

Automated choroidal segmentation of 1060 nm OCT in healthy and pathologic eyes using a statistical model

Vedran Kajić,^{1,2,3,*} Marieh Esmaelpour,^{4,1} Boris Považay,¹ David Marshall,²
Paul L. Rosin,² and Wolfgang Drexler^{1,3}

¹Center for Medical Physics and Biomedical Engineering, Medical University Vienna,
General Hospital Vienna 4L, Waehringer Guertel 18-20, A-1090 Vienna, Austria

²School of Computer Science, Cardiff University, 5 The Parade, Cardiff, CF24 3AA, UK

³School of Optometry and Vision Sciences, Cardiff University, Maindy Road, Cardiff, CF24 4LU, UK

⁴Ludwig Boltzmann Institute of Retinology and Biomicroscopic Laser Surgery, Department of Ophthalmology, Rudolf
Foundation Clinic, Vienna, Austria

*kajicv1@cf.ac.uk

Abstract: A two stage statistical model based on texture and shape for fully automatic choroidal segmentation of normal and pathologic eyes obtained by a 1060 nm optical coherence tomography (OCT) system is developed. A novel dynamic programming approach is implemented to determine location of the retinal pigment epithelium/ Bruch's membrane /choriocapillaris (RBC) boundary. The choroid-sclera interface (CSI) is segmented using a statistical model. The algorithm is robust even in presence of speckle noise, low signal (thick choroid), retinal pigment epithelium (RPE) detachments and atrophy, drusen, shadowing and other artifacts. Evaluation against a set of 871 manually segmented cross-sectional scans from 12 eyes achieves an average error rate of 13%, computed per tomogram as a ratio of incorrectly classified pixels and the total layer surface. For the first time a fully automatic choroidal segmentation algorithm is successfully applied to a wide range of clinical volumetric OCT data.

© 2011 Optical Society of America

OCIS codes: (170.4500) Optical coherence tomography; (100.0100) Image processing; (100.3008) Image recognition, algorithms and filters; (170.4580) Optical diagnostics for medicine.

References and links

1. W. Drexler and J. G. Fujimoto, *Optical Coherence Tomography: Technology and Applications* (Springer, 2008), Vol. 1.
2. M. Esmaelpour, B. Povazay, B. Hermann, B. Hofer, V. Kajić, K. Kapoor, N. J. Sheen, R. V. North, and W. Drexler, "Three-dimensional 1060-nm OCT: choroidal thickness maps in normal subjects and improved posterior segment visualization in cataract patients," *Invest. Ophthalmol. Vis. Sci.* **51**(10), 5260–5266 (2010).
3. W. R. Green and S. N. Key 3rd, "Senile macular degeneration: a histopathologic study," *Trans. Am. Ophthalmol. Soc.* **75**, 180–254 (1977).
4. D. S. McLeod and G. A. Luty, "High-resolution histologic analysis of the human choroidal vasculature," *Invest. Ophthalmol. Vis. Sci.* **35**(11), 3799–3811 (1994).
5. S. H. Sarks, "Ageing and degeneration in the macular region: a clinico-pathological study," *Br. J. Ophthalmol.* **60**(5), 324–341 (1976).
6. Z. Q. Yin, T. J. Vaegan, T. J. Millar, P. Beaumont, and S. Sarks, "Widespread choroidal insufficiency in primary open-angle glaucoma," *J. Glaucoma* **6**(1), 23–32 (1997).
7. M. Esmaelpour, B. Povazay, B. Hermann, B. Hofer, V. Kajić, S. L. Hale, R. V. North, W. Drexler, and N. J. Sheen, "Mapping choroidal and retinal thickness variation in type 2 diabetes using three-dimensional 1060-nm optical coherence tomography," *Invest. Ophthalmol. Vis. Sci.* **52**(8), 5311–5316 (2011).
8. D. Cabrera Fernández, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Opt. Express* **13**(25), 10200–10216 (2005).
9. A. K. Mishra, P. W. Fieguth, and D. A. Clausi, "Decoupled active contour (DAC) for boundary detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(2), 310–324 (2011).
10. J. Molnár, D. Chetverikov, D. Cabrera DeBuc, W. Gao, and G. Somfai, "Layer extraction in rodent retinal images acquired by optical coherence tomography," *Mach. Vis. Appl.* (2011).

11. M. Mujat, R. Chan, B. Cense, B. Park, C. Joo, T. Akkin, T. Chen, and J. de Boer, "Retinal nerve fiber layer thickness map determined from optical coherence tomography images," *Opt. Express* **13**(23), 9480–9491 (2005).
12. M. V. Sarunic, A. Yazdanpanah, E. Gibson, J. Xu, Y. Bai, S. Lee, H. U. Saragovi, and M. F. Beg, "Longitudinal study of retinal degeneration in a rat using spectral domain optical coherence tomography," *Opt. Express* **18**(22), 23435–23441 (2010).
13. K. A. Vermeer, J. van der Schoot, H. G. Lemij, and J. F. de Boer, "Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images," *Biomed. Opt. Express* **2**(6), 1743–1756 (2011).
14. M. A. Mayer, J. Horneegger, C. Y. Mardin, and R. P. Tornow, "Retinal Nerve Fiber Layer Segmentation on FD-OCT Scans of Normal Subjects and Glaucoma Patients," *Biomed. Opt. Express* **1**(5), 1358–1383 (2010).
15. F. Rossant, I. Ghorbel, I. Bloch, M. Paques, and S. Tick, "Automated segmentation of retinal layers in OCT imaging and derived ophthalmic measures," in *IEEE International Symposium on Biomedical Imaging: from Nano to Macro, 2009. ISBI '09* (IEEE, 2009), pp. 1370–1373.
16. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," *Opt. Express* **18**(18), 19413–19428 (2010).
17. D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements from optical coherence tomography using a Markov boundary model," *IEEE Trans. Med. Imaging* **20**(9), 900–916 (2001).
18. Q. Yang, C. A. Reisman, Z. Wang, Y. Fukuma, M. Hangai, N. Yoshimura, A. Tomidokoro, M. Araie, A. S. Raza, D. C. Hood, and K. Chan, "Automated layer segmentation of macular OCT images using dual-scale gradient information," *Opt. Express* **18**(20), 21293–21307 (2010).
19. A. Yazdanpanah, G. Hamarneh, B. Smith, and M. Sarunic, "Intra-retinal layer segmentation in optical coherence tomography using an active contour approach," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2009* (Springer-Verlag, London, UK, 2009), pp. 649–656.
20. M. K. Garvin, M. D. Abramoff, R. Kardon, S. R. Russell, X. Wu, and M. Sonka, "Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search," *IEEE Trans. Med. Imaging* **27**(10), 1495–1505 (2008).
21. D. Tolliver, Y. Koutis, H. Ishikawa, J. S. Schuman, and G. L. Miller, "Unassisted segmentation of multiple retinal layers via spectral rounding," in *ARVO 2008 Annual Meeting* (Association for Research in Vision and Ophthalmology, 2008), poster.
22. R. J. Zawadzki, S. S. Choi, S. M. Jones, S. S. Oliver, and J. S. Werner, "Adaptive optics-optical coherence tomography: optimizing visualization of microscopic retinal structures in three dimensions," *J. Opt. Soc. Am. A* **24**(5), 1373–1383 (2007).
23. B. Povazay, B. Hermann, B. Hofer, V. Kajić, E. Simpson, T. Bridgford, and W. Drexler, "Wide-field optical coherence tomography of the choroid in vivo," *Invest. Ophthalmol. Vis. Sci.* **50**(4), 1856–1863 (2009).
24. P. Thevenaz and M. Unser, "A pyramid approach to sub-pixel image fusion based on mutual information," in *International Conference on Image Processing, 1996. Proceedings* (1996), vol. 1, pp. 265–268.
25. I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," *IEEE Signal Process. Mag.* **22**(6), 123–151 (2005).
26. I. Perkon, A. Košir, J. Tasič, and M. Diamond, "Whisker detection as a shortest path problem," presented at Visual Observation and Analysis of Animal and Insect Behavior, Istanbul, Aug. 22, 2010.
27. V. Kajić, B. Povazay, B. Hermann, B. Hofer, D. Marshall, P. L. Rosin, and W. Drexler, "Robust segmentation of intraretinal layers in the normal human fovea using a novel statistical model based on texture and shape analysis," *Opt. Express* **18**(14), 14730–14744 (2010).
28. M. Scholz, M. Fraunholz, and J. Selbig, "Nonlinear principal component analysis: neural network models and applications," in *Principal Manifolds for Data Visualization and Dimension Reduction* (Springer, 2007), pp. 44–67.
29. M. Scholz, F. Kaplan, C. L. Guy, J. Kopka, and J. Selbig, "Non-linear PCA: a missing data approach," *Bioinformatics* **21**(20), 3887–3895 (2005).
30. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 681–685 (2001).
31. J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vis. Comput.* **22**(10), 761–767 (2004).
32. A. Vedaldi and B. Fulkerson, "VLFeat: an open and portable library of computer vision algorithms" (2008), <http://www.vlfeat.org/>.

1. Introduction

Optical coherence tomography (OCT) [1] is a biomedical imaging method using coherent infrared light that enables high resolution, three-dimensional (3D) sub-surface insight into living tissue utilizing white-light interferometry. The simple optical access to the light sensitive retina and choroid, which is challenging to accomplish by other high resolution methods, lead to wide clinical application of this imaging technique. Currently available OCT systems can obtain an axial resolution of less than 3 μm utilizing the wavelengths at 800 nm, and enabling imaging of the retinal and choroidal microstructure using the 1060 nm wavelength band. A typical scan takes less than 10 seconds and produces a large volumetric data set constitutes of a stack of high resolution cross-sectional tomograms (B-scans).

The choroid is a densely vascularized layer under the retinal pigment epithelium's (RPE). Its deeper boundary is formed by the sclera, the outer fibrous shell of the eye. The choroids structure and thickness are linked to many parameters such as age or axial eye length [2] but also seem to have a strong relation to pathologies: histology of choroidal thickness (ChT) reveals an increase that correlates with a higher density of vessels in the superficial choroidal layers in open-angle glaucoma, atrophy of the choroidal capillary structure, and neovascularization in the choriocapillaris [3–6]. ChT-maps may give valuable insight in pathologies and therapy monitoring. Previously used manual segmentation [2,7] proved too time consuming to be practical in a clinical setting and presented a great difficulty when working with numerous data sets needed for statistical analysis. However, the inhomogeneity within the layers of choroidal OCT-tomograms presents a significant challenge to existing segmentation algorithms, including those successfully implemented for retinal layers. To the best of authors knowledge, no segmentation method has been published so far that successfully automatically segments the choroid from 1060 nm OCT without manual intervention.

Previous OCT segmentation algorithms do not perform well with respect to the choroidal segmentation requirements. Retinal segmentation algorithms vary depending on the number of layers to be segmented and on their robustness in the presence of speckle, shadows, morphologic irregularities (i.e. vessels, physiologic structural changes at the fovea and optic nerve head) and pathological changes in the tissue. In general most of the existing approaches tend to be very sensitive to noise or missing data and often rely on well-contrasted boundaries or uniform layer structure. Intensity threshold based algorithms utilize simple analysis along the depth scan's intensity profiles or find borders by investigating intensity gradients and are frequently confused by missing data, which leads to non-physiologic results [8–15]. Thus, they are unsuitable for choroidal segmentation.

Boundary tracking methods utilize Markov models or dynamic programming (i.e. Dijkstra's algorithm) [16–18]. Even though more robust than standard column-wise thresholding methods, they still rely on connecting 1D points. That makes this class sensitive to weak signal conditions and noise, thus detected layer boundaries can easily drift off from the real ones.

An active contour approach has been used for retinal segmentation [19], minimizing a region term and assuming consistent image intensity within one layer, with added terms for shape prior and a regularization term to enforce smoothness. The initial position of the boundaries is user initialized, followed by the least squares optimization. Apart from this approach being semi-automatic rather than fully user independent, it also assumes that layers are of uniform brightness, which is not true in case of choroid segmentation. It is also not likely to perform well without prior knowledge of the likely shape of choroid, judging on our experiments using level sets.

Graph cuts have been applied to retinal segmentation [20]; the optimal graph cut was performed with weights representing both edge and regional information. Similarly, spectral rounding [21], a graph-partitioning algorithm that is based on the eigenvector calculation to determine the oscillation steps that represent the retinal edges, performed well in retinal segmentation, considering that no a priori information was available to the algorithm, dividing an image iteratively along the oscillation boundary. However, similarly as with active contour approaches, it is unlikely that this algorithm can extract heterogeneous structures, such as the choroid, without using some structural information.

A semi-automatic algorithm [22] has been used for OCT segmentation which requires user to paint the areas of interest in any slice of the volume. A support vector machine (SVM) was used for segmentation with a feature vector that contained intensity, location, mean of the neighborhood, standard deviation and gradient magnitude. However, for a high clinical throughput a fully automatic approach is preferable. Thus, all of the aforementioned methods suffer from one or more of the following disadvantages: they segment only the most prominent layers, do not exhibit robustness in noisy and varied cases and/or require manual intervention of the operator. The proposed method of the present paper uses training data

obtained from manual segmentations by human operators to build a statistical model that is able to actively learn and determine the plausible solutions in a low signal, noisy environment. Unlike most of the other approaches, it operates without relying on boundary edge information, instead it uses regional information. That is essential to choroidal segmentation as the CSI boundary is frequently very weak or nonexistent. During the learning stage parameters defining the lower choroidal boundary (CSI) of a statistical model are extracted so that it best fits the training data. This approach offers greater flexibility over using fixed constraints on layer smoothness, since it learns from the data what amount of variability is possible and in what regions, while on the other hand, it constrains the model to a plausible space of states.

2. Materials and methods

High-speed 3D OCT-imaging at 1060 nm was performed with less than 2.5 mW at the cornea, well below the maximum power limit for 10-second exposure [23]. Three-dimensional OCT volumes were acquired with 15 to 20 μm transverse resolution, approximately 7 μm axial resolution, and 512 voxels per depth scan (A-scan). Scans across a 36° by 36° field with 512 by 512 A-scans at 47,000 A-scans/s were centred on the fovea and resulted in to 70 frames/s (B-scans/s). The spectrometer utilized a line camera with 1024 px (Goodrich SU-LDH2) clocked at 92 kHz linerate, which led to 512 voxel in each depth scan between the zero delay and the maximal depth. The signal-loss of -14dB along the depth scan was counteracted by placing the closer, less attenuated side of the tissue towards the zero delay. For registration and noise reduction ImageJ software was used [24]. The OCT volume was averaged in both transverse directions within a field of approximately 1° , to remove speckle and increase sensitivity. Axial choroidal thickness (ChT) was defined as the distance between the center of the signal peaks originating from the RPE/Bruch's membrane/choriocapillaris (RBC) complex and the choroidal-sclera interface (CSI). Before the segmentation process, dual-tree complex wavelet (DTCW) denoising is applied to the data. The denoising algorithm exhibits good performance, while being computationally efficient [25]. This reduces the speckle present and thus makes the subsequent segmentation tasks easier.

A total of 871 B-scans, 691 from diabetes 1 and 180 from advanced pathologies, were manually segmented and evaluated against the proposed algorithm. We have used 7 manually segmented stacks, uniformly sampled, of diabetes type 1 patients. 128 B-scans per eye were available of which 100 (apart from two stacks that had 98 and 93 B-scans segmented respectively) were manually segmented by an expert. 28 B-scans at stack start (14) and end (14) were not segmented due to very low signal in that region. The algorithm was also evaluated on 5 different pathologies with 20 slices manually segmented per eye (apart from one stack which had full 100). Testing the performance on various pathologies, that often exhibit very weak RBC boundaries combined with choroid of changed shape and structure, demonstrates versatility and robustness of the proposed algorithm. The algorithm was trained on a completely independent data set of normal choroidal data.

2.1. The algorithm overview

The pre-processing stage is performed for both the training step and the segmentation of the unseen data. Once the variation parameters have been learned from the manually segmented training data, they can be used to drive the model to perform segmentation of unseen data. The actual segmentation process is essentially an optimization run that adjusts the model parameters in order to minimize the objective function that defines the difference between the model and a given unseen image that is to be segmented. The overview is shown in Fig. 1.

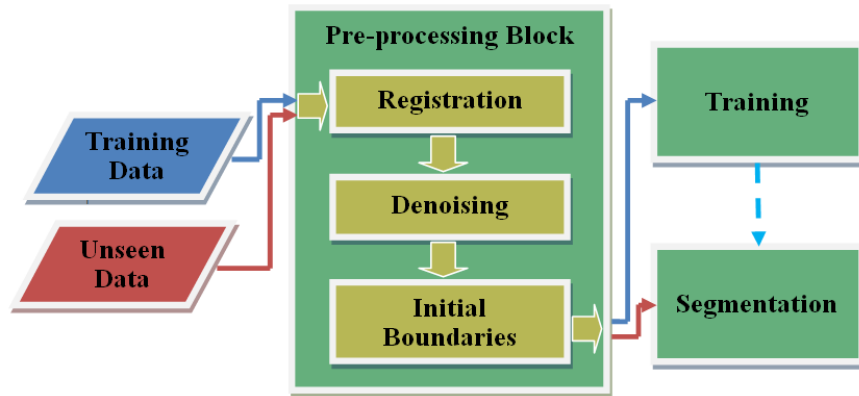


Fig. 1. Algorithm overview: both the training and unseen data pass through the pre-processing block. The registration is used only for averaging in z-direction as a despeckling procedure, while the segmentation is independent for each B-scan. A statistical model is subsequently constructed that captures the variance in the training data, which can be then used (dotted line) to segment unseen data.

2.2. Initial Boundaries Detection

To successfully segment internal limiting membrane (ILM) and RBC boundaries of choroidal data several challenges have to be overcome. RBC defines the top choroidal boundary and is used as initialization for the statistical model that subsequently finds the CSI (choroidal sclera interface) boundary. ILM is relatively easy to find and is of clinical use, thus it is segmented as well.

The RBC is usually almost at the same position as RPE, but in cases of RPE detachment it is often not (i.e., Fig. 10 below). The OCT data obtained was from a 1060 OCT system by scanning over wide angle [23], which causes low signal strength on the image edges, especially towards the stack start and end, where the retina is narrow (Fig. 2a), the right side of the image). Another problem were various pathologies that present different challenges; the boundaries can be discontinuous, also their shape is highly unpredictable and of highly changing curvature, so it is impossible to efficiently parameterize them, using a small set of parameters, with polynomials or other models that can be used for fitting. The third problem is that the data itself is often coming from different OCT systems and pre-processing pipelines; obviously it is highly beneficial if the correct segmentation can be achieved without retuning the algorithm parameters, thus a robust approach is required.

To meet all these goals an algorithm has been devised that consists of several basic steps, however special care was taken to keep the number and importance of free parameters as small as possible. This makes the algorithm intuitive and general while avoids over specializing.

First non-signal areas at the edges of an input image (Fig. 2 a) are found by performing Canny edge filtering using large sigma (32) and high threshold (0.35). The parameters were heuristically chosen to extract only the major edges in an image. Columns on the sides where there is no edge detected are considered to be non-signal areas. The first column with an edge detected is considered to be the start of the signal area; this is determined both from the left and the right side (Fig. 2 b).

A modified Dijkstra's algorithm is used to find a minimal path from the left to the right edge of an image, representing the RPE boundary, as it is almost always the brightest and longest one (in the sense that it extends furthest to the sides of the image). It is important to note that this algorithm has been used only to find the RPE which is a high intensity boundary, unlike some previous approaches which used path finding algorithms to find all the retinal boundaries, making the segmentation sensitive to any changes in the input data.

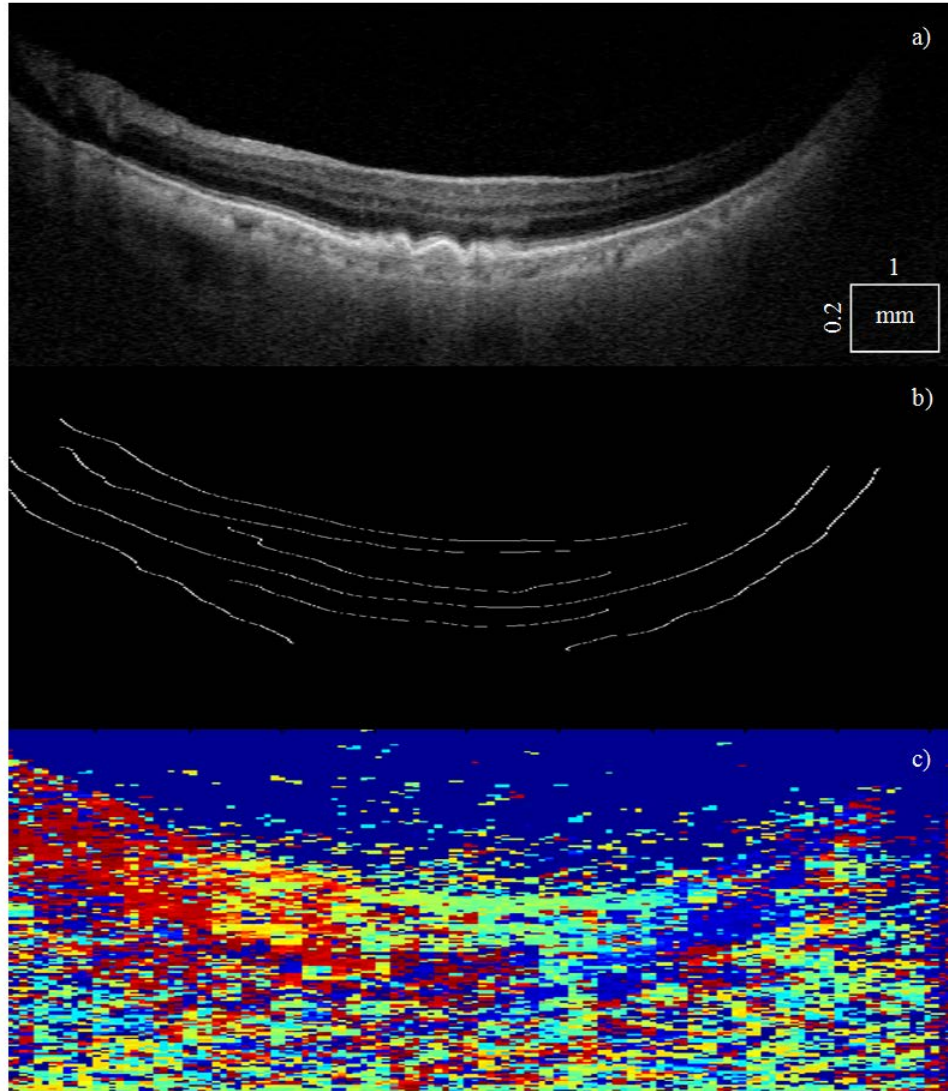


Fig. 2. Pre-processing steps for finding ILM and RBC: a is the original image; b is obtained after Canny edge filtering; c shows color coded dominant wavelet domain orientations (from dark red “descending” (−90 deg) angle, to dark blue “ascending” (90 deg)).

A graph is constructed, where each node represents the inverted pixel intensity in an image (i.e. the dark pixels become bright and bright pixels dark). Since the path to be found needs to be minimal, traversing bright pixels will mean going through low intensity values in the inverted image. A diagonal block connectivity matrix \mathbf{A} consists of $m * n$ rows, where m is height and n is width of an image \mathbf{I} . Thus, every row represents connectivity of certain pixel to pixels in the next column of the original image. Considering that the boundary cannot track backwards or skip columns, matrix \mathbf{A} has only m columns, as each pixel is connected only to the next column of the original image. In this case it was assumed that plausible steps from column to column do not exceed 10 pixels, thus jmp was set to 10, Eq. (1). Vector \mathbf{c} represents connectivity of a particular node (pixel) to other nodes and it is defined within the jmp range, with the rest of pixels set to 0 (no connection):

$$\mathbf{A} = \begin{pmatrix} \mathbf{c}_{11} & & & & 0 \\ \vdots & \ddots & & & \\ 0 & & \mathbf{c}_{1m} & & 0 \\ \vdots & & & \ddots & \\ 0 & & & & \mathbf{c}_{n1} & 0 \\ \vdots & & & & & \ddots \\ 0 & & & & & & \mathbf{c}_{nm} \end{pmatrix}, \quad (1)$$

$$\mathbf{c}_{ij} = 1 - \mathbf{I}(x, j+1), \max(0, i-jmp) < x < \min(n, i+jmp).$$

To increase robustness of this approach it was expanded in two ways. Firstly, derivatives were added to prevent unrealistic boundary jumps caused perhaps by just slightly more favorable intensity information.

Second, a complex dual-tree wavelet decomposition [25], being multi-scale, orientation sensitive and translation invariant, making it more efficient than standard edge enhancement methods, such as Sobel, Prewitt, Roberts or Canny, was used to obtain the edge orientation information (Fig. 2 c). Smooth dominant directions were calculated from -90 (dark red) to 90 (dark blue) degrees, which was applied in the modified algorithm to obtain the third term expressing how well the path is aligned with the real edges in an image. This prevents path drifting over high intensity areas that possibly have completely different orientation, thus producing a non-realistic path.

Unlike working with static weights and having a complicated workaround the problem that each node can be part of multiple paths and thus cannot be assigned an optimal weight using static values [26], the problem has been solved by making an elegant extension to the optimal path algorithm itself, while preserving the same time complexity.

Equations (2) show the modified algorithm distance function; sc is the current node index, corresponding to any pixel in the original image, xc and yc are x and y coordinates obtained of the node sc , \mathbf{y} are the neighboring column's connected nodes (pixels), yp is sc 's parent's y value (inherently always known from the path history), \mathbf{y}'' is the second derivative of the potential path, while $\overline{\mathbf{y}'}$ is the derivative mean used as an approximation to obtain the angles \mathbf{j} of the potential path. Weights \mathbf{ew}_1 and \mathbf{ew}_2 indicate how close to each of the two neighboring out of 6 total discrete orientations possible path is. \mathbf{IW} is the tensor containing directional wavelet coefficients (amplitudes) for each of the six directions, that can be index using the potential path directions \mathbf{j} . n is the number of wavelet levels used, for our application it was set to 4.

$$d(sc) = w_1 \mathbf{A}(sc, \mathbf{y}) + w_2 |\mathbf{y}''| + w_3 \left(1 - \frac{\mathbf{ew}_1 \odot \sqrt[n]{\mathbf{IW}(yc, xc, \mathbf{j}_1)} + \mathbf{ew}_2 \odot \sqrt[n]{\mathbf{IW}(yc, xc, \mathbf{j}_2)}}{\sum \sqrt[n]{\mathbf{IW}(yc, xc, \phi_i)}} \right)$$

$$xc = \left\lfloor \frac{sc-1}{M} \right\rfloor$$

$$yc \equiv sc-1 \pmod{M}$$

$$\mathbf{y} = \mathbf{A}(sc, i) > 0$$

$$\mathbf{y}'' = (\mathbf{y} - yc - (yc - yp)) = (\mathbf{y} - 2 * yc + yp)$$

$$\overline{\mathbf{y}'} = \frac{\mathbf{y} - yc + (yc - yp)}{2} = \frac{\mathbf{y} - yp}{2}$$

$$\mathbf{ew}_1 = \frac{\mathbf{j}_2 \cdot \mathbf{j}}{\mathbf{j}_2 \cdot \mathbf{j}_1}$$

$$\mathbf{ew}_2 = \mathbf{1} - \mathbf{ew}_1$$

The weights w of the distance measure's three terms have been set heuristically to 1, 0.08 and 0.2 respectively; the algorithm is robust enough that the exact weight values are not very important. For some other application another values might be more appropriate. Having the distance measure defined between each pair of nodes in the input image, the task of finding the optimal path is simply that of finding a minimum sequence of distances from end to end.

After the initial estimate of RPE boundary has been found (Fig. 3 a), a convex hull is fitted to the points that comprise the RPE boundary (Fig. 3 b). A formal definition of the convex hull for a set of points X is the minimal area convex set containing X .

Fitting the convex hull allows for an efficient estimation of which regions should be interpolated over. The convex edges are always left intact, while concave regions (intuitively dents) are considered for interpolation. Following the efficient representation of the problem allowed for a simple statistical measure to be developed. In general, any segment which has a high boundary curvature, or is significantly shifted from the estimate of the convex shape is likely to be caused by pathological detachment of the RPE (i.e. drusen). Thus, an assumption can be made that the position of the RPE does not mark the start of the choroid, while in most cases a cubic interpolation gives a good estimate of the choroid. While the RBC estimate produced by the interpolation over these nonconvex regions is quite close to the boundary determined by experts, it is not always completely correct. For the simplicity and rarity of these occurrences, the algorithm was found to be performing sufficiently well even with simple interpolation. However, a more complex approach can be devised. After finding the cubic spline estimate of the start of the choroid region, a neighborhood of the spline can be considered for a second run of the path finding algorithm. This will remove the area where the detached RPE is located, allowing for a more precise boundary estimate. That boundary is often very weak, and in some cases it is even marked by a black edge rather than a white one. Thus, for that run of the path finding algorithm, a more complex edge detector would have to be considered, instead of simply using intensity values. A straight forward approach would be

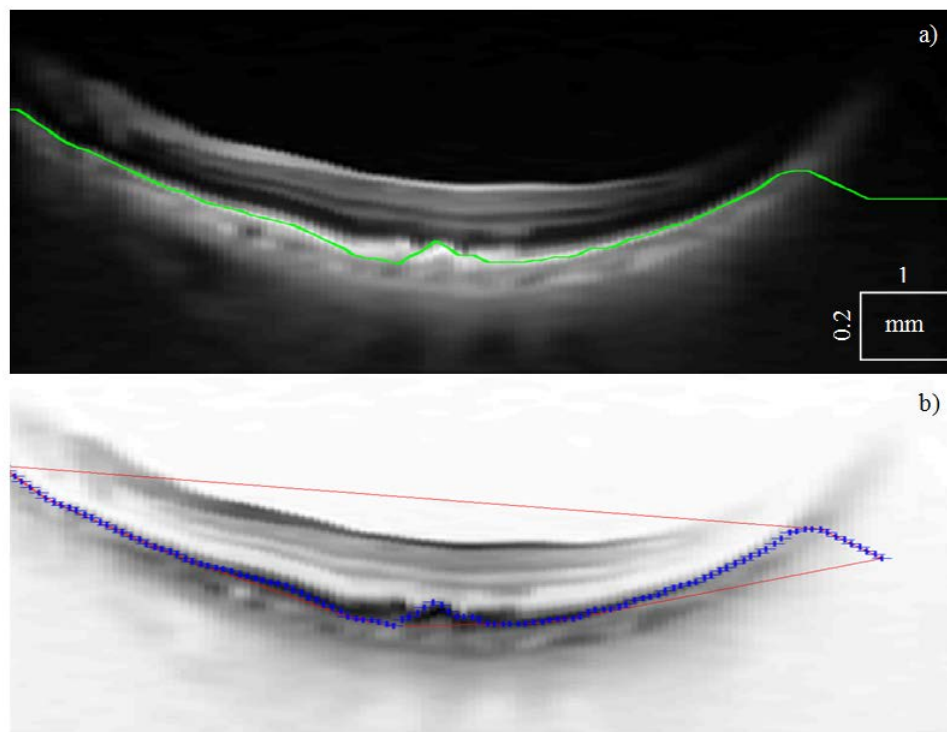


Fig. 3. RPE estimate (green line and blue dots) and convex hull (red line).

to use a trained texture descriptor to obtain edge probabilities. If encoded as an intensity image, the same shortest path procedure could be subsequently applied.

Once the RBC boundary has been found, the neighborhood and the image areas below it are removed (current position-50 pixels, to the bottom of the image). In the next step this image is used to find the real ILM position. Columns in which no edge was found signify gaps in the ILM, as the removal of the image signal around and below the RPE will result in no edge detection in the whole column (ILM and RPE being the brightest boundaries). The final ILM boundary is then found using piecewise cubic spline interpolation.

At the non-signal areas of the image, RBC is set as straight, while ILM is set to the value of the RPE (Fig. 4). Additionally, the program returns the non-signal area positions, so that the statistical model can use it later to find the choroidal-scleral interface, as at the non-signal areas of the image that boundary is set to the RBC value as well and is excluded from the optimization run (thus noise at the sides does not confuse the optimization function).

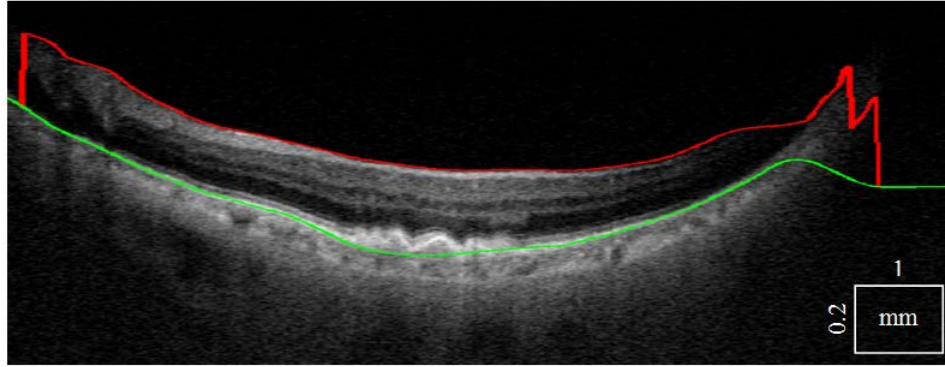


Fig. 4. The final result: ILM (red) and RBC (green).

2.3. Model building

After the pre-processing stage the statistical model, similar to the one already applied to the robust foveal segmentation [27], is first trained on a set of manually segmented images and can be then applied to the unseen data. Its main advantage is that knowledge of the problem can be used to resolve the confusion caused by structural complexity, provide tolerance to noisy or missing data, and provide a means of labeling the recovered structures. The basic idea is to perform supervised learning by applying knowledge of the expected shapes of structures, their spatial relationships, and their textural appearance to restrict the automated system to plausible interpretations. Supervised learning is a type of machine learning based on training data, which consists of pairs of input objects, and desired outputs. The task of the supervised learner is to predict the value of the function for any valid input object after having seen a number of training examples. To be useful, a model needs to be specific, representing only legal examples of the modeled object.

From the manually segmented images we extract the shape features and for each image we put all the extracted shape features into a vector. If we have m training images, for each layer (n layers) we get one vector of offsets \mathbf{v} per layer, per image of width w , which stacked together for all the layers define \mathbf{x} . All of the manual segmentations then comprise the matrix \mathbf{X} Eq. (3).

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_m \end{pmatrix} = \begin{pmatrix} \mathbf{v}_{11} & \cdots & \mathbf{v}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{v}_{m1} & \cdots & \mathbf{v}_{mn} \end{pmatrix}, \quad (3)$$

$$\mathbf{v}_{ij} = [\text{off}_1 \dots \text{off}_w].$$

Shape features that are used are sparsely sampled distances (at 26 positions) of the boundaries from the RBC. Thus we have 26 spatial features as there is only one boundary being modeled.

Statistical models can reproduce specific patterns of variability in shape and texture by analyzing the variations in shape across the training set. The key step of the statistical model training is the dimensionality reduction of the large set of features from the training data set. The reason for dimensionality reduction is to make feasible the computational cost of the optimization method that is used to fit the model to the real data later on. The idea behind this concept is to find statistical dependencies between the produced features and reduce the dimensionality of the space by identifying only a certain number of the most prominent properties in the data set, represented by the most important eigenvectors.

Principal component analysis (PCA) is the standard vector space transform technique used to reduce multidimensional data sets to lower dimensions for analysis. It operates by calculating the eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix. Hopefully a relatively small number of eigenvectors with greatest eigenvalues will describe the original data well. If \mathbf{X} is the original data matrix, as defined in Eq. (3), after the decomposition, where \mathbf{W} and \mathbf{V}^T are unitary matrices while Σ is a diagonal matrix, we can select only L principal components and in that way project the data into a reduced dimensionality space to get \mathbf{Y} , Eq. (4):

$$\begin{aligned}\mathbf{X} &= \mathbf{W}\Sigma\mathbf{V}^T \\ \mathbf{Y} &= \mathbf{W}_L^T \mathbf{X}\end{aligned}\quad (4)$$

However, rather than PCA, we used neural network based dimensionality reduction since it offers nonlinear eigenvectors (components) and therefore can reduce the space more compactly if the data is nonlinearly distributed than the linear representation obtained by PCA [28]. The shape features proved to be nonlinear and thus we obtained a more compact representation using nonlinear dimensionality reduction. A Neural network (NN) is a mathematical or computational model based on principles found in biological neural networks. It consists of an interconnected group of artificial neurons and processes information where each connection between neurons has a weight, with the weights modulating the value of the connection. The training phase is performed to modify the weights until the network implements a desired function. Once training has completed, the network can be applied to data that was not part of the training set. A special type of neural network (inverse) [29] can be used to perform dimensionality reduction on the training feature set that is produced which contains missing values. Missing values occur when no data value is stored for the variable in the current observation. The generating function Φ_{gen} produces larger dimensionality data \mathbf{X} from the parameters \mathbf{z} (lower dimension, equivalent to \mathbf{W}_L in PCA) Eq. (5). The extraction function Φ_{extr} does the reverse:

$$\begin{aligned}\Phi_{gen} : \mathbf{z} &\rightarrow \hat{\mathbf{X}} \\ \Phi_{extr} : \mathbf{X} &\rightarrow \mathbf{z}\end{aligned}\quad (5)$$

We reduced the dimensionality of the original spatial feature space from 26 to 8. This number of features allowed for an efficient optimization in the subsequent steps, while still preserving the original data variation well.

Our approach is based on a similar concept to the Active Appearance Model (AAM). For completeness, it will be first explained how the basic AAM model is constructed, followed by an explanation of how the proposed statistical model differs from that concept. An Active Appearance Model (AAM) is a model capable of synthesizing new images of the object of interest by finding the model parameters which generate a synthetic image as close as possible to the target image [30]. An AAM will, based on learned shape variation, generate a new image with a texture learned from the texture variation and then compute the distance between the synthesized and the given image that is to be segmented. \mathbf{x} is the shape vector (which is

normalized by subtracting the mean shape and rescaling, Eq. (6)) and \mathbf{g} is the texture vector obtained by the texture generating function G from an image \mathbf{I} and the shape vector (it is also normalized) Eq. (7).

$$\mathbf{x} \rightarrow (\mathbf{x} - \mu(\mathbf{x})) / \sigma(\mathbf{x}) \quad (6)$$

$$\mathbf{g} = G(\mathbf{x}, \mathbf{I}) \quad (7)$$

Function $S(\mathbf{s})$ produces new shape vectors by adding the shape parameters \mathbf{s} multiplied by the shape matrix \mathbf{Q}_s (a matrix of sorted eigenvectors learned from the training data, usually produced by PCA decomposition) to the mean shape vector $\bar{\mathbf{x}}$ Eq. (8). The same procedure is used to generate new texture vectors, where \mathbf{Q}_g is the texture matrix and \mathbf{t} texture parameters.

$$\begin{aligned} \mathbf{x} &= S(\mathbf{s}) = \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{s} \\ \mathbf{g} &= T(\mathbf{t}) = \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{t} \end{aligned} \quad (8)$$

The second stage of the algorithm is based on fitting a model for each independently used A-scan (depth-scan) to further improve the accuracy. This stage starts from the position defined by the result of the first stage B-scan fitting. The A-scan model is trained on offsets produced by back projecting the manual segmentation data using the main B-scan model and computing the boundary offsets $aOff$ between the back projections and the original segmentations Eq. (9) (n is the number of layers and u is the number of A-scans from all the images in the given segment). The improvement after the A-scan optimization can be seen in Fig. 5.

$$\mathbf{A} = \begin{pmatrix} aOff_{11} & \dots & aOff_{1n} \\ \vdots & \ddots & \vdots \\ aOff_{u1} & \dots & aOff_{un} \end{pmatrix} \quad (9)$$

2.4. The Model applied to the choroidal segmentation

Choroidal segmentation is an ideal application for the proposed statistical model, as the choroid is notoriously difficult to detect using standard approaches that have been applied so far to the foveal segmentation. There are three main reasons for that:

- the choroid is usually deep underneath the retina and thus OCT signals can degrade significantly, especially if the choroid is thick due to multiple scattering
- the CSI, its lower boundary is, unlike that of foveal layers, often very weak or in some locations invisible
- the texture is inconsistent, comprising large and small vessels

Essentially the choroid is more defined by its various vessel features as a textured region, rather than an area clearly separated by a strong boundary. All these factors make it an ideal application for a statistical model. The statistical model already assumes certain shape information and is thus far likely to perform poorly due to, at segments missing boundary and weak signal. Secondly, it considers the area bounded by the lower boundary being optimized, unlike many other segmentation methods that only try to track or in other ways segment the boundary itself without any knowledge about the area that that boundary defines.

The statistical model relies on initial boundary algorithm that defines the ILM and RBC and provides the offsets for the non-signal area of an image. The task for the model is then to find CSI boundary (the layer between the RBC and CSI). ILM is not used by the model, but is clinically useful to compare produced choroid thickness maps to the foveal shape.

The pre-processing and statistical model construction were done similarly to the procedure described in [27], thus only algorithm elements that differ will be described in more detail.

There are two distinctions. The first one is that the stacks are not registered by finding the position of the foveal pit, as it would be too unreliable to determine the exact location utilizing the same method as for the normal retinal segmentation. Dealing with pathologies means that there is no foveal pit, that part can be a bulge (i.e. AMD). Thus the stacks are simply used as they are, which does not cause any problems since the input to the shape training model is simpler, consisting of only one boundary (RBC), unlike with the foveal segmentation which had many more. The second difference is that the texture information is not included in the statistical model; texture is however used in form of a blob detector, but how that information is used by the objective function is fixed.

Given the RBC position the model is fitted to the image area below the RBC. A simple objective function is given: it tries to optimize the ratio of area covered by blobs (choroid being an area containing many vessels) versus the total area. Blobs are computed using a maximally stable extremal regions method (MSER [31]). An extremal region Q is a region for which all the elements are either larger (maximum intensity region) or smaller (minimum intensity region) than the neighboring elements (pixels).

The concept more simply can be explained by thresholding. All the pixels below a given threshold can be considered black and all those above or equal white. In a progressive sequence of thresholded images the first one would be completely white, then black spots corresponding to local intensity minima would appear, then grow larger and eventually merge, until the whole image would be black. The set of all connected components in the sequence is the set of all extremal regions. A formal definition of maximally stable extremal region is given below:

Let $Q_1, \dots, Q_{i-1}, Q_i, \dots$ be a sequence of nested extremal regions $Q_i \subset Q_{i+1}$. Extremal region Q_i is maximally stable if and only if $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$ has a local minimum at i . Δ is a parameter of the method. Intuitively it can be understood as a region that stays stable through a relatively large number of thresholds, i.e., it is a blob with a strong boundary, making it suitable for vessel cross-section detection.

Over a large range of thresholds, the local binarization is stable in certain regions, and has the properties listed below:

- Invariance to affine transformation of image intensities
- Covariance to adjacency preserving (continuous) transformation on the image domain
- Stability, as only regions whose support is nearly the same over a range of thresholds are selected.
- Multi-scale detection without any smoothing involved, both fine and large structure is detected

For this application the last two properties are important. The MSER implementation used is part of VLFeat, cross-platform open source collection of vision algorithms [32]. It bundles a MATLAB toolbox, a portable C library and a number of command line utilities. Several parameters that the MSER function takes will be explained:

- Delta: already mentioned in the MSER definition
- MaxVariation: defines the maximum variation (absolute stability score) of the regions
- MinDiversity: the minimum diversity of the region. When the relative area variation of two nested regions is below this threshold, then only the most stable one is selected
- MinArea and MaxArea: these parameters define the minimum and maximum allowed areas of detected regions as percentage of the total area of an image

For this application the parameters were heuristically chosen as follows: 1, 0.25, 0.7, 10/imageArea, 500/imageArea. Given the RBC boundary and a MSER filtered image $F_{MSER}(\mathbf{I})$, the optimizer, using the objective function, looks for the minimal value by changing the shape of the lower boundary. The objective function is defined as

$$f(\mathbf{s}) = -\frac{G_{CH}(S(\mathbf{s}), F_{MSE}(\mathbf{I}))}{\mathbf{A}_{CH} + \mathbf{A}_{PCH}} \quad (10)$$

It maximizes the ratio of the choroidal area covered in blobs against the total area of the choroid (A_{CH}) and the post choroidal region (A_{PCH}), taken to be an area from the choroid boundary plus 50 pixels, thus it is a constant. For the model initialization we have simply used the mean of the model which, having a robust objective function, proved to be sufficient. The A-scan optimization seeks black to white transition, given the fact that the choroid is usually darker than the post-choroidal region (though this boundary is often very weak or non-existent):

$$f_A(\mathbf{s}_A) = G_{DW}(S(\mathbf{s}_A), \mathbf{I}) - G_{AW}(S(\mathbf{s}_A), \mathbf{I}) \quad (11)$$

G_{DW} gives the average intensity of the choroidal region, CSI boundary minus 50 pixels, while G_{AW} of the post choroidal region, CSI plus 50 pixels. The final boundary is then filtered using a 1D median filter of size 11 pixels, with the size set heuristically.

3. Results and discussion

For evaluation, automatic segmentation results were compared to manual segmentation performed by an experienced examiner. A basic error measure was used, computed for a boundary i separately (in this case there is only one layer and two boundaries: RBC and CSI) and from these we compute error measures for an entire B-scan or for an individual layer, Eq. (12). E_B (Basic) is the basic error measure that defines the number of misclassified pixels.

$$^i E_B = \sum_{j=1}^{j=w} |yAut_{ij} - yRef_{ij}| \quad (12)$$

To compute error for layer k separately, only the two boundaries that define a layer are added (offsets) and divided by the sum of the layer area as given by the automatic segmentation (A_A) and the layer area as given by the reference segmentation (A_R), Eq. (13). This is used, instead of summing up across all boundary errors, to normalize for double counting of misclassified pixels, as each layer is bounded by two boundaries.

$$E = \frac{\sum_{i=k}^{i=k+1} ^i E}{A_A + A_R}, \quad 0 < k < b \quad (13)$$

For clinical application, it would be useful for the operator to decide whether the obtained results are reliable. As part of future work, a confidence measure could be introduced, based on the values returned by the objective function after the optimization step. Large values are proportional to the low confidence in the boundary positions determined by the model fitting.

In Table 1 results are presented for the seven diabetes type 1 eyes. In general, diabetes type 1 choroid is close to the choroidal shape and structure of healthy eyes. Error rates are even lower than what we have obtained with retinal segmentation [27]. In Fig. 5 a typical B-scan is shown from a diabetes type 1 patient, before and after the A-scan optimization. Notice the good performance even in the central region where the signal is weak.

Table 1. Error values on seven eyes with diabetes type 1, expressed as percentage of misclassified pixel relative to the layer area. The total average value is computed as an average of all the patient error values, and is independent of the number of B-scans per stack.

Error Type	A (100 B-scans)	B (100 B-scans)	C (100 B-scans)	D (98 B-scans)	E (100 B-scans)	F (93 B-scans)	G (100 B-scans)	Total
E_B	14.1	9.2	19.4	15.8	13.4	15.7	16.1	14.8

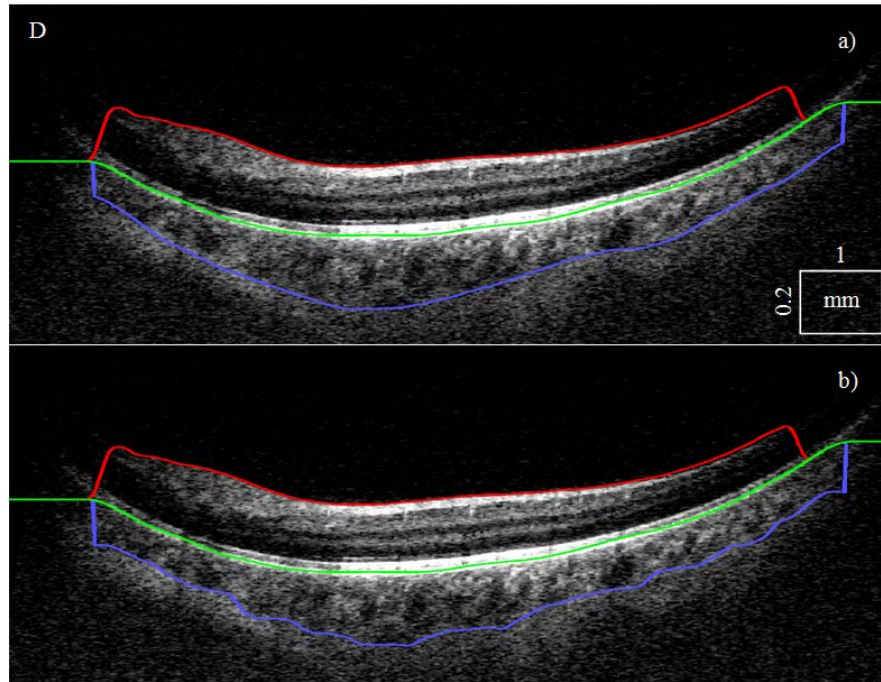


Fig. 5. A slice from a typical diabetes type 1 eye (patient D), before (a) and after (b) the A-scan optimization. ILM (red), RBC (green), CSI (purple).

In Table 2, five different pathologies are evaluated. The results are, surprisingly, even better than, close to normal, diabetes type 1 eyes (Table 1), especially considering the difficulty of segmenting such varied cases. One likely reason for that is the fact that the five pathologies were pre-processed slightly differently, with an average filtering of neighboring B-scans performed after the registration (to remove speckle). Another reason is that the manual segmentation was performed in the B-scan range of 20 to 100 (every fourth, 20 B-scans in total), and not 15 to 114 (diabetes type 1 patients, 100 B-scans in total), thus the error rates that are larger in slices (B-scans) close to the start and end of stack, where the signal is lower, contributed to the larger error value of diabetes type 1 data segmentations.

Table 2. Error values on 5 eyes with various pathologies, about 20 slices manually segmented per stack, expressed as percentage of misclassified pixel relative to the layer area. The total average value is computed as an average of all the patient error values, and is independent of the number of B-scans per stack.

Error Type	H (20 B-scans)	I (20 B-scans)	J (100 B-scans)	K (20 B-scans)	L (20 B-scans)	Total
E_B	5.6	12.8	10.1	13.0	15.3	11.3

Case H is an eye with a central serous chorioretinopathy. A lesion is present that affects the outer retina above the RBC complex. A deep, difficult to delineate, choroid is another characteristic of the pathology. This eye does not present a serious challenge for RBC segmentation (only slight RPE thinning in the middle), but determining CSI is difficult as the signal becomes very weak due to the thickness of the choroid. In Fig. 6 a representative B-scan is shown after the final A-scan optimization.

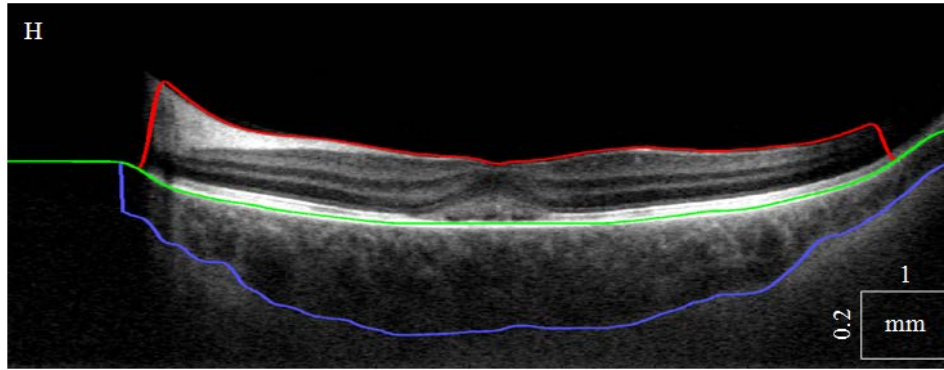


Fig. 6. The algorithm performs well despite the thinning of the RPE in the middle and deep, low signal choroid.

Case I is a terminal wet age related macular degeneration (AMD). The RBC complex has an irregular shape and thickness; however the dynamic programming approach we have introduced performs well. In Fig. 7 a representative B-scan is shown after the final A-scan optimization.

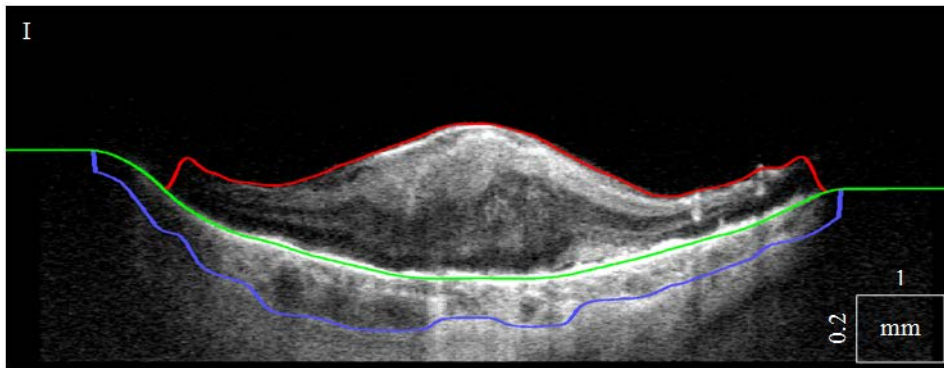


Fig. 7. The choroid is very thin and has a strong variation in contrast.

Case J is an eye with RPE atrophy as a result of dry age related AMD with a strong signal in the choroid due to reduced melanin concentration in the RPE. The choroid is extremely thin. The bright area in the choroid presents a challenge for RBC detection as well as for the

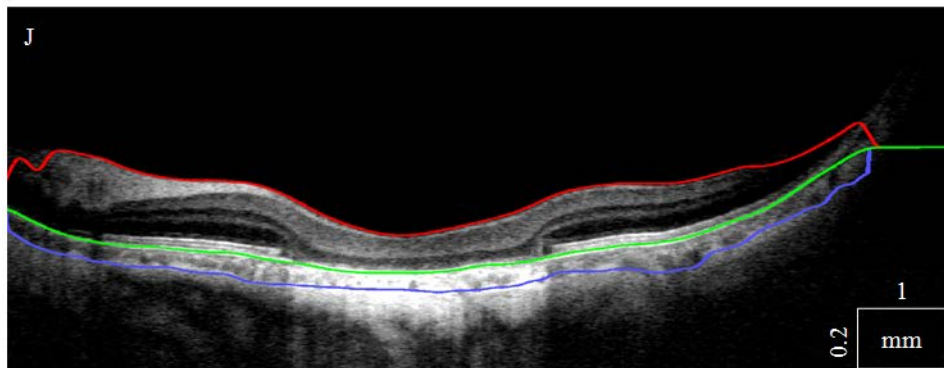


Fig. 8. RPE/RBC and CSI are found despite the extremely thin choroid with a strong signal in the middle as a result of the RPE atrophy.

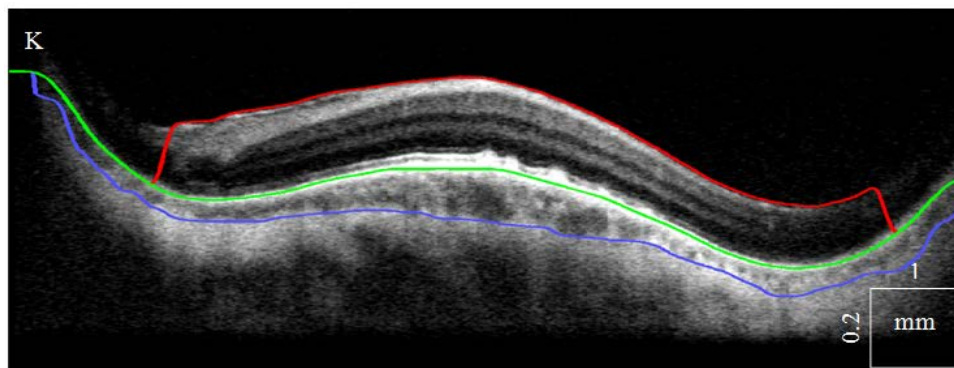


Fig. 9. RPE and RBC have an uncommon “inverted” shape.

statistical model (determining CSI), as the choroidal structure is significantly different from a normal choroid. However, as we can observe, the model is general enough to perform well even in that case. In Fig. 8 a representative B-scan is shown after the final A-scan optimization.

Case K shows an eye with deposits underneath the RPE (drusen). The RBC complex has an unusual inverted shape. A simple parametric approach for representing RBC, such as low order polynomial, would fail in this case. The proposed algorithm still produces reliable results. In Fig. 9 a representative B-scan is shown after the final A-scan optimization.

Case L is another AMD with a very difficult RBC to delineate. In Fig. 10 a representative B-scan is shown after the final A-scan optimization.

In our foveal segmentation paper [27] we collected large, manually segmented data sets via Amazon Mechanical Turk (AMT) service. Since the segmentation of choroid, and especially pathologies, required too much expertise to use AMT, there is no inter-observer variability measure. By comparing the obtained error rates to the foveal inter-observer variability it became obvious that the error rates are within that range. The average of the total error values from Table 1 and Table 2 is 13%.

In Fig. 11 six thickness maps are presented. For brevity, we have chosen to show only one diabetes type 1 eye as the other diabetes type 1 eyes are rather similar. However, we have included all 5 pathologies. Due to almost non-existing signal at the start and end of the stacks (volumes) sometimes a few of the first and last B-scans were left out.

Our model based approach uses the variation obtained from the training set and imposes those constraints when segmenting an unseen image. This ensures that the segmentation will

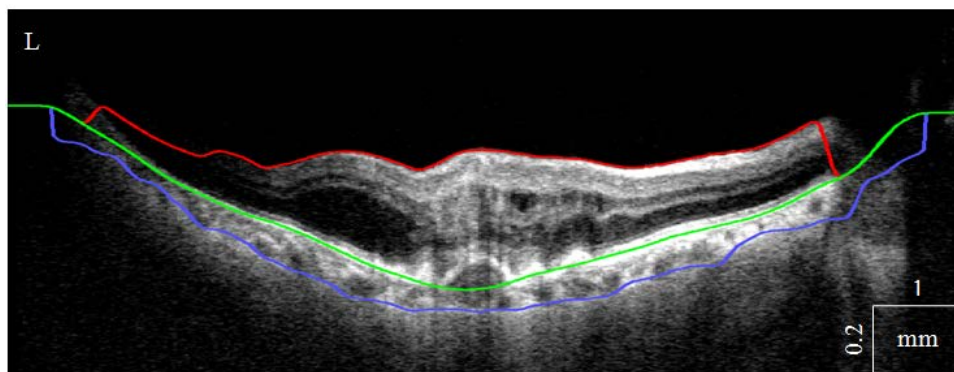


Fig. 10. The algorithm correctly interpolated over the drusen, finding a close approximation of the RBC.

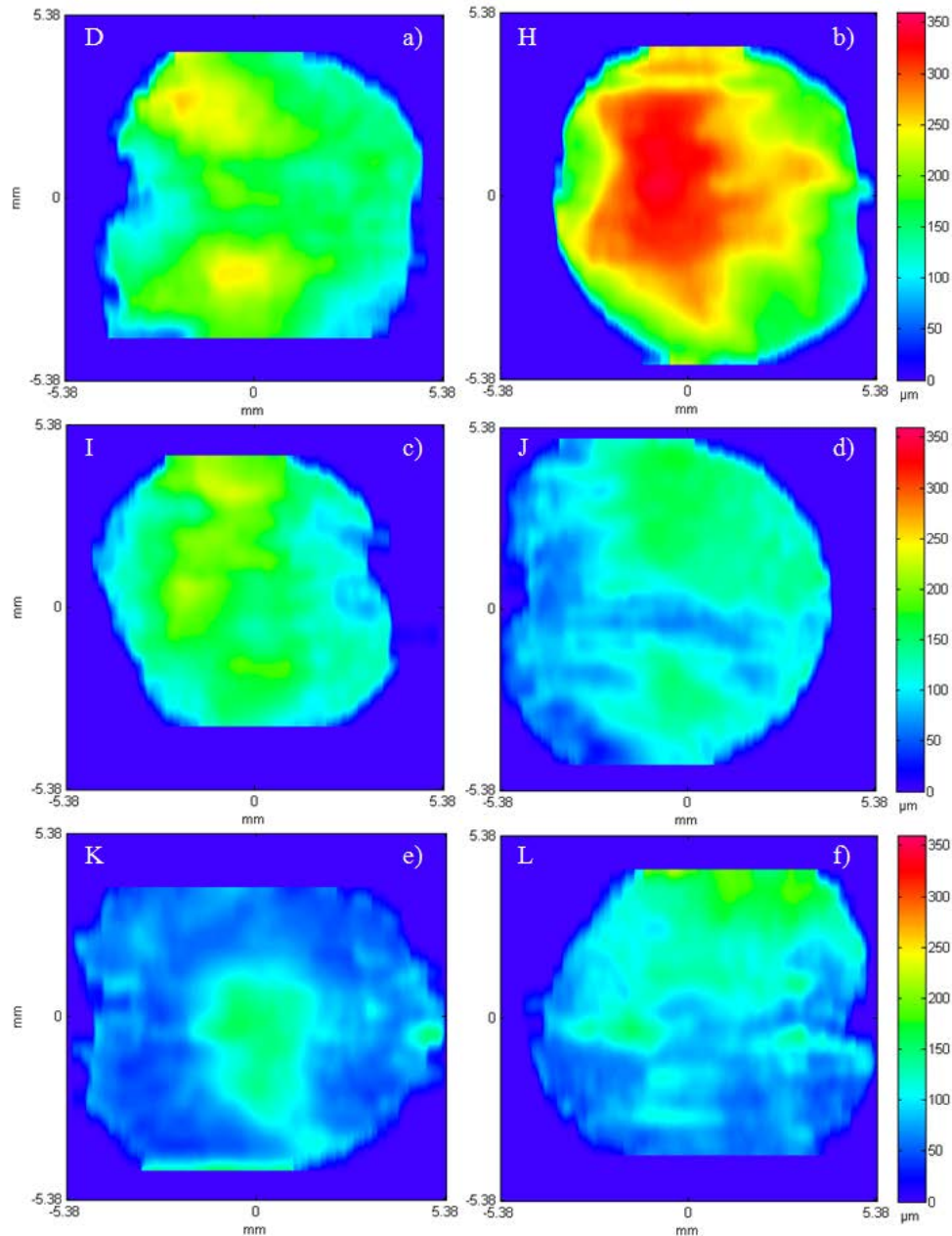


Fig. 11. Choroidal Thickness maps of a diabetes type 1 patient (D, subfigure a) and 5 pathologies (H-L, subfigures b to f).

likely be close to the ground truth and less sensitive to noise. However, it is extremely important to have a large, representative training set that includes all possible variation. Overall the novel algorithm segments choroid accurately as compared to the human expert segmentation, with the error rate 13%. In case that a precise and robust stack registration algorithm is used in the pre-processing stage, our algorithm could be implemented in full 3D and would be even more robust and accurate.

The current processing time is about 30 seconds per slice, using non-optimized MATLAB implementation. For our application, close to real time processing is not essential as the

results obtained are used for subsequent scientific analysis to obtain trends in various choroidal conditions. For clinical use, however, using optimized C implementation and a trivially parallelized approach, this could be brought down to approximately same time for the whole stack, using commodity hardware such as a GPU card.

4. Conclusion

We have proposed an algorithm for automatic segmentation of the choroid based on a statistical model. Before applying the statistical model, the top choroidal boundary (RBC) has to be determined, as it used to initialize the model. A novel method was developed for that purpose according to high robustness requirements in order to work well in cases of severe pathologies and signal degradation. It is a shortest path dynamic programming method that was extended to use the path derivative information and edge orientation with the same time complexity. It is also applicable to a general class of computer vision problems.

The choroidal automatic segmentation algorithm has been thoroughly tested and evaluated against a large, manually segmented data set obtained from a 1060nm OCT system and proved highly robust in wide-field scans even in the presence of artifacts and pathologies. It is the first time that a large, representative choroidal data set has been automatically segmented. We conclude that our algorithm successfully demonstrated reliable performance under difficult conditions. Clinically, fully automated segmentation of the choroid is essential in making medically useful the possibilities given by the method of high resolution, high speed OCT.

Acknowledgments

This research was supported in part by the BBSRC, Cardiff University; DTI grant (OMICRON); AMR grant (AP1110); Medical University Vienna, European Union project FUN OCT (FP7 HEALTH, contract no. 201880), FWF-NFN 'Photoacoustic imaging in biology and Medicine', OENB Jubiläumsfondsprojekt (14294), CARL ZEISS Meditec Inc., Femtolasers GmbH and the Christian Doppler Society (Christian Doppler Laboratory 'Laser development and their application in medicine').