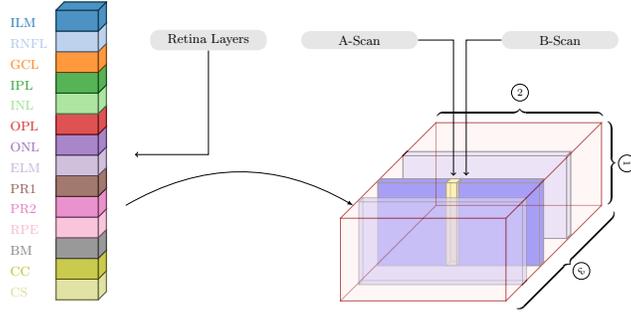


Fig. 4.1: OCT volume acquisition: ① is the A-scan axis (single A-scan is marked yellow). Multiple A-scans taken in rapid succession along axis ② form a two-dimensional B-scan (single B-scan is marked blue). The complete OCT volume is formed by repeating this procedure along axis ③. A list of retina layers and membranes we expect to find in every A-scan is shown on the left.

$\{By: y \in \mathbb{R}_+^c\}$ which is equivalent to $Q(w_j - w_i) \in \mathbb{R}_+$ with the matrices

$$B_{ij} := \begin{cases} -1 & \text{if } i = j \\ 1 & \text{if } i - j = 1 \\ 0 & \text{else} \end{cases}, \quad Q_{i,j} = \begin{cases} 1 & \text{if } i \geq j \\ 0 & \text{else} \end{cases} \quad (4.1)$$

4.2 Ordered Assignment Flow

Likelihoods as defined in (2.1) emerge by lifting $-\frac{1}{\rho}D_{\mathcal{F}}$ regarded as Euclidean gradient of $-\frac{1}{\rho}\langle D_{\mathcal{F}}, W \rangle$ to the assignment manifold. It is our goal to encode order preservation into a generalized likelihood matrix $L_{\text{ord}}(W)$. To this end, consider the assignment matrix $W \in \mathcal{S}^N$ for a single A-scan consisting of N voxels. We define the related matrix $Y(W) \in \mathbb{R}^{N(N-1) \times c}$ with rows indexed by pairs $(i, j) \in [N]^2$, $i \neq j$ in fixed but arbitrary order. Let the rows of Y be given by

$$Y_{(i,j)}(W) = \begin{cases} Q(w_j - w_i) & \text{if } i > j \\ Q(w_i - w_j) & \text{if } i < j \end{cases}. \quad (4.2)$$

By construction, an A-scan assignment W is ordered exactly if all entries of the corresponding $Y(W)$ are nonnegative. This enables to express the ordering constraint on a single A-scan in terms of the energy objective

$$E_{\text{ord}}(W) = \sum_{(i,j) \in [N]^2, i \neq j} \phi(Y_{(i,j)}(W)). \quad (4.3)$$

where $\phi: \mathbb{R}^c \rightarrow \mathbb{R}$ denotes a smooth approximation of $\delta_{\mathbb{R}_+^c}$. In our numerical experiments, we choose

$$\phi(y) = \left\langle \gamma \exp\left(-\frac{1}{\gamma}y\right), \mathbf{1} \right\rangle \quad (4.4)$$

with a constant $\gamma > 0$. Suppose a full OCT volume assignment matrix $W \in \mathcal{W}$ is given and denote the set of submatrices for each A-scan by $C(W)$. Then order preserving assignments consistent with given distance data $D_{\mathcal{F}}$ in the feature space \mathcal{F} are found by minimizing the energy objective

$$E(W) = \langle D_{\mathcal{F}}, W \rangle + \sum_{W_A \in C(W)} E_{\text{ord}}(W_A) . \quad (4.5)$$

We consequently define the generalized likelihood map

$$L_{\text{ord}}(W) = \exp_W (-\nabla E(W)) = \exp_W \left(-\frac{1}{\rho} D_{\mathcal{F}} - \sum_{W_A \in C(W)} \nabla E_{\text{ord}}(W_A) \right) \quad (4.6)$$

and specify a corresponding assignment flow variant.

Definition 2 (Ordered Assignment Flow). *The dynamical system*

$$\dot{W} = R_W S(L_{\text{ord}}(W)), \quad W(0) = \mathbb{1}_{\mathcal{W}} \quad (4.7)$$

evolving on \mathcal{W} is called the ordered assignment flow.

By applying known numerical schemes [27] for approximately integrating the flow (4.7), we find a class of discrete-time image labeling algorithms which respect the physiological cell layer ordering in OCT data. In chapter 5, we benchmark the simplest instance of this class, emerging from the choice of geometric Euler integration.

5 Experimental Results and Discussion

OCT-Data. In the following sections, we describe experiments performed on a set of volumes with annotated OCT B-scans extracted by a spectral domain OCT device (Heidelberg Engineering, Germany). Further, we always assume an OCT volume in question to consist of N_B B-scans, each comprising N_A A-scans with N voxels.

While raw OCT volume data has become relatively plentiful in clinical settings, large volume datasets with high-quality gold-standard segmentation are not widely available at the time of writing. By extracting features which represent a given OCT scan *locally* as opposed to incorporating global context at every stage, it is our hypothesis that superior generalization can be achieved in the face of limited data availability. This is most expected for pathological cases in which global shape of cell layers may deviate drastically from seen examples in the training data. Our approach consequently differs from common deep learning methods which explicitly aim to incorporate global context into the feature extraction process. Utilization of shape prior limits the methods ability to generalize to unseen data if large deviation from the expected global shape seen in training is present.

Prototypes on \mathcal{P}^d . For applying the framework introduced in Section (2), we interpret covariance features (3.6) as data points $f_i \in \mathcal{F}_n$ evolving on the natural metric space (3.1a) and model each retina tissue indexed by $l \in \{1, \dots, C\}$ with a random variable S_l taking values $\{S_l^k\}_{k=1}^{N_l}$. To generalize the retina layer detection to multiple OCT data sets instead of just using a single prototype (3.13), we partition the samples $\{S_l^k\}_{k=1}^{N_l}$ into K_l disjoint sets $\{S_{1_l}^l, \dots, S_{K_l}^l\}$ with representatives $\{\tilde{S}_l^1, \dots, \tilde{S}_l^{K_l}\}$. These are serving as prototypes f_j^* , $j \in \mathcal{J}$ which are determined offline for each $l \in \{1, \dots, 14\}$ as the minimal expected loss measured by the Stein divergence (3.12) according to K -means like functional

$$E_{p_l}(S^l) = \sum_{j=1}^{K_l} p(j) \sum_{S_i \in S_j} \frac{p(i|j)}{p(j)} D_S(S_l^i, \tilde{S}_l^j), \quad p(i, j) = \frac{1}{N_l}, \quad p(j) = \frac{N_j}{N_l}, \quad (5.1)$$

with marginals $p_l(j) = \sum_{i=1}^{N_j} p_l(j|S_l^i)$ and using Algorithm (1) for mean retrieval. The experimental results discussed next illustrate the relative influence of the

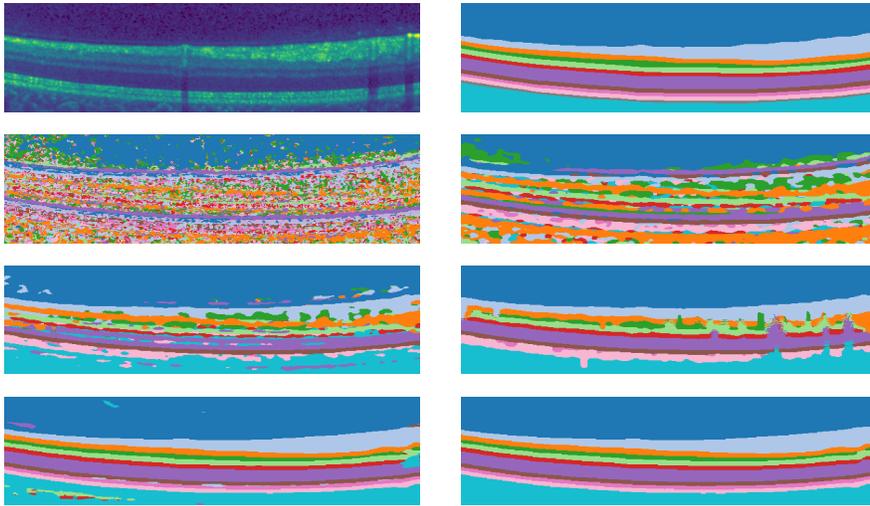


Fig. 5.1: **From top to bottom. 1st row:** One B-scan from OCT-volume showing the shadow effects with annotated ground truth on the right. **2nd row:** Nearest neighbor assignment based on prototypes computed with Stein divergence and result of the segmentation returned by the basic assignment flow (Section (2)) on the right. **3rd row:** Illustration of the proposed *layer-ordered* volume segmentation based on covariance descriptors with ordered volume segmentation for different $\gamma = 0.5$ on left and $\gamma = 0.1$ on the right (cf. Eq. (4.4)). **4th row:** Illustration of local rounding result extracted from Res-Net and the result of ordered flow on the right.

covariance descriptors and regularization property of the ordered assignment flow, respectively. Throughout, we fixed the grid connectivity \mathcal{N}_i for each voxel

$i \in \mathcal{I}$ to $3 \times 5 \times 5$. Figure (5.1), second row, illustrates a typical result of nearest neighbor assignment and the volume segmentation *without* ordering constraints. As the second row shows, the high texture similarity between the choroid and GCL layer yields wrong predictions resulting in violation of biological retina ordering through the whole volume which cannot be resolved with the based assignment flow approach given in Section (2). In third row of Figure (5.1), we plot the *ordered* volume segmentation by stepwise increasing the parameter γ defined in (4.4), which controls the ordering regularization by means of the novel generalized likelihood matrix (4.6). The direct comparison with the ground truth remarkably shows how the ordered labelings evolve on the assignment manifold while simultaneously giving accurate data-driven detection of RNFL, OPL, INL and the ONL layer. For the remaining critical inner layers, the local prototypes extracted by (5.1) fail to segment the retina properly, due to the presence of vertical shadow regions originating from the scanning process of the OCT-data.

CNN Features. In addition to the covariance features in Section 3, we compare a second approach to local feature extraction based on a convolutional neural network architecture. For each node $i \in [n]$, we trained the network to directly predict the correct class in $[c]$ using raw intensity values in \mathcal{N}_i as input. As output, we find a score for each layer which can directly be transformed into a distance vector suitable as input to the ordered assignment flow (4.7) via (4.6). The specific network used in our experiments has a ResNet architecture comprising four residually connected blocks of 3D convolutions and ReLU activation. Model size was hand-tuned for different sizes of input neighborhoods, adjusting the number of convolutions per block as well as corresponding channel dimensions. In particular, labeling accuracy is increased for detection of RPE and PR2 layers, as illustrated in the last row of figure (5.1).

Evaluation. To assess the segmentation performance of our proposed approach, we compared to the state of the art graph-based retina segmentation method of 10 intra-retinal layers developed by the Retinal Image Analysis Laboratory at the Iowa Institute for Biomedical Imaging [16,1,11], also referred to as the IOWA Reference Algorithm. We quantify the region agreement with manual segmentation regarded as gold standard. Specifically, we calculate the DICE similarity coefficient [7] and the mean absolute error for segmented cell layer within the pixel size of $3.87 \mu\text{m}$ compared to human grader on an OCT volume consisting of 61 B-scans reported in Table 1. To allow a direct comparison to the proposed segmentation method, the evaluation was performed on layers summarized in Table 1. We point out that in general our method is not limited to any number of segmented layers if ground truth is available and further performance evaluations though additional comparison with the method proposed in [19] will be included in the complete report of the proposed approach which is beyond scope of this paper. The OCT volumes were imported into OCTExplorer 3.8.0 and segmented using the predefined Macular-OCT IOWA software.

Both methods detect the RNFL layer with high accuracy whereas for the underlying retina tissues the automated segmentation with ordered assignment flow indicates the smallest mean absolute error and the highest Dice similarity

index underpinning the superior performance of order preserving labeling in view of accuracy.

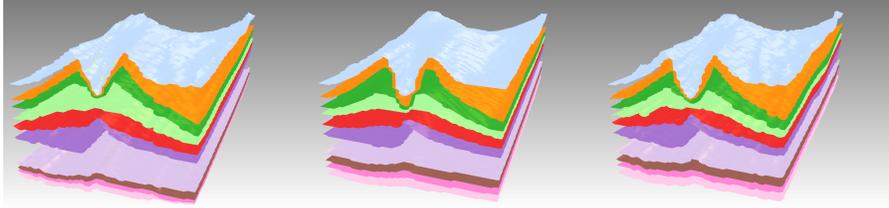


Fig. 5.2: Visualization of segmented intraretinal surfaces: **Left:** IOWA layer detection of 10 boundaries, **Middle:** Proposed labeling result based on local features extraction, **Right:** Ground truth. For a quantitative comparison: see Table 1.

	DICE index	Mean absolute error	IOWA	DICE index	Mean absolute error
OAF	0.9962	3.5920	RNFL	0.9906	2.8290
RNFL	0.8390	1.3091	GCL	0.7933	2.1063
GCL	0.8552	4.0340	IPL	0.7148	5.0753
IPL	0.8714	6.0180	INL	0.7696	6.0090
INL	0.8886	4.5345	OPL	0.8510	5.4852
OPL	0.9070	1.9550	ONL+ELM+PR1	0.8374	7.0928
ONL+ELM+PR1	0.9784	2.6511	PR2+RPE	0.9006	12.4891
PR2+RPE					

Table 1: Mean absolute error measures are given in pixels (1 pixel = 3.87 μm). **Left:** Dice and mean absolute error of the proposed approach. **Right:** Resulting metrics achieved by IOWA reference algorithm. All numbers demonstrate the superior performance of our novel order-preserving labeling approach.

6 Conclusion

In this paper we presented a novel, fully automated and purely data driven approach for retina segmentation in OCT-volumes. Compared to methods [16] [9] and [19] that have proven to be particularly effective on tissue classification with a priori known retina shape orientation, our ansatz merely relies on local features and yields ordered labelings which are directly enforced through the underlying geometry of statistical manifold. Consequently, by building on the feasible concept of spatially regularized assignment [22], the ordered flow (2) possesses the potential to be extended towards the detection of pathological retina changes and vascular vessel structure, which is the objective of our current research.

Acknowledgement. We thank Dr. Stefan Schmidt and Julian Weichsel for sharing with us their expertise on OCT sensors, data acquisition and processing. In addition, we thank Prof. Fred Hamprecht and Alberto Bailoni for their guidance in training deep networks for feature extraction from 3D data.

References

1. Abramoff, M.D., Garvin, M.K., Sonka, M.: Retinal Imaging and Image Analysis. *IEEE Reviews in Biomedical Engineering* 3, 169–208 (2010)
2. Antony, B., Abramoff, M., Lee, K., Sonkova, P., Gupta, P., Kwon, Y., Niemeijer, M., Hu, Z., Garvin, M.: Automated 3-D Segmentation of Intraretinal Layers from Optic Nerve Head Optical Coherence Tomography Images. *Progress in Biomedical Optics and Imaging - Proc.SPIE* 7626, 249 – 260 (2010)
3. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Geometric Means in a Novel Vector Space Structure on Symmetric Positive Definite Matrices. *SIAM Journal on Matrix Analysis and Applications* 29(1), 328–347 (2007)
4. Åström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image Labeling by Assignment. *Journal of Mathematical Imaging and Vision* 58(2), 211–238 (2017)
5. Bauschke, H.H., Borwein, J.M.: Legendre Functions and the Method of Random Bregman Projections. *J. Convex Analysis* 4(1), 27–67 (1997)
6. Censor, Y.A., Zenios, S.A.: *Parallel Optimization: Theory, Algorithms, and Applications*. Oxford Univ. Press, New York (1997)
7. Dice, L.R.: Measures of the Amount of Ecologic Association Between Species. *Ecology* 26(3), 297–302 (1945)
8. Duan, J., Tench, C., Gottlob, I., Proudlock, F., Bai, L.: New Variational Image Decomposition Model for Simultaneously Denoising and Segmenting Optical Coherence Tomography Images. *Physics in medicine and biology* 60, 8901–8922 (11 2015)
9. Dufour, P.A., Ceklic, L., Abdillahi, H., Schroder, S., De Dzanet, S., Wolf-Schnurrbusch, U., Kowal, J.: Graph-Based Multi-Surface Segmentation of OCT Data Using Trained Hard and Soft Constraints. *IEEE Transactions on Medical Imaging* 32(3), 531–543 (March 2013)
10. Fang, L., Cunefare, D., Wang, C., Guymer, R., Li, S., Farsiu, S.: Automatic Segmentation of Nine Retinal Layer Boundaries in OCT Images of Non-Exudative AMD Patients Using Deep Learning and Graph Search. *Biomed. Opt. Express* 8(5), 2732–2744 (May 2017)
11. Garvin, M.K., Abramoff, M.D., Wu, X., Russell, S.R., Burns, T.L., Sonka, M.: Automated 3-D Intraretinal Layer Segmentation of Macular Spectral-Domain Optical Coherence Tomography Images. *IEEE Transactions on Medical Imaging* (9), 1436–1447 (2009)
12. Haeker, M., Abramoff, M., Wu, X., Kardon, R., Sonka, M.: Use of Varying Constraints in Optimal 3-D Graph Search for Segmentation of Macular Optical Coherence Tomography Images. *MICCAI* 10, 244–51 (02 2007)
13. Hashimoto, M., Sklansky, J.: Multiple-Order Derivatives for Detecting Local Image Characteristics. *Computer Vision, Graphics, and Image Processing* 39(1), 28 – 55 (1987)
14. He, Y., Carass, A., Liu, Y., Jedynek, B., Solomon, S., Saidha, S., Calabresi, P., Prince, J.: Deep Learning Based Topology Guaranteed Surface and MME Segmentation of Multiple Sclerosis Subjects from Retinal OCT. *Biomed. Opt. Express* 10(10), 5042–5058 (Oct 2019)
15. Higham, N.: *Functions of Matrices: Theory and Computation*. SIAM (2008)
16. Kang, L., Xiaodong, W., Chen, D.Z., Sonka, M.: Optimal Surface Segmentation in Volumetric Images-A Graph-Theoretic Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(1), 119–134 (Jan 2006)

17. Liu, X., Cao, J., Fu, T., Pan, Z., Hu, W., Zhang, K., Liu, J.: Semi-Supervised Automatic Segmentation of Layer and Fluid Region in Retinal Optical Coherence Tomography Images Using Adversarial Learning. *IEEE Access* 7, 3046–3061 (2019)
18. Novosel, J., Vermeer, K.A., de Jong, J.H., Wang, Z., van Vliet, L.J.: Joint Segmentation of Retinal Layers and Focal Lesions in 3-D OCT Data of Topologically Disrupted Retinas. *IEEE Transactions on Medical Imaging* 36(6), 1276–1286 (June 2017)
19. Rathke, F., Schmidt, S., Schnörr, C.: Probabilistic Intra-Retinal Layer Segmentation in 3-D OCT Images Using Global Shape Regularization. *Medical Image Analysis* 18(5), 781–794 (2014)
20. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *MICCAI*. pp. 234–241. Springer International Publishing, Cham (2015)
21. Roy, A., Conjeti, S., Karri, S., Sheet, D., Katouzian, A., Wachinger, C., Navab, N.: ReLayNet: Retinal Layer and Fluid Segmentation of Macular Optical Coherence Tomography Using Fully Convolutional Networks. *Biomed. Opt. Express* 8(8), 3627–3642 (Aug 2017)
22. Schnörr, C.: Assignment Flows. In: Grohs, P., Holler, M., Weinmann, A. (eds.) *Variational Methods for Nonlinear Geometric Data and Applications*, pp. 235–260. Springer (2020)
23. Song, Q., Bai, J., Garvin, M.K., Sonka, M., Buatti, J.M., Wu, X.: Optimal Multiple Surface Segmentation With Shape and Context Priors. *IEEE Transactions on Medical Imaging* 32(2), 376–386 (Feb 2013)
24. Sra, S.: Positive Definite Matrices and the S-Divergence. *Proceedings of the American Mathematical Society* 144(7), 2787–2797 (2016)
25. Tuzel, O., Porikli, F., Meer, P.: Region Covariance: A Fast Descriptor for Detection and Classification. In: *Proc. ECCV*. pp. 589–600 (2006)
26. Yazdanpanah, A., Hamarneh, G., Smith, B.R., Sarunic, M.V.: Segmentation of Intra-Retinal Layers From Optical Coherence Tomography Images Using an Active Contour Approach. *IEEE Transactions on Medical Imaging* 30(2), 484–496 (Feb 2011)
27. Zeilmann, A., Savarino, F., Petra, S., Schnörr, C.: Geometric Numerical Integration of the Assignment Flow. *Inverse Problems* 36(3), 034004 (33pp) (2020)
28. Zern, A., Zeilmann, A., Schnörr, C.: Assignment flows for data labeling on graphs: Convergence and stability. *CoRR* abs/2002.11571 (2020)