Automated segmentation of intramacular layers in Fourier domain optical coherence tomography structural images from normal subjects

Xusheng Zhang
Siavash Yousefi
Lin An
Ruikang K. Wang
Automated segmentation of intramacular layers in Fourier domain optical coherence tomography structural images from normal subjects

Xusheng Zhang, Siavash Yousefi, Lin An, and Ruikang K. Wang

University of Washington, Department of Bioengineering, Seattle, Washington 98195
Beijing Institute of Technology, School of Optoelectronics, Beijing 100081, China
University of Washington, Department of Electrical Engineering, Seattle, Washington 98195

Abstract. Segmentation of optical coherence tomography (OCT) cross-sectional structural images is important for assisting ophthalmologists in clinical decision making in terms of both diagnosis and treatment. We present an automatic approach for segmenting intramacular layers in Fourier domain optical coherence tomography (FD-OCT) images using a searching strategy based on locally weighted gradient extrema, coupled with an error-removing technique based on statistical error estimation. A two-step denoising preprocess in different directions is also employed to suppress random speckle noise while preserving the layer boundary as intact as possible. The algorithms are tested on the FD-OCT volume images obtained from four normal subjects, which successfully identify the boundaries of seven physiological layers, consistent with the results based on manual determination of macular OCT images.

Keywords: optical coherence tomography; segmentation; macular layers; weighted gradient; statistical error removing; normal subjects.

1 Introduction

Optical coherence tomography (OCT), which was first introduced in the early 1990s, is a noninvasive, label-free, and high-resolution imaging modality that can provide depth-resolved tomograms of biological tissue morphology in vivo. The recent extension of OCT into Fourier domain, i.e., Fourier domain optical coherence tomography (FD-OCT), has accelerated its pace in biomedical imaging applications, particularly in ophthalmology. OCT is now being widely used in ophthalmic research and has proven clinically useful for diagnosing a variety of retinal diseases.

The morphology and determination of the thicknesses of retinal nerve fiber and ganglion cell layers (NFL and GCL) using OCT has been used to diagnose glaucoma. The morphological structures and the junction boundary of the inner/outer segment (IS/OS) layers have been reported to indicate the diagnosis of retinitis pigmentosa. The quantitative assessment of retinal pigment epithelium (PRE) in OCT tomograms is also useful in diagnosing some age-related retinal diseases such as macular drusen, polyoidal choroidal vasculopathy (PCV), and choroidal neovascularization (CNV). Therefore, automated extraction of useful information from numerous OCT retina tomograms is clinically valuable. Segmentation, in which different intraretinal layers are identified and separated from each other, is a critical task for reliable quantitative evaluation of retina structures.

Fully automated segmentation of intraretinal layers in OCT is challenging because many factors affect the sharpness of layer boundaries and homogeneity of each layer. For example, the non-Gaussian multiplicative neighborhood-correlated speckle noise, superimposed on structural images, reduces the image contrast near the layer boundaries. Because of the light absorption in red blood cells, relatively large retinal blood vessels often cast shadows below them that cause discontinuities in the OCT structural images. The motion artifacts due to involuntary head and eye movement are also inevitable during patient examination. Also, the diseased retinal structures are likely to vary among different patients.

Several methods have been proposed in the literature for intraretinal layer segmentation of OCT structural images. Koozekanani et al. proposed an automated algorithm using edge detection kernel and Markov model to segment retinal boundaries, from which the thicknesses of retinal layers were then obtained. Subsequently, other methods were reported based on detection of the reflectivity changes in retinal substructures using intensity and gradient of the microstructures or using a deformable spline algorithm. The goal of retinal segmentation is to accurately detect the edges and boundaries between retinal substructures. Edge detection is a classical problem in the field of image processing that is based on evaluation of image intensity and its derivatives (gradient and Laplacian). However, edge-detection-based segmentation suffers from image flaws, such as speckle noise, intensity discontinuities, and fusion or disruption of layer boundaries. Consequently, more robust complicated techniques have been proposed to improve the effectiveness of retinal segmentation. Fernández et al. utilized a sophisticated denoising solution that applied nonlinear complex diffusion filtering along with coherence-enhancing diffusion filtering techniques, and then located slope peaks on the enhanced images. However, their algorithm was computationally expensive and pathology-dependent, and required modifications in practice. Mishra et al. proposed a two-step kernel-based optimization scheme that first located...
the approximate location of retinal layers using a gradient-based adaptive vector-valued kernel function and then optimized the segmentation using dynamic programming-based force balance equation to identify noisy layer boundaries. But their method was only tested on rodent models and not in human retina. Mayer et al.\textsuperscript{26} proposed a segmentation algorithm based on the minimization of an energy function consisting of gradient and local smoothing terms. However, the iterative process to minimize the energy depends on the accurate choice of initial values that have to be determined by a heuristic rule. Vermeer et al.\textsuperscript{27,28} presented a support vector machine pixel-classification algorithm using the intensity and gradient information of each pixel for layers’ identity and a level set method to smooth the boundaries. However, their algorithm required training on manually segmented data, which can be subjective. The same problem holds for the work of Kajić et al.,\textsuperscript{29} who employed a statistical model based on texture and shape. Although Kajić’s method was described to be highly robust to speckle noise, the prior knowledge on segmentation was important and the results were highly influenced by the pathology variations. All of the aforementioned methods have varying levels of success and are either computationally expensive or subjective.

In this paper, we demonstrate an automatic segmentation of intraretinal layers in the macular region of healthy human subjects. This method applies a two-step predenoising filtering procedure within both B-frame images and en-face planes, after alignment of A-lines. After the denoising procedure, a weighted gradient segmentation is utilized to initially identify the boundaries, and then least-square regression with a given statistic confidence level is employed to eliminate the gross erroneous points and smooth the detected retinal layer boundaries. Compared to the existing methods in the literature, our technique is computationally efficient and does not depend on training data sets and prior knowledge from sample datasets.

2 Experimental Setup

To determine the accuracy and repeatability of the proposed algorithm, 30 three-dimensional (3-D) FD-OCT images from the macular region of the right eyes of four healthy volunteers were acquired. The OCT imaging on normal subjects was approved by the Institutional Review Board at the University of Washington. Figure 1 shows the schematic system setup used in this study, which is similar to that described previously.\textsuperscript{30,31} Briefly, the system is capable of providing both structural and blood flow perfusion images within retina. It used a superluminescent diode with a central wavelength of 842 nm and a full width at half maximum (FWHM) bandwidth of 46 nm, delivering an axial resolution of \(\sim 7 \mu m\) in air. In the sample arm of the interferometer, an objective lens with 50 mm focal length was used to achieve \(\sim 16 \mu m\) lateral resolution. In the imaging, each B-scan was formed by acquiring 500 A-scans with a spacing of \(\sim 6 \mu m\) between adjacent A-scans, which covered a total lateral scan range of \(\sim 3 \text{mm}\). To suppress the speckle noise, five B-scans were acquired at the same position and averaged to form one final B-frame. In the \(y\)-direction, 200 B-frames over \(\sim 3.0 \text{mm}\) on the tissue were captured, meaning a spacing of \(\sim 15 \mu m\) between adjacent B-frames.

3 Methodology

The schematic diagram of segmentation steps is shown in Fig. 2 where the abbreviations are: ILM—inner limiting membrane, RPE—retinal pigment epithelium, ONL—outer nuclear layer, IS/OS—inner/outer segment, NFL—nerve fiber layer, GCL—ganglion cell layer, IPL—inner plexiform layer, INL—inner nuclear layer, OPL—outer plexiform layer.

3.1 RPE Layer Boundary Detection

In cross-sectional structural images of retina, the vitreous body and choroid occupy a large space that is not necessary for segmentation. To reduce the processing time and limit the search space for the layers’ boundaries, the retinal borderlines, i.e., the anterior and posterior borders, were first identified. The identification of retinal borderlines is straightforward due to the high intensity contrast among vitreous, retina, and choroid. In order to detect the retinal borderlines, first the B-frame images were calibrated by removing a background intensity level and then blurred heavily by 7-by-7 median filter followed by Gaussian
filter of size 15 pixels and standard deviation of 2. This filtering operation is considered to have less impact on large gradient-based retinal borderlines detection. In general, the RPE complex boundary exhibits the largest change in refractive index in every A-lines, therefore by identifying the maximum values in each A-lines, the RPE boundary layer can be estimated. However, due to the interference of noise, shadows below the large vessels, relatively high scattering intensity in NFL, and other uncertainties, the determined points are not always representing the RPE boundary. Fabritius et al. presented a method to identify the erroneous pixels by applying an automatic binarization algorithm following a top-hat filtering operation. However, their method is not always effective, especially when the retinal images are significantly inclined and no prior knowledge of morphological filtering shape parameter is known.

### 3.1.1 Iterative statistical regression method

In order to correct the estimated RPE boundary, we performed a statistical regression method to eliminate the grossly erroneous identifications. After finding the points with maximum intensities in every A-lines and estimating the RPE boundary, a polynomial along the transverse coordinate was fitted to the points. In healthy human retina, the RPE layer generally has a small curvature, thus a second or a little higher order(cubic) polynomial is sufficient to estimate that layer. The fitted quadratic polynomial is used in this paper to estimate the RPE boundary layer. Then, a confidence interval with 92% confidence level was defined around the RPE estimate. The confidence interval was defined on the estimation error between the actual points and the fitted curve. The upper and lower boundaries of the interval were the points in the normalized error histogram (probability distribution function), in which the area under the histogram between these intervals was equal to the confidence level. In other words, if $E(x)$ is the estimation points, the confidence interval $[L(x), U(x)]$ can be defined such that

$$P[L(x) < E(x) < U(x)] = \text{confidence level},$$

where the $P[\cdot]$ represents probability. The confidence level was derived experimentally such that most of erroneous values would fall outside the interval. Then, the points outside the interval were removed and another polynomial was fitted to the points with maximum intensities in the confidence interval. This process was repeated until there would be no change in the polynomial. The remained points with maximum intensity values indicated the RPE boundary.

Figure 3 shows a typical procedure for estimating the RPE where the points with maximal intensities in every A-lines are identified. The arrows point to the erroneous locations of the RPE. The solid smooth line is the second-order polynomial fit to the identified points and the two dashed lines indicate the upper and lower boundaries of the confidence interval (i.e., confident limits). Then, the identified points with maximum intensities between the dashed lines were accepted and those outside were removed and corrected. The process was repeated until no significant change was observed on the confidence interval [Fig. 3(b) through 3(d)]. The corrected points in Fig. 3(d) indicate the RPE. Note that the least-squares polynomial regression method with statistical estimation analysis would also be applicable for detecting some other layered boundaries discussed in this paper.

### 3.2 Vitreous/ILM and RPE/Choroid Boundary Detection

Once the RPE layer was determined, the vitreous/ILM and RPE/choroid boundaries can be easily identified. The vitreous/ILM boundary was defined as the points with the greatest rise in contrast in the region above the RPE layer, and the RPE/choroid boundary was defined as the points with greatest contrast drop in the region beneath the RPE layer. A sixth-order polynomial was fitted to smooth the boundary lines after the erroneous data being removed by statistical regression method. Figure 4 demonstrates the typical detection of the retinal borderlines near the fovea.

At this stage, all A-scans in the cross-sectional images were aligned to make the RPE/choroid boundary to form a straight line. By aligning all RPE/choroid boundaries, the retinal natural motion along the depth caused by microsaccade could be effectively removed so that filtering in the en-face plane (perpendicular to the A-line direction) could be performed. The regions corresponding to vitreous and choroid were also trimmed off and removed from the images.

### 3.3 Intramacular Layers Segmentation

#### 3.3.1 Denoising intramacular layers

The intramacular layers are closely spaced and the intensity contrast in between these layers is relatively low, which makes the

![Fig. 3 Iterative RPE boundary detection. (a) Error points (shown by arrows). Then, a low-order polynomial is fit (solid smooth line) and a confidence interval with 92% of confidence level (two dashed lines) is defined. The estimated points inside the interval are accepted. (b) to (d) The process is repeated until the interval has no significant variation.](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics)
segmentation and denoising challenging. Because of the textural properties of these layers, anisotropic diffusion filters have been used to denoise the images before segmentation\textsuperscript{20,23,26} while preserving the edges of layer boundaries. However, these filters are computationally expensive.

We performed a denoising procedure before segmenting intramacular layers. First, a 9-by-9 Gaussian filter with standard deviation of 2 was performed along the en-face plane (x-y plane). Note that the window size should be moderately selected in which no significant curvature could be found with the retinal layers, and the standard deviation was experimentally determined according to the degree for blurring. Then, a 3-by-7 Gaussian filter with standard deviation of 1.6 was implemented for each B-frame (y-z plane). Herein we used such filtering parameters as to doing lighter blurring in the depth direction than the slow (en-face) direction. In addition to its computational efficiency, this bidirectional filtering procedure is efficient to suppress the noise and preserve boundary edges. The performance of the proposed filtering technique in suppressing noise is illustrated in Fig. 5 and compared with the conventional anisotropic diffusion\textsuperscript{32}

\[
\frac{\partial I}{\partial t} = \text{div}[c(x,y,t)\nabla I] = c(x,y,t)\Delta I + \nabla c \cdot \nabla I, \quad (2)
\]

where \(\Delta, \nabla, \text{div}(\ldots)\) denote the Laplacian, gradient, and divergence operators, respectively. \(c(x,y,t)\) means the diffusion function that controls the rate of diffusion. The diffusion function \(c(x,y,t)\) depends on the magnitude of the gradient of the image gray and it should be a monotonically decreasing function

\[
c(\nabla I) = \frac{1}{1 + (\|\nabla I\|/K)^{2}}, \quad (3)
\]

where the constant \(K\) controls the sensitivity to edges and is usually chosen experimentally by manually or histogram-based noise estimator. The anisotropic diffusion was implemented in 20 iterations with a sensitivity constant \(K = 30\). Figure 5(a) shows a raw B-frame of the retina. Figure 5(b) and 5(c) show the denoised results using the anisotropic diffusion and the proposed method on that B-frame, respectively. Also, Fig. 5(d) and 5(e) show the intensity profiles of two selected A-lines from the B-frame. It can be seen that the performance of the proposed method is comparable to that of the anisotropic diffusion in suppressing the noise for segmentation of retinal layers.

### 3.3.2 Segmentation of ONL and IS/OS/RPE

After denoising, the ONL layer was identified first as the depth location with minimal backscattering intensity in retina for each A-line so that the macular could be divided into two different regions, as the second-top line illustrated in Fig. 6. Thereafter, IS/OS and OS/RPE boundaries were detected by the maximal contrast rise and drop, respectively, with the search area being limited between the ONL layer and RPE layer. Since ONL, ELM, and IS layers are located very close to each other and their absolute and relative backscattering intensities are low, their boundaries cannot be identified with our current system. Figure 6 shows the segmentation of ONL and IS/OS/RPE layers.

### 3.3.3 Segmentation of the layers between NFL and ONL

Detecting boundaries between NFL and ONL layers is very challenging because of the blood vessels and their shadows and the variations of their thickness at different locations. In the foveal region, the layers would merge together and become increasingly indistinguishable. Therefore, simple edge detector algorithms, such as Prewitt, Laplacian of Gaussian, Canny method,\textsuperscript{33} often fail to effectively detect the boundaries, and
thus are not practical. Figure 7(a) shows the raw results with many local errors from the gradient extrema detection algorithm, where the boundaries of NFL/GCL, IPL/INL, INL/OPL are detected as the depth positions of two significant intensity drops and the greatest intensity rise, respectively. Erroneous identification of boundary occurs due to the influence of vessels and noise, and the absence of clear layer structure near the fovea region. The top and bottom borderline of macula, and the ONL layer line are smoothed by polynomial fitting. It should be noted that in the images from our OCT system, the boundary between GCL and IPL is almost imperceptible even to manual inspection, so it will not be detected in our algorithm.

In healthy human retina, the physical surfaces of the retinal layers are relatively smooth and locally continuous. A good method should make use of some neighborhood 3-D information to guide the segmentation. Based on this consideration, we propose an algorithm that evaluates the weighted gradient extrema along each A-line within limited search region with neighborhood-based optimization strategy employed to avoid potential erroneous determination.

\[
    z(i,j) = \arg\min_{z_1, z_2, \ldots, z_K} \left\{ w_j(\delta_{i,j,k}) \cdot \nabla I(z_{i,j,k}) \right\}, \tag{4}
\]

where each local extreme of intensity gradient \( \nabla I(z) \) along the \((i,j)\)th depth scan line is imposed by an unimodal weight function \( w(z) \), and \( z_{i,j,k} \) represents the \( K \) locations in depth with maximal or minimal intensity gradients. The subscript \( ij \) in \( w_j(z) \) means that the weight function centroids are variable against lateral scanning positions \((i,j)\). We used a normalized Gaussian function with standard deviation of 5 pixels coupled with a sign function as the weight function, which generated a minus weight coefficient for maximal gradient and a plus one for minimal gradient. The standard deviation of Gaussian weight function is determined from expert analysis with the consideration of axial resolution. The boundaries were searched one by one within the limited region between the top NFL and ONL layer. When performing search on each boundary, the algorithm sorted the weighted gradient extremes in each A-line and considered the position in depth with minimal weighted value to be the boundary position in this A-line. Considering the smoothness and no abrupt spike of retinal layer boundary, the centroid of the unimodal weight function imposed on the later A-line comes from the mean depth position of the corresponding boundary in the previous sub-region of the 5-by-5 pixels neighborhood. However, due to the lack of previous scanning information, the first three lines and frames have to depend on weightless gradient extremes. So they should be located a little far away from the fovea to avoid the absence of clear layer structures. Nevertheless there still existed a few cross-layer errors in the fovea subregion after performing the searching method. This can be simply corrected because there are no physical tissue layers that can cross each other in retina, so these errors only mean the merge of two layers. Figure 7(b) shows the segmentation outcome after correction according to this search strategy on boundaries. The search was limited to the region between the previously segmented vitreous/NFL line and ONL line.

4 Results and Discussion

4.1 Segmentation Results

The acquired macular structure movie was first decomposed to BMP frames by the program. The described method in Sec. 3 was implemented using MATLAB (The MathWorks, Inc.) M-file code. The program runs on a personal computer (64 bit OS, Intel Core i7 CPU at 2.7 GHz, and 8 GB RAM) and took about 2.5 min to complete the whole 3-D image volume (448 × 500 × 200) pixels detection of seven layer boundaries and two layer location lines. If carefully optimized using a more efficient language, for example C++ programming language, it can perform with dramatically reduced processing time.

The segmented results of seven layer boundaries (VITREOUS/NFL, NFL/GCL, IPL/INL, INL/OPL, IS/OS, OS/RPE, RPE/CHOROID) and two layer location lines (ONL, RPE) are illustrated in Fig. 8 for one of the typical 3-D datasets. Sixth-order polynomial fittings are applied to smooth the boundaries. It can be observed that the effect of blood vessels and their corresponding shadows below them is minimized [see Fig. 8(a)]. The fifth and eighth lines, approximately showing the positions of ONL and RPE, are used as the separating line.
lines in segmentation algorithm to limit the search areas. The GCL/IPL boundary is almost indetectable in the images from our OCT system due to the limited sensitivity, so we only get the GCL + IPL complex layer. This complex layer is just located between the boundaries of NFL/GCL and IPL/INL. Figure 9 shows the 3-D view of the segmentation.

To confirm the accuracy of our method, we applied the algorithm to the datasets from four volunteers in our lab. In addition, an independent physician separately reviewed the 3-D images, and manually segmented 32 B-frames randomly extracted from the datasets. The absolute mean and standard deviation of the differences between manual and automatic estimates were calculated and shown in Table 1. The frames with poor image contrasts or obvious motion artifacts were first discarded before segmentation. Typically about 1% of frames were eliminated in our experiments.

4.2 Discussion

We applied our algorithm to more than 30 3-D data volumes captured from four volunteers and obtained a promising success. The influence of the common blood vessels and their corresponding shadows on the final results is minimized. However, we found that if the blood vessels are relatively large, which are located exactly at the first several A-lines, this situation would usually result in a failure of the segmentation algorithm. This is due to the fact that our proposed algorithm requires some knowledge of the successful segmentation of previous neighborhood lines. However, for the first few A-lines there is no prior knowledge, but instead, the algorithm is solely dependent upon their own layer structural information. Also, the large eye motion may break the local 3-D continuousness and smoothness of layers in tomograms, which would affect the robustness of segmentation algorithm. Thus, careful measures should be exercised in order to reduce large motion artifacts.

It should be mentioned that the method of statistical error estimation based on polynomial fitting imposes a global regularity assumption on the geometry of the retinal layers. So, it is usually limited to the cases where the OCT images cover a small field of view (~3 mm in this study). Also, when the polynomial fitting is used to smooth the boundaries, the temporal raphe in the NFL would be smeared out, and some cross-layer errors occur in the fovea as well.

Although some spatial treatment has been applied in denoising and boundary segmentation in the present work, more 3-D information from neighboring A-lines and B-scans could further be utilized for guiding segmentation or correcting errors. Processing of layered structure image-set in a completely 3-D perspective will be a promising way for accurate and robust segmentation.

5 Conclusions

We have presented an automated segmentation method that is based on the combined use of the locally weighted gradient extrema search strategy with the statistic estimation polynomial regression technique. The locally weighted gradient extrema search strategy was designed to obtain the smooth and continuous boundaries with less abrupt jumpiness in the neighborhoods, and the probability-based polynomial regression method was used to selectively eliminate the potential gross errors in the detection of boundaries. In the segmentation, a two-step denoising treatment in both B-frame images and en-face planes was employed to remove random speckle noises while preserving the boundary details. The method has been applied to a reasonable number of macular scans acquired from four healthy subjects. The segmentation results were found to be consistent with those by manual segmentation. In addition, the thickness color maps of multiple layers were generated based on the segmentation results and the thickness measurements of different layers were also in agreement with those reported in the prior literatures. Future study will focus on testing the effectiveness of the proposed segmentation algorithm on the 3-D OCT scans from representative patients.

Acknowledgments

This work was supported in part by research grants from the National Institutes of Health (R01EB009682), and an unlimited fund from Research to Prevent Blindness (Research Innovation Award).

Table 1 Differences between manual and automatic segmentations.

<table>
<thead>
<tr>
<th>Layers</th>
<th>Mean difference (Pixels)</th>
<th>Standard deviation (Pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VITREOUS/ILM</td>
<td>−0.85</td>
<td>0.83</td>
</tr>
<tr>
<td>NFL/GCL</td>
<td>−0.77</td>
<td>1.30</td>
</tr>
<tr>
<td>IPL/INL</td>
<td>−0.09</td>
<td>1.06</td>
</tr>
<tr>
<td>INL/OPL</td>
<td>0.22</td>
<td>1.12</td>
</tr>
<tr>
<td>IS/OS</td>
<td>0.07</td>
<td>0.79</td>
</tr>
<tr>
<td>OS/RPE</td>
<td>−0.66</td>
<td>0.77</td>
</tr>
<tr>
<td>RPE/CHOROID</td>
<td>1.06</td>
<td>0.70</td>
</tr>
</tbody>
</table>

References


