Validated Automatic Segmentation of AMD Pathology Including Drusen and Geographic Atrophy in SD-OCT Images

Stephanie J. Chiu,¹ Joseph A. Izatt,¹,² Rachelle V. O’Connell,² Katrina P. Winter,² Cynthia A. Toth,¹,² and Sina Farsiu¹,²

PURPOSE. To automatically segment retinal spectral domain optical coherence tomography (SD-OCT) images of eyes with age-related macular degeneration (AMD) and various levels of image quality to advance the study of retinal pigment epithelium (RPE)+drusen complex (RPEDC) volume changes indicative of AMD progression.

METHODS. A general segmentation framework based on graph theory and dynamic programming was used to segment three retinal boundaries in SD-OCT images of eyes with drusen and geographic atrophy (GA). A validation study for eyes with nonneovascular AMD was conducted, forming subgroups based on scan quality and presence of GA. To test for accuracy, the layer thickness results from two certified graders were compared against automatic segmentation results for 220 B-scans across 20 patients. For reproducibility, automatic layer volumes were compared that were generated from 0° versus 90° scans in five volumes with drusen.

RESULTS. The mean differences in the measured thicknesses of the total retina and RPEDC layers were 4.2 ± 2.8 and 3.2 ± 2.6 μm for automatic versus manual segmentation. When the 0° and 90° datasets were compared, the mean differences in the calculated total retina and RPEDC volumes were 0.28% ± 0.28% and 1.60% ± 1.57%, respectively. The average segmentation time per image was 1.7 seconds automatically versus 3.5 minutes manually.

CONCLUSIONS. The automatic algorithm accurately and reproducibly segmented three retinal boundaries in images containing drusen and GA. This automatic approach can reduce time and labor costs and yield objective measurements that potentially reveal quantitative RPE changes in longitudinal clinical AMD studies. (ClinicalTrials.gov number, NCT00734487.) (Invest Ophtalmol Vis Sci. 2012;53:53–61) DOI:10.1167/iovs.11-7640

From the ¹Department of Biomedical Engineering, Duke University, Durham, North Carolina; and the ²Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina.

Supported in part by the American Health Assistance Foundation. The A2A SD-OCT Study was funded in part by Genentech Grant IST-44005, with clinical imaging equipment support from Biotoptigen and Alcon Laboratories.

Submitted for publication March 28, 2011; revised August 22, 2011; accepted October 1, 2011.

Disclosure: S.J. Chiu, P.; J.A. Izatt, P.; R.V. O’Connell, None; K.P. Winter, None; C.A. Toth, Alcon (C, F), Genentech (C, F), Biotoptigen (F), Physical Sciences Inc. (C), P.; S. Farsiu, P.

Corresponding author: Stephanie J. Chiu, Department of Biomedical Engineering, Duke University, Durham, NC 27708; stephanie.chiu@duke.edu.
the inner limiting membrane (ILM), inner aspect of the RPEDC, and outer aspect of Bruch’s membrane (Fig. 1). We then propose an algorithm that automatically segments these layer boundaries. In parallel with other graph theory-based algorithms, this algorithm is in part based on our previously proposed graph theory and dynamic programming technique used to segment eight retinal layer boundaries, which has been verified to be accurate and reliable in normal adult eyes.34 In this article, we also present the additional algorithmic steps that are required to apply this segmentation framework to eyes with nonneovascular AMD and subsequently validate it for accuracy and reproducibility.

**METHODS**

Extending our previous segmentation framework to AMD eyes required a) the establishment of guidelines for segmenting images with AMD pathology, b) metrics for image quality, c) adaptation of the graph theory and dynamic programming framework to handle images with drusen and GA, and d) assessment of the segmentation results through accuracy and reproducibility studies.

**Proposed Manual Segmentation Guidelines for AMD Pathology**

Before manual segmentation and algorithm development, we constructed a set of qualitative guidelines based on previous literature, expertise from the Duke OCT Reading Center, and representative images, to trace layer boundaries on images with nonneovascular AMD pathology. Guidelines and example images were used as a reference for manual segmentation to maintain a consistent and unbiased interpretation between certified graders. Practice sessions for manual segmentation were also performed on training data sets based on the guidelines. These guidelines are listed as follows:

1. We isolate the RPE and drusen complex (denoted RPEDC) by delineating the inner aspect of the RPE plus drusen material and the outer aspect of Bruch’s membrane.

Sarks et al. have shown progression in AMD by correlating basal linear and basal laminar deposits of the RPE to greater amounts of membranous debris associated with clinically evident drusen and pigmentary changes on color funduscopic measurements. More recently, Zweifel et al. have shown subretinal deposits in reticular drusen. Thus, in particular for macular SD-OCT datasets with nonneovascular AMD, we believe that a measure of the RPEDC volume containing all drusen material, whether above (Fig. 2B) or below the RPE (Fig. 2A), would be a more useful measure of disease. Such a metric, which includes the RPE and small deposits of drusen material rather than only large collections of debris, should therefore differentiate normal aging from pathologic AMD processes. This hypothesis will be tested in the longitudinal AREDS2 Ancillary Spectral Domain Optical Coherence Tomography (A2A SD-OCT) study with age-matched controls.

**FIGURE 1.** SD-OCT image of an eye with intermediate AMD and the target layers segmented. (A) An unsummed (raw), high-quality, foveal B-scan with a 6.50-μm lateral pixel resolution and a 3.24-μm axial pixel resolution. (B) Manual segmentation of the image in (A), delineating the inner aspect of the inner limiting membrane (ILM) in blue, inner aspect of the RPE + drusen complex (RPEDC) in green, and outer aspect of Bruch’s membrane in yellow. These boundaries isolate the total retina (blue to green) and the RPEDC (green to yellow). (C) Automatic segmentation of the image in (A). Comparison of the total retinal thickness in (B) versus (C) yielded a mean thickness error of 0.9 μm, maximum error of 22.7 μm, and 2% of A-scans with a >5-pixel difference. Respective values for the RPEDC are 2.2 μm, 19.4 μm, and 1%. In both (B) and (C), note the exclusion of the photoreceptor outer segments from the RPEDC layer and the near convergence of the green and yellow lines at the site of focal GA.

**FIGURE 2.** Example of features to include in the RPEDC from eyes with intermediate AMD. (A) Sub-RPE drusen (under the asterisks) and (B) a subretinal drusenoid deposit (under the asterisks), both of which are included in the RPEDC.

54 Chiu et al. IOVS, January 2012, Vol. 53, No. 1

Downloaded from iovs.arvojournals.org on 03/29/2019
Automatic Segmentation of SD-OCT Images with AMD

II. We include all hyperreflective material contiguous with the RPE as part of the RPEDC, excluding the following.
A. Material over a nearly absent RPE with a width narrower than the azimuthal pixel resolution (Fig. 3B).
B. Indistinguishable dim or shadowy features over a nearly absent RPE (Fig. 3C).

We include all forms of drusen, such as sub-RPE drusen (Fig. 2A) and subretinal drusenoid deposits (Fig. 2B), in the RPEDC due to the implications outlined in guideline I. While hyperreflective foci have been suggested to indicate disease progression, we chose not to include these foci as part of the RPEDC because they represent cells that have migrated away from (and are not contiguous with) the RPE. The inner border of the RPEDC was distinguished from the overlying hyperreflective IS-OS band when present, as demonstrated in Figure 1C.

We do not include narrow particulate (Fig. 3B) or dim material (Fig. 3C) over regions where the RPE is nearly absent (Fig. 3A) since they may represent residual drusen material or degenerated neurosensory cells. To determine whether the RPE is nearly absent, we qualitatively assess the thickness of the RPE and use hyperreflectivity in the underlying choroid as a supporting indicator of geographic atrophy (GA). For small, particulate material, we selected the minimum resolution to be equivalent to the azimuthal pixel resolution (distance between B-scans) to attain isotropic resolution, because in our experiments the azimuthal pixel resolution was lower than the lateral (distance between A-scans) and axial pixel resolutions (depth resolution). In this study, 67 μm was used as the minimum resolution.

Automatic Layer Segmentation Algorithm

We base our new three-retinal-layer boundary segmentation algorithm for SD-OCT images with AMD pathology on the generalized graph theory and dynamic programming framework that we previously introduced for normal retina. An outline of the new algorithm flow (Fig. 4) highlights the key components needed to adapt this method for images with drusen and GA, and an overview of the steps involved are described in the subsequent paragraphs.

![Flowchart diagram](https://via.placeholder.com/150)

**Image Downsampling.** To reduce the overall computation time, we first downsample the image by a factor two in both dimensions using bi-cubic interpolation and anti-aliasing. This step can be ignored for images with low resolution, or if the computational complexity is of no concern.

**NFL-OPL and IS-RPE Separation.** There are two distinct hyperreflective regions in a filtered SD-OCT image of the retina: the region bounded by the NFL and outer plexiform layer (denoted NFL-OPL complex) and the region containing the inner segment–outer segment junction (IS-OS), RPE, and drusen (denoted IS-RPE complex). For retinal images with AMD, the pathology may result in a merging of the NFL-OPL and IS-RPE complexes. If these two regions are not separated before segmenting, then it is possible for the ILM boundary and the inner boundary of the RPEDC to be mistaken for each other due to similarities in their characteristics.

We therefore generate a binary mask of the image to isolate the NFL-OPL and IS-RPE hyperreflective complexes, by smoothing the image with an 11-pixel Gaussian filter with a standard deviation of 11 pixels, extracting the edges with a [-1;1] high-pass filter (using MATLAB notation: `Tor: The MathWorks, Natick, MA), normalizing the image to range from 0 to 1, generating a binary mask using a threshold of 0.5 on the normalized image, opening any gaps in the clusters using a 3 × 3 pixel structuring element, removing connected clusters smaller than 200 pixels, and closing any remaining gaps using the same structuring element.

Once the mask is generated, we delineate the boundaries of the two white bands corresponding to the two NFL-OPL and OS-RPE complexes using graph theory and dynamic programming. We generate two vertical gradient adjacency matrices—a black-to-white and a white-to-black matrix—using the [-1;1] and [1;1] edge filters and set all negative values to 0. After automatic endpoint initialization, we segment the four boundaries in the image. We achieve this by twice searching for a black-to-white edge to locate the upper boundaries of the two white bands and twice searching for a white-to-black edge to locate the two lower boundaries of the white bands. To ensure the same edge is not cut again, we exclude already delineated nodes from the graph when cutting subsequent edges. The result is a pilot estimate of the ILM, inner RPEDC, and Bruch’s membrane boundaries.
Image Flattening. Next, we flatten the image based on the convex hull of the estimated inner RPEDC boundary, by shifting the columns of the image up or down until the estimated convex hull lies on a flat line. To prevent introducing any new border artifacts, columns of the flattened image without pixel intensity information are assigned intensity values corresponding to the mirror image of the values in the valid regions of the same column.

Calculating Graph Weights. We then create adjacency matrices based on the flattened image. To delineate the ILM, weights assigned to a dark-light adjacency matrix are calculated based on equation 1:

\[ w_{ab} = \text{normalize}(- (g_a^{DL} + g_b^{DL}), 0, 1) + w_{\text{norm}}. \]  

(1)

where \( w_{ab} \) is the edge weight connecting nodes \( a \) and \( b \). \( g_a^{DL} \) and \( g_b^{DL} \) are the vertical dark-to-light gradients of the image at nodes \( a \) and \( b \), respectively, and \( w_{\text{norm}} \) the minimum weight of the graph, is \( 1 \times 10^{-5} \). The normalize\((x, y, z)\) notation indicates a normalization of the values \( x \) to range from \( y \) to \( z \).

For Bruch’s membrane, a separate adjacency matrix is used with weights calculated based on equation 2.

\[ w_{ab} = \text{normalize}(- (g_a^{LD} + g_b^{LD}), 2, 4) + \text{normalize}(- (g_a^{LN} + g_b^{LN}), 0, 2) + \text{normalize}(- d_{ab}, 2, 4) + w_{\text{norm}}. \]  

(2)

where \( g_a^{LD} \) and \( g_b^{LD} \) are the vertical light-to-dark gradients of the image at nodes \( a \) and \( b \), respectively, and \( d_{ab} \) is the Euclidean distance from node \( a \) to node \( b \).

For the inner RPEDC boundary, the dark-light adjacency matrix from equation 1 is used along with a third adjacency matrix calculated based on image intensity, as shown in equation 3.

\[ w_{ab} = \text{normalize}(- (\lambda_a + \lambda_b), 0, 1) + w_{\text{norm}}. \]  

(3)

Limiting the Search Region and Finding the Shortest Path. After calculating the graph weights, we automatically initialize the endpoints and cut the layer boundaries using Djikstra’s algorithm to find the shortest path. We cut the ILM using the dark-light adjacency matrix in a search region ranging from the top of the image to the inner boundary of the RPEDC estimated from the binary mask.

Repeat for Subsequent Layer Boundaries. To segment Bruch’s membrane, we first tentatively cut the inner boundary of the RPEDC using the dark-light adjacency matrix in a search region limited by the inner RPEDC and Bruch’s membrane boundaries estimated by the binary mask. To find the final cut for Bruch’s membrane, we use the second adjacency matrix with combined weights from equation 2 and the tentative inner boundary of the RPEDC as the upper search limit.

Study Dataset Resolutions

<table>
<thead>
<tr>
<th>Study Site</th>
<th>Devers</th>
<th>Duke</th>
<th>Emory</th>
<th>NEI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axial FWHM resolution in retina, ( \mu )m</td>
<td>4.54</td>
<td>4.38</td>
<td>4.56</td>
<td>4.56</td>
</tr>
<tr>
<td>Axial pixel resolution in retina, ( \mu )m/pixel</td>
<td>3.21</td>
<td>3.23</td>
<td>3.06</td>
<td>3.24</td>
</tr>
<tr>
<td>Lateral pixel resolution, ( \mu )m/pixel</td>
<td>6.60</td>
<td>6.54</td>
<td>6.58</td>
<td>6.50</td>
</tr>
<tr>
<td>Azimuthal pixel resolution, ( \mu )m/pixel</td>
<td>68.2</td>
<td>67.0</td>
<td>69.8</td>
<td>65.0</td>
</tr>
<tr>
<td>Scan width, mm</td>
<td>6.60</td>
<td>6.54</td>
<td>6.58</td>
<td>6.50</td>
</tr>
<tr>
<td>Scan length, mm</td>
<td>6.82</td>
<td>6.70</td>
<td>6.98</td>
<td>6.50</td>
</tr>
</tbody>
</table>

Rectangular volumetric scans were acquired with 1000 A-scans and 100 B-scans at all sites. The axial full-width at half-maximum (FWHM) resolution of the SD-OCT system in retina, axial image pixel spacing resolution in retina, lateral pixel resolution (distance between A-scans), and azimuthal pixel resolution (distance between B-scans) varied by site.

Table 2. Volume Quality Metrics

<table>
<thead>
<tr>
<th>Allowable Characteristics</th>
<th>Volume Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low resolution or saturation</td>
<td>Good, good, fair</td>
</tr>
<tr>
<td>Blinking artifacts within frames 20–60</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Motion or loss of fixation within frames 20–60</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Complex conjugate artifact</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Imaging system scan artifact</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Tilt, clipping, blank frames</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Ungradable</td>
<td>✓ ✓ ✓</td>
</tr>
</tbody>
</table>

The characteristics allowed for a particular volume quality. For example, a volume containing blinking artifacts could not be considered as a high- or low-quality volume; therefore, it was excluded from the validation study. A volume exhibiting imaging system scan artifacts, on the other hand, did not affect volume quality.

To refine the inner boundary of the RPEDC, we first estimate the RPE by cutting it using the intensity-based adjacency matrix. We limit the search region to the inner RPEDC boundary estimated by the binary mask and the final cut for Bruch’s membrane. We then recut the inner boundary of the RPEDC using the RPE as the lower boundary and 10 \( \mu \)m above the RPE as the upper boundary of the search region.

Unflattening and Upsampling the Layer Boundaries. Last, we unflatten and upsample the cuts by reversing the flattening and downsampling processes, resulting in the original retinal image with three automatically detected layer boundaries.

Study Dataset

For this study, we considered rectangular volumes with nonneovascular AMD under the A2A SD-OCT study, which was registered at clinicaltrials.gov and approved by the institutional review boards (IRBs) of the four A2A SD-OCT clinics (Devers Eye Institute, Duke Eye Center, Emory Eye Center, and the National Eye Institute). The study complied with the Declaration of Helsinki, and informed consent was obtained from all participants.

In the A2A SD-OCT study, volumetric scans were acquired using the SD-OCT imaging systems from Bioptigen, Inc. (Research Triangle Park, NC) located at the four clinic sites. For each patient across all sites, 0° and 90° rectangular volumes centered at the fovea with 1000 A-scans and 100 B-scans were captured for one eye. The scan sizes and the axial, lateral, and azimuthal resolutions varied slightly by site, and are specified in Table 1. The eye length was not measured. For this study, we included volumes from all four clinical sites to validate algorithm performance for images acquired at slightly varying axial resolutions and by different clinical operators.

As part of the A2A SD-OCT study, each volume was graded for quality by graders certified by the Duke Advanced Research in Spectral Domain OCT Imaging (DARSI) group. In addition to an overall scoring of good, fair, or poor, they assessed these volumes for the following characteristics: (1) foveal centration (a fovea located approximately at the center of the volume); (2) presence of low resolution or saturation; (3) presence of artifacts produced by subject blinking; (4) presence of artifacts produced by eye motion or loss of fixation; (5) presence of complex conjugate artifacts; (6) scan artifacts arising from the imaging system; (7) tilt, clipping, or blank frames; and (8) ungradable. We used these existing scores in our study to classify the volumes as high quality, low quality, or excluded from the study based on the criteria in Table 2. Volumes with motion or loss of fixation artifacts, for example, could not be categorized as high-quality, because they result in inaccurate retinal layer volume measurements. Likewise, we ex-
TABLE 3. Validation Study Volume Selection Criteria

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, n</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Volumes per patient, n</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Total volumes, n</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pathology</td>
<td>Drusen</td>
<td>Drusen</td>
<td>Drusen + GA</td>
</tr>
<tr>
<td>Volume quality</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Scan direction (0°/90°)</td>
<td>Both</td>
<td>Either</td>
<td>Either</td>
</tr>
</tbody>
</table>

All 10 volumes selected for group 1 were used in the Reproducibility analysis, while five volumes from each of the four groups were used in the automatic versus manual segmentation analysis.

RESULTS

Automatic versus Manual Segmentation Analysis

A total of 220 B-scans from 20 volumes were selected for this analysis. Five of these 20 volumes comprised one randomly selected volume from each patient in group 1, and the remaining 15 volumes were those selected from groups 2 to 4 (defined in Table 3). The 11 B-scans from each volume were chosen as follows, with F denoting the B-scan number containing the foveal center: F, F ± 2, F ± 5, F ± 10, F ± 15, and F ± 20.

Two DARI-certified graders performed manual segmentation of the retina by drawing three layer boundaries (inner aspect of the ILM, inner aspect of the RPEDC, and outer Bruch’s membrane) using customized software with a graphic user interface (GUI). During manual segmentation, no outside consultation or communication between graders was allowed. We then performed automatic segmentation using the algorithm described earlier, which was implemented in MATLAB (The MathWorks).

After segmentation, B-scans were cropped by 20% on each side to achieve equal axial and azimuthal lengths in the selected volume. The mean thickness difference between the automatic and manual segmentation of a predetermined (the more senior) grader was calculated for each B-scan. The absolute mean difference and standard deviation across all B-scans were then computed and compared between the automatic and manual segmentation. We also determined the maximum error and the percentage of A-scans with an error >5 pixels (note that the axial resolution varied by site, and therefore the 5 pixels was not converted to the 15.3-16.2-μm range). The same comparison was then conducted between the two manual graders, to estimate intergrader variability.

Reproducibility Analysis

We automatically segmented all B-scans in the 10 volumes from group 1 using the developed software to delineate the inner ILM, inner RPE, and outer Bruch’s membrane boundaries. Based on these segmentation results, we measured the volume of the RPEDC and the total retina (defined in Fig. 1) in millimeters for the region enclosed in a 4-mm-diameter circle centered at the fovea.

We chose a 4-mm-diameter circle to match the automatic versus manual analysis, where we examined the inner 60% of a 6.5- to 7.0-mm volume (Table 1). Using the lateral and azimuthal pixel resolutions of each volume, we summed the total number of pixels enclosed between the upper and lower boundaries of the layer across all A-scans within the circle, to produce the pixel volume for the layer of interest. We calculated the millimeter volume of a pixel by multiplying the axial, lateral, and azimuthal pixel resolutions (Table 1) and then multiplied the pixel volume by this factor. To determine the reproducibility of our segmentation algorithm, we compared the percent difference in the measured volumes between the 0° and 90° scans of the same eye at the same visit.

Reproducibility Analysis

The total retina and RPEDC volumes and the percentages of difference in volume between the 0° and 90° datasets are reported in Table 5. The table shows that the calculated volumes of the total retina and RPEDC measured on a 0° volumetric scan and equivalent 90° scan differed on average by 0.28% ± 0.28% and 1.60% ± 1.57%, respectively.

Performance

We coded the algorithm (MATLAB; The MathWorks), resulting in an average computation time of 1.7 seconds per image (512 × 1000 pixels) on a laptop computer with a 64-bit operating system, a CPU at 1.73 GHz (Core i7; Intel, Mountain View, CA), a 7200 rpm hard drive, and 16 GB of RAM. This time includes the overhead required for reading and writing operations. Manual segmentation took an average time of 3.5 minutes per image.

DISCUSSION

Despite the establishment of predefined segmentation guidelines and practice sessions for manual segmentation on training data sets, two certified graders did not achieve perfect agreement when delineating the layer boundaries (Table 4, column 1). Implementing even more explicit guidelines for manual segmentation may improve agreement, but this will not eliminate the inherent intraobserver variability and differences between manual tracings. Also note that although we excluded RPEDC material over a nearly absent RPE with a minimum lateral width equal to the azimuthal pixel resolution (67 μm in...
this study), future investigators may employ a fixed width to improve uniformity across clinical studies.

Results show that our algorithm automatically segmented the total retina and RPEDC in eyes with intermediate AMD with accuracy comparable to that of a second human grader (Table 4, column 1 versus 2). A low-quality volume did not significantly reduce the segmentation accuracy (Table 4, volume groups 1 vs. 2 and 3 vs. 4), illustrating the algorithm’s robustness for images of various levels of quality. Future study across a dataset of several hundred eyes with intermediate AMD may reveal new segmentation challenges that occur infrequently and thus may not have been identified in this series. We currently do not know the range of changes in RPEDC volume associated with disease progression or how these compare to color fundus photographs, and therefore we cannot be certain of the accuracy required for predictive volume measurements. RPEDC volume measurements from SD-OCT imaging will hopefully provide greater accuracy in assessing drusen load compared to the common technique of mentally summing the area of drusen visible on color fundus photographs.40

**TABLE 4. Automatic Versus Manual Segmentation Results**

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Quality</th>
<th>Volume Group</th>
<th>Retinal Layer Boundary</th>
<th>Mean Error ± SD (μm)</th>
<th>Max Error; Error &gt;5 Pixels (μm; %)</th>
<th>Mean Error ± SD (μm)</th>
<th>Max Error; Error &gt;5 Pixels (μm; %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drusen (110 images)</td>
<td>High</td>
<td>1</td>
<td>Total retina</td>
<td>3.2 ± 2.3</td>
<td>40; 5.5</td>
<td>2.5 ± 1.8</td>
<td>52; 2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>RPEDC</td>
<td>4.1 ± 3.1</td>
<td>49; 5.8</td>
<td>3.0 ± 2.1</td>
<td>52; 3.2</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>2</td>
<td>Total retina</td>
<td>3.3 ± 2.3</td>
<td>67; 10.0</td>
<td>3.7 ± 2.3</td>
<td>67; 7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>RPEDC</td>
<td>4.5 ± 3.5</td>
<td>70; 12.5</td>
<td>2.8 ± 2.0</td>
<td>70; 7.8</td>
</tr>
<tr>
<td>GA (110 images)</td>
<td>High</td>
<td>3</td>
<td>Total retina</td>
<td>4.6 ± 3.9</td>
<td>103; 12.8</td>
<td>5.0 ± 3.0</td>
<td>103; 11.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>RPEDC</td>
<td>4.8 ± 4.3</td>
<td>100; 13.5</td>
<td>4.1 ± 3.5</td>
<td>90; 11.3</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>4</td>
<td>Total retina</td>
<td>2.7 ± 2.3</td>
<td>75; 10.6</td>
<td>5.6 ± 3.0</td>
<td>97; 10.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RPEDC</td>
<td>4.4 ± 3.2</td>
<td>71; 13.4</td>
<td>3.0 ± 2.2</td>
<td>100; 7.8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>Total retina</td>
<td>3.4 ± 2.9</td>
<td>103; 9.7</td>
<td>4.2 ± 2.8</td>
<td>103; 8.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RPEDC</td>
<td>4.5 ± 3.5</td>
<td>100; 11.3</td>
<td>3.2 ± 2.6</td>
<td>100; 7.5</td>
</tr>
</tbody>
</table>

Differences (error) in retinal layer thickness segmentation between two certified manual graders for 220 B-scans (Column 1), as compared to the thickness differences between the automatic segmentation and one certified manual grader of the same 220 B-scans (Column 2). The maximum difference and the percentage of A-scans with a difference greater than 5 pixels are also reported.

![SD-OCT images](image)

**FIGURE 5.** SD-OCT images of eyes with intermediate AMD, without and with automatic segmentation. (A) A high-quality image with both large and small drusen, (B) the segmented image of (A), (C) a low-quality image with small deposits of drusen material, (D) the segmented image of (C), (E) a high-quality image demonstrating an extensive area of GA with irregular reflectivity from outer retinal structures, (F) the segmented image of (E), (G) a low-quality image with an area of GA and an overlying small spot of hyperreflectivity that was not included as REPDC, and (H) the segmented image of (F).
AMD. The algorithm was marginally less accurate for volumes con-
composing the RPE without drusenoid material (under asterisk).
An SD-OCT image from volume group 3 with atrophy of the RPE
and a hyperreflective choroid typical of GA. (D) The automated algo-
rightly segmented hyperreflective structures within the
 choroid as the RPEDC (under bracket).

Our measurement of the RPEDC builds from the known pathophysiology and morphology of AMD and should be useful in
testing hypotheses of disease progression. The term drusen has
been based on yellow spots visible on ophthalmoscopy, and has been recorded with color fundus photographs. They
contain a wide range of materials, including lipids, lipopro-
teins, amyloid, collagen, proteins associated with inflamma-
tion, and degradation products. Although drusen can be
composed of basal laminar deposits (internal to the RPE), basal
linear deposits (external to the basal lamina of the RPE), and
apical or subretinal deposits (reticular drusen), the difference
between aging processes and the onset of AMD remains con-
troversial. Each of these deposits has been implicated
in the pathogenesis of AMD, and it would appear clinically
relevant to identify the early onset of changes in the RPE
associated with AMD. Although large drusen can be readily
segmented from the RPE, small drusen deposits in the early
stages of disease, depending on the pattern of reflectivity,
would likely initially produce a change in RPE volume followed
by a subsequent appearance of distinct drusen as the deposits
enlarge. Thus, because of our interest in identifying RPE and
drusen pathology associated with early AMD, we pursued
RPEDC measurement to capture the full extent of early disease
and chose to compare this to an aged non-AMD control pop-
ulation. This will be important when paired with measures-
ments of the neurosensory retina to investigate the timing of
druses. This long segmentation time was largely attributable
to the difficulty in segmenting the irregularly shaped inner
border of the RPEDC and in distinguishing the RPE and drusen
from extraneous material, such as hyperreflective foci and
drusenoid remnants over GA. Future studies will include a
more in-depth analysis on a larger pool of data and will identify
common automated drusen segmentation errors similar to the
identifications made in other studies.

Even with these limitations, our algorithm segmented
drusen of various shapes and sizes (Fig. 5B), images of signifi-
cantly low quality (Fig. 5D), RPE and drusen in the presence of
GA (Fig. 5F), and retina with irregular curvatures (Fig. 5H).
Furthermore, the <5% difference in measured layer volume,
when comparing 0° and 90° scans of the same eye (Table 5),
attests to the reproducibility of the automatic measurements.
Differences in the measured layer volume may partially be
attributable to the fact that the volumes were unregistered.

Not only did the algorithm segment these images accurately
and reproducibly, but also efficiently. On average, a certified
grader could draw three boundaries on a single B-scan in 3.5
minutes. This long segmentation time was largely attributable
to the difficulty in segmenting the irregularly shaped inner
border of the RPEDC and in distinguishing the RPE and drusen
from extraneous material, such as hyperreflective foci and
drusenoid remnants over GA. Future studies will include a
more in-depth analysis on a larger pool of data and will identify
common automated drusen segmentation errors similar to the
identifications made in other studies.

The clinical implications of these results are encouraging
for large-scale ophthalmic studies, since they suggest that this
automatic segmentation algorithm can efficiently and repro-
cducibly segment the total retina and RPEDC. Furthermore, for
clinical studies with a wide range of image quality, our algo-
rithm is capable of accurately segmenting images of lower
quality. Last, automatic segmentation of the RPEDC contributes
to the progress in drusen quantification, which is especially
important in AMD studies. However, note that the algorithm
segments all drusen types, including soft drusen, cuticular
drusen, and subretinal drusenoid deposits. While soft drusen
and subretinal drusenoid deposits have been shown to be
significant indicators of AMD progression, cuticular
drusen are considered by some as not being associated with
AMD. Our future studies will include the development of
automated drusen classification techniques to segment drusen
types that are specific to a particular disease.

Table 5. Reproducibility Analysis Results

<table>
<thead>
<tr>
<th>Patient</th>
<th>Volume (mm³)</th>
<th>Volume Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Retina</td>
<td>RPEDC</td>
</tr>
<tr>
<td>1</td>
<td>3.45</td>
<td>3.45</td>
</tr>
<tr>
<td>2</td>
<td>3.56</td>
<td>3.57</td>
</tr>
<tr>
<td>3</td>
<td>3.74</td>
<td>3.71</td>
</tr>
<tr>
<td>4</td>
<td>3.46</td>
<td>3.45</td>
</tr>
<tr>
<td>5</td>
<td>3.48</td>
<td>3.49</td>
</tr>
<tr>
<td>Mean</td>
<td>3.54</td>
<td>3.53</td>
</tr>
<tr>
<td>SD</td>
<td>0.12</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Measured layer volumes for 0° and 90° datasets of the same patient and their calculated percentage difference.
Validation of our proposed algorithm was limited to intermediate AMD and was not tested for disease processes such as neovascular AMD, vitreoretinal pathologies, or proliferative diabetic retinopathy. Algorithmic modification, extension of application, and assessment of the performance in eyes exhibiting pathologies outside of nonneovascular AMD is part of our ongoing work. Furthermore, while only volumes with high or low quality were considered in our validation study, this does not imply that the algorithm necessarily err for volumes excluded from the study. These volumes were excluded due to missing retinal data. All such volumes will be included in our future studies identifying common segmentation and acquisition errors on a broader pool of data.

In summary, we developed a fully automatic algorithm to segment three retinal boundaries with a performance comparable to that of manual graders. The algorithm performed reliably for images containing drusen and GA and for images of various levels of quality and yielded reproducible measurements of layer volumes for the same eye. Our automatic approach can reduce time and labor costs and yield an objective evaluation for the study of AMD in future clinical studies.

Acknowledgments

The authors thank Stefanie G. Schuman (Director of Grading for the A2A SD-OCT study) for her contribution in developing the segmentation guidelines for AMD pathology, and Ramiro Maldonado, Michelle McCall, and Neeru Sarin for their contributions to the validation studies.

References

1. Bressler NM. Age-related macular degeneration is the leading cause of blindness. JAMA. 2004;291:1900–1901.
32. Garvin MK, Abramoff MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-


