Order Preserving and Shape Prior Constrained Intra-retinal Layer Segmentation in Optical Coherence Tomography

Fabian Rathke[1], Stefan Schmidt[2], and Christoph Schnörr[1,2]

[1] University of Heidelberg, Image & Pattern Analysis Group (IPA)
[2] Heidelberg Collaboratory for Image Processing (HCI)

Abstract. We present a probabilistic approach to the segmentation of OCT scans of retinal tissue. By combining discrete exact inference and a global shape prior, accurate segmentations are computed that preserve the physiological order of intra-retinal layers. A major part of the computations can be performed in parallel. The evaluation reveals robustness against speckle noise, shadowing caused by blood vessels, and other scan artifacts.

1 Introduction

Over the last years Optical Coherence Tomography (OCT) has become a key technique for non-invasive diagnostic retina imaging. By measuring the backscattering of light, OCT enables to produce high-resolution 2-D and 3-D scans of retinal tissues. Quantitative measurement of the intra-retinal layers plays a central role for the early diagnosis of diseases like glaucoma or age-related macular degeneration. Since manual segmentation is tedious and time-consuming, there is a high demand for automated algorithms – see Fig. 1.

Related Work. The literature on segmentation is vast. Closely related work includes *heuristics* to perform 1-D edge detection and to connect candidate points across image columns into consecutive boundaries, e.g. [1]. Another series of papers reformulate the segmentation problem as a *graph cut* problem, see [2] and references therein. By construction of the graph, the ordering of layers as well as the smoothness of boundaries are enforced. Edge weights are defined in terms of simple pixel-wise intensity and gradient based features. The approach takes several hours per 3-D volume. Recently, *active contour* approaches [3] were applied to OCT-segmentation. A functional is minimized that enforces circular boundaries (shape prior) and constant intensity within layers as well as smooth boundaries. Albeit being very fast (1s per 2-D scan), this approach requires the user to manually select the parameters.

Contribution. We present a novel probabilistic approach to the OCT segmentation problem. A global shape prior and a local appearance model are combined in a discrete graphical model which is solved for each image column separately,

G. Fichtinger, A. Martel, and T. Peters (Eds.): MICCAI 2011, Part III, LNCS 6893, pp. 370–377, 2011.

thus allowing for order preserving, exact and fast parallel inference. Few updates of the model parameters by iterative conditioning suffice to incorporate the global shape prior, trained offline using ground truth data. Evaluations using independent test data yield accurate segmentations and show robustness against strong speckle noise, shadowing caused by blood vessels and other scan artifacts.

Organization. In the next section, we present an OCT image model comprising a local appearance model and a global shape prior. We describe in Sect. 3 the coupling of both models. Sect. 4 introduces the dataset we used for evaluation and reports empirical results. We conclude in Sect. 5.

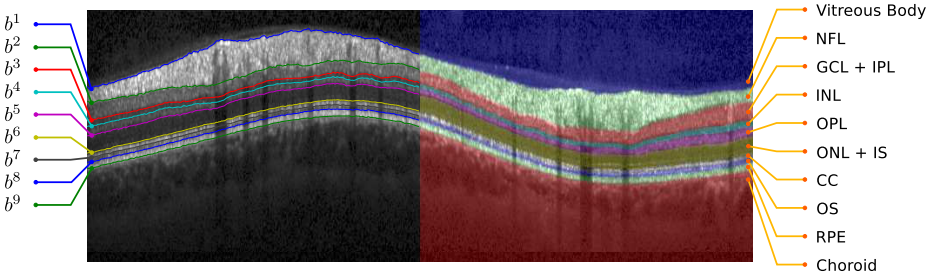

Fig. 1. An OCT B-Scan with the segmentation output of our approach: The right half shows labels $l^1, \ldots, l^{10}$ in different colors, the left half depicts the corresponding boundaries $b^1, \ldots, b^9$. The full names for layers $l^2, \ldots, l^9$ are: nerve fiber layer (NFL), ganglion cell layer and inner plexiform layer (GCL + IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer and inner segment (ONL + IS), connecting cilia (CC), outer segment (OS), retinal pigment epithelium (RPE).

2 An OCT Image Model

We model

- in Sect. 2.1 the *local appearance* in a given $N \times M$ image I of boundaries $b^1, \ldots, b^9$, and corresponding layers $l^2, \ldots, l^9$ located in between, and layers l^1, l^{10} located above b^1 and below b^9, respectively (all shown in Fig. 1);
- in Sect. 2.2 a *global joint shape prior* for all boundaries $b^1, \ldots, b^9$.

2.1 Local Appearance Model

We model the appearance of pixel values I_{ij} in terms of Gaussian Markov Random Fields (GMRFs) [4] for the corresponding patches $s(i, j)$ of size 3×15 around pixel (i, j).

For each class $k \in \{l^1, \ldots, l^{10}, b^1, \ldots, b^9\}$, we draw 1000 sample patches from labeled training images and convert them into vectors $s_i^k, i = 1, \ldots, 1000$, of size 1×45. Using these empirical data, we estimate for each k a class-specific density $\mathcal{N}(s; \mu_k, \Theta_k^{-1})$ with mean parameter μ_k and *sparse* precision matrix Θ_k, by applying a lasso penalty [5] as a regularizer.

We assign class variables m_{ij} to all pixels (i, j). Given an image I, we *define* the class-conditional likelihood of I_{ij} – with slight abuse of notation – to be the likelihood of the corresponding patch

$$p(I_{ij}|m_{ij} = k) := \mathcal{N}(s(i,j)|\mu_k, \Theta_k^{-1}), \quad k \in \{l^1, \ldots, l^{10}, b^1, \ldots, b^9\}. \tag{1}$$

2.2 Global Shape Model

As model of the typical shape variation of layers due to both biological variability as well as to the image formation process[1], we adopt a joint Gaussian distribution of the continuous height values of all boundaries $\{b^1, \ldots, b^9\}$ for all image columns j, that we denote by the $9M$-dimensional vector $b = (b_j^n)_{n=1,\ldots,9;\, j=1,\ldots,M}$. Hence,

$$p(b) = \mathcal{N}(b; \mu, \Sigma)\,. \tag{2}$$

We regularize the estimation of this high-dimensional model by Probabilistic Principal Component Analysis (PPCA) [6] with a preset number k_b of eigenmodes, which yields the representation $\Sigma = WW^T + \sigma^2 I$ with a low-rank matrix W. Given the spectral decomposition of the empirical covariance matrix estimate $\tilde{\Sigma} = U\Lambda U^T$, with eigenvalues $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_{9M})$ arranged in decreasing order, the maximum likelihood estimates of the shape prior parameters σ^2, W are given by $\sigma^2 = (9M - k_b)^{-1} \sum_{i=k_b+1}^{9M} \lambda_i$ and $W = U_b(\Lambda_b - \sigma^2 I)^{1/2}$, where U_b and Λ_b denote the submatrices of U and Λ corresponding to the k_b largest eigenvalues.

3 Model Fusion and Inference

We fuse our models of appearance and shape

- in Sect. 3.1 to obtain a discrete graphical model, for which globally optimal inference (determining positions of layer boundaries) can be efficiently done in parallel for all image columns, taking the order of layers and *marginal* shape prior knowledge into account;
- by incorporating the full global shape prior knowledge *across* image columns in two alternative ways, as described in Sect. 3.2.

In a preprocessing step, we compute local class variable distributions based on the local appearance model (1),

$$p(m_{ij}|I_{ij}) = \frac{p(I_{ij}|m_{ij})p(m_{ij})}{\sum_{m_{ij}} p(I_{ij}|m_{ij})p(m_{ij})} = \frac{p(I_{ij}, m_{ij})}{p(I_{ij})}, \tag{3}$$

using a uniform prior $p(m_{ij})$.

[1] For circular scans, a wave-like distortion pattern is observed due to the conic scanning geometry and the spherical shape of the retina, which we capture statistically rather than modelling it explicitly.

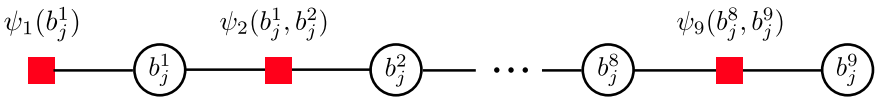

Fig. 2. The factor graph corresponding to the undirected graphical model in (4)

3.1 Order Preserving Inference and Marginal Shape Priors

For each image column $j \in \{1, \ldots, M\}$, we separately set up graphical models of the form

$$p(b_j|I_j) = \frac{1}{Z}\psi_1(b_j^1)\prod_{n=2}^{9}\psi_n(b_j^n, b_j^{n-1}), \tag{4}$$

in order to infer row positions of layer boundaries $b_j^n \in \{1, \ldots, N\}$, for all $n = 1, \ldots, 9$, conditioned on given column intensities $I_j := (I_{1j}, \ldots, I_{Nj})^T$. Here, we adopt common Markovian conditional independency assumptions, leading to the factorization (4) that is graphically depicted in Fig. 2. As a consequence, we can infer in parallel for all columns, and very efficiently, layer boundary positions b_j^n.

The factors in (4) are given by

$$\psi_1(b_j^1 = i) = p(m_{ij} = b^1|I_{ij})\, p(b_j^1 = i) \tag{5a}$$

$$\psi_n(b_j^n = i, b_j^{n-1} = l) = p(m_{ij} = b^n|I_{ij})\, p(b_j^n = i|b_j^{n-1} = l), \tag{5b}$$

and computed using (3) and by marginalizing out all columns but j using the global shape prior (2). Notice that the latter is a trivial operation for GMRFs [4] enabling the factorization (4) and, in turn, efficient inference.

Reading out the continuous shape prior at *discrete* row positions, and ignoring probability mass assigned to shape configurations that *violate* the natural order of layers, seems somewhat crude. Yet, this performs surprisingly well as we will demonstrate below and, after all, is computationally very efficient.

The output of the parallel inference process are expected layer positions

$$\hat{b}_j^n = \mathbb{E}[b_j^n], \qquad n = 1, \ldots, 9, \quad j = 1, \ldots, M, \tag{6}$$

computed using the distributions (4) and the *sum-product* algorithm for marginalization.

3.2 Iterative Conditioning and Global Shape Prior

An obvious shortcoming of the parallel approach (6) is that the global shape prior (2) is only exploited by marginal distributions of b_j, separately for all columns j, in (5).

In order to overcome this drawback, we investigated two different approaches.

1. The first one simply projects $\hat{b}$ given by (6) onto the latent PPCA subspace underlying the global shape prior (2) – see [6] for details – to obtain our final estimate $\hat{b}_{\text{proj}}$.

2. Alternatively, the second approach iteratively updates the observation probabilities $p(m_{ij}|I_{ij})$ in (5) using the evidence (6). Specifically, using the global shape prior (2), we compute marginals for all columns j by conditioning on (6) for all other columns $-j$, and update the observation probabilities by adding weighted pseudo-observations

$$p(m_{ij} = b^n|I_{ij}) \;\leftarrow\; p(m_{ij} = b^n|I_{ij}) + \frac{p_{\text{best}} - p_{\text{pred}}}{p_{\text{best}}}\,\delta_{i,\tilde{b}_j^n}, \qquad (7)$$
$$n = 1,\ldots,9, \qquad i = 1,\ldots,N$$

where $\delta_{i,\tilde{b}_j^n}$ is the Kronecker delta, and p_{pred} and p_{best} are conditional marginals evaluated for the *prediction* $\hat{b}_j^n$ and the (in terms of the shape model) *best* boundary $\tilde{b}_j^n = \text{argmax}_{i\in\{1,\ldots,N\}}\, p(b_j^n = i|\hat{b}_{-j})$ respectively.

Subsequently, we again infer (6) based on the terms (5) updated by (7), and iterate this process until $\hat{b}$ converges. Experiments show that this happens after few steps, and thus does not compromise computational efficiency.

4 Evaluation and Discussion

This section reports the performance of the approaches described in Sect. 2 and 3. To this end, we evaluated the following models:

1. *Uniform Prior*: $p(b_j^n = i|b_j^{n-1} = l) = \text{const.}, \forall\, i > l$, used in order to inspect how performance degrades *without* shape prior information.
2. *Shape Prior (SP)*: The graphical model described in Sect. 3.1.
3. *SP + Projection*: Post-processing $\hat{b}$ by projection onto the PPCA subspace, described in Sect. 3.2, alternative 1.
4. *SP + Iteration*: Post-processing $\hat{b}$ by iterative conditioning, described in Sect. 3.2, alternative 2.

Data Aquisition. Our dataset comprises 80 circular B-scans from 80 healthy persons with given ground truth labeling. The dataset was acquired by using a Heidelberg Engineering Spectralis HRA+OCT device. Each scan has a diameter of 3.4mm and consists of $M = 768$ A-scans of depth resolution $3.87\mu m$/pixel ($N = 496$ pixels).

Evaluation. For each boundary we compute the unsigned, signed and squared distances between the estimate $\hat{b}^n$ and the manual segmentation $\tilde{b}_j^n$ (ground truth),

$$E^n_{\text{unsgn}} = \sum_{j=1}^{M} \hat{b}_j^n - \tilde{b}_j^n, \qquad E^n_{\text{sgn}} = \sum_{j=1}^{M} |\hat{b}_j^n - \tilde{b}_j^n|, \qquad E^n_{\text{sq}} = \sum_{j=1}^{M} (\hat{b}_j^n - \tilde{b}_j^n)^2, \quad (8)$$

where M indicates the number of image columns (A-scans). The errors of the whole segmentation are computed as

$$E_* = (9M)^{-1} \sum_{n=1}^{9} E_*^n, \qquad * \in \{\text{unsgn}, \text{sgn}, \text{sq}\}. \qquad (9)$$

A 10-fold *nested* cross-validation over the lasso parameter (Sec. 2.1) was performed: First, the dataset was split into 10 subsets. For each subset (the *hold-out* set), the optimal parameter was determined by an *inner* 5-fold cross-validation on the remaining 9 subsets. Given this parameter, a model was trained using all 9 subsets and performance was evaluated on the hold-out set. Fig. 3 reports results in terms of E_{unsgn}, E_{sgn} and E_{sq} with error bars indicating standard deviation. Fig. 4 shows characteristic segmentation results for all four approaches in comparison.

Segmentation with the uniform prior performs worst, indicating the necessity of using a shape prior. Adding prior statistical information about the relative distances of layer boundaries (model 2. listed above) within each column boosts performance significantly. For many columns this approach finds the true boundary positions. Nevertheless, for columns with specific appearance, e.g. caused by blood vessels, segmentation may fail (see Fig. 4b). This reveals the lack of communication across image columns. Adding either PPCA projection (model 3.) or iterative conditioning (model 4.) as a post-processing step can resolve such issues (Fig. 4c and d).

As an illustration, Fig. 5 depicts the addition of pseudo-observations (7) for boundaries 6 and 9 in column j for the first round of iterative conditioning. Depending on the distance between $\tilde{b}_j^n$ (marked as green points) and $\hat{b}_j^n$, strong (boundary 6) or weak (boundary 9) pseudo-observations are added. As a consequence, estimates $\hat{b}_j^n$ that rely on weak observations and/or differ significantly from $\tilde{b}_j^n$ are corrected by the shape prior. While the iterative approach is working well for the lower boundaries b^6 - b^9, it may happen that estimates for b^2 - b^5 are too smooth, as shown for b^2 in Fig. 4d, where the shape prior somewhat tends to overrule the appearance model.

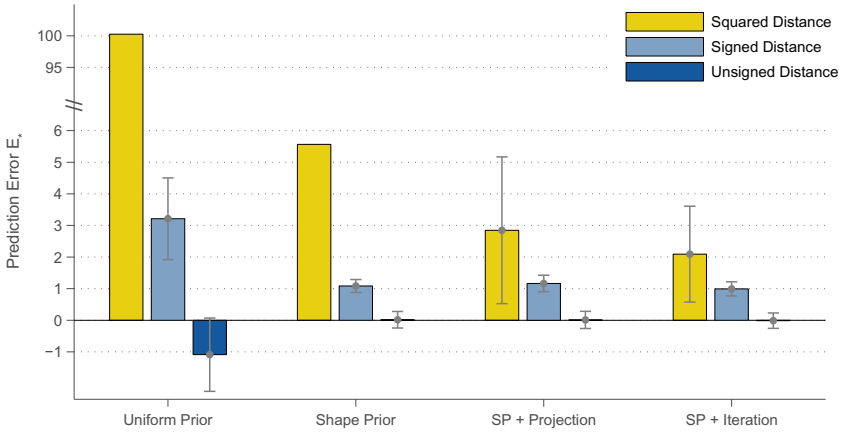

Fig. 3. Error means and standard deviation for all approaches 1., . . . , 4., evaluated for all 80 scans by means of a 10-fold nested cross-validation. The standard deviations of the squared error for the uniform prior and the shape prior model are 128.34 and 8.75 respectively.

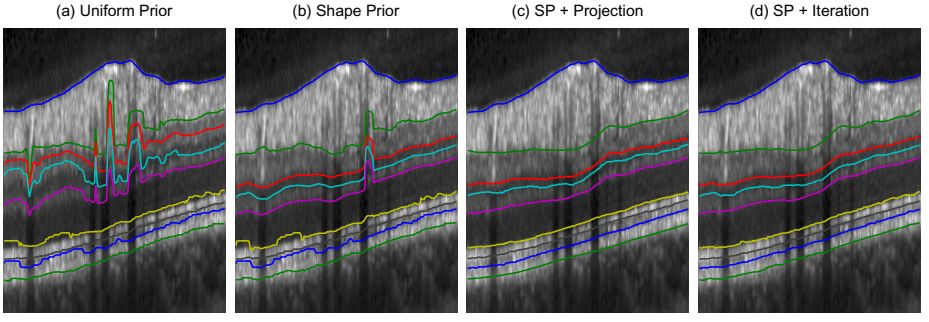

Fig. 4. Close-up view of segmentation results for the approaches $1., \ldots, 4.$. Left panel: no shape prior. Adding the shape prior significantly improves segmentation performance, especially for boundaries $b^2 - b^5$, but still may fail locally (2nd panel from left). Iteratively enforcing the full shape prior leads to good results (3rd and 4th panel).

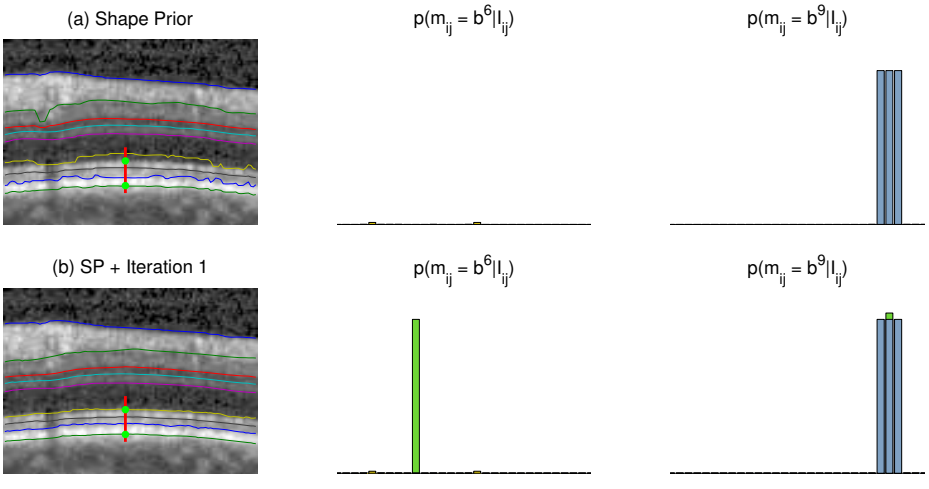

Fig. 5. The left panels show segmentations before (a) and after (b) the first step of iterative conditioning. Green points indicate $\tilde{b}_j^n$ for $n = 6, 9$ (see Sect. 3.2). The middle and right panels show the corresponding observation probabilities of both boundaries for the red marked segment of column j. Green bars indicate the magnitude of pseudo-observations added in (7), being stronger if the estimate $\hat{b}_j^n$ differs significantly from $\tilde{b}_j^n$.

Our current implementation in Matlab requires $\sim$ 15 seconds per 2D-scan on an Intel Core 2 Quad Q9550 with 2.83 Ghz.

5 Conclusion and Further Work

This work presented a novel probabilistic model for the segmentation of circular OCT scans. It combines a local appearance model and a global shape model using

a graphical model that enables very efficient computational inference. Comparison of different model variants has shown that utilizing the full shape prior performs best. The approach works without user interaction, is robust against appearance artifacts, and returns accurate segmentation results.

Our future work will investigate computationally tractable extensions that enable horizontal coupling of image columns not only by the appearance model but also by the shape prior. Furthermore, we will apply our approach to other types of OCT scans, e.g. 3D volumes or scans from the fovea region. The extension to 3D is straightforward: The shape prior as well as the texture model can be trained on 3D patches/boundaries, while the graphical model in Sec. 3 can be augmented to include columns (A-scans) of all slices in the volume. Finally, we will focus on discriminating non-/pathological scans and the development of statistical tools for assessing such images locally.

Acknowledgments. The authors would like to thank Heidelberg Engineering, for providing the OCT dataset as well as their employees Karsten Müller, Dr. Thomas Fendrich and Dr. Tilman Otto for helpful discussions.

This work has been supported by the German Research Foundation (DFG) within the programme "Spatio-/Temporal Graphical Models and Applications in Image Analysis", grant GRK 1653.

References

1. Koozekanani, D., Boyer, K., Roberts, C.: Retinal thickness measurements from optical coherence tomography using a Markov boundary model. IEEE T. Med. Imaging 20(9), 900–916 (2001)
2. Garvin, M., Abramoff, M., Wu, X., Russell, S., Burns, T., Sonka, M.: Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE T. Med. Imaging 28(9), 1436–1447 (2009)
3. Yazdanpanah, A., Hamarneh, G., Smith, B.R., Sarunic, M.V.: Segmentation of intraretinal layers from optical coherence tomography images using an active contour approach. IEEE T. Med. Imaging 30(2), 484–496 (2011)
4. Rue, H., Held, L.: Gaussian Markov Random Fields: Theory and Applications. Chapman & Hall/CRC, London (2005)
5. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9(3), 432–441 (2008)
6. Tipping, M.E., Bishop, C.M.: Probabilistic principal component analysis. J. R. Stat. Soc. 61(3), 611–622 (1999)