

# Geodesic Graph Cut Based Retinal Fluid Segmentation in Optical Coherence Tomography

Hrvoje Bogunović<sup>1</sup>, Michael D. Abramoff<sup>2,1,3</sup>, and Milan Sonka<sup>1,2</sup>

<sup>1</sup>Department of Electrical and Computer Engineering

<sup>2</sup>Department of Ophthalmology and Visual Sciences

The University of Iowa, Iowa City, Iowa, USA

<sup>3</sup>Department of Veterans Affairs, Iowa City, USA

{hrvoje-bogunovic,michael-abramoff,milan-sonka}@uiowa.edu

**Abstract.** Age-related macular degeneration (AMD) is a leading cause of blindness in developed countries. Its most damaging form is characterized by accumulation of fluid inside the retina, whose quantification is of utmost importance for evaluating the disease progression. In this paper we propose an automated method for retinal fluid segmentation from 3D images acquired with optical coherence tomography (OCT). It combines a machine learning approach with an effective segmentation framework based on geodesic graph cut. After an image preprocessing step, an artificial neural network is trained based on textural features to assign to each voxel a probability of belonging to a fluid. The obtained probability maps are used to compute minimal geodesic distances from a set of identified seed points to the remaining unassigned voxels. Finally, the segmentation is solved optimally and efficiently using graph cut optimization. The method is evaluated on a clinical longitudinal dataset consisting of 30 OCT scans from 10 patients taken at 3 different stages of treatment. Manual annotations from two retinal specialists were taken as the gold standard. The segmentation method achieved mean precision of 0.88 and recall of 0.83, with the combined  $F_1$  score of 0.85. The segmented fluid volumes were within the measured inter-observer variability. The results demonstrate that the proposed method is a promising step towards accurate quantification of retinal fluid.

## 1 Introduction

Age-related macular degeneration (AMD) is the primary cause of blindness in elderly populations as it directly affects the central retina, deteriorating the sharp vision. The most damaging form is neovascular AMD, which is caused by the abnormal growth of choroidal blood vessels that leak fluid. This results in fluid-filled regions inside the retina called symptomatic exudate associated derangements (SEAD). An effective treatment exists in the form of injecting anti-VEGF (vascular endothelial growth factor) agents, which suppress vessel growth and enable regression of the SEAD. Currently, the treatment frequency is primarily guided by the amount of SEAD, which is subjectively assessed from images noninvasively acquired with optical coherence tomography (OCT). Thus,

having an accurate quantitative estimate of the SEAD sizes is of utmost importance for effective clinical management of patients.

With the widespread use of OCT in ophthalmology, SEAD segmentation has become of increasing interest recently. The first attempt was semi-automatic [1], where a deformable model was manually initialized and grown within a SEAD. Later, a fully automated approach was proposed in [2], where they combined a k-nearest neighbor (k-NN) classification and a graph cut segmentation. Very recently, several automated methods were presented, which followed the generally identical strategy [3–5]. The retina is first over-segmented using thresholding [3], Split-Bregman [4], or watershed [5], after which a classifier is trained or a set of heuristic rules are used to reject false positive detections. More success has been reported in identifying fluid under the retina, which is typically solved by dual-surface segmentation, as the fluid pockets can be found subsequently by thresholding the resulting thickness between the two surfaces [6]. However, despite the recent efforts, SEAD segmentation remains a challenging problem due to low signal-to-noise ratio (SNR) of OCT, SEAD variability in size, shape and location, and interobserver variability between retinal specialists.

In this paper we combine a machine learning approach with an efficient segmentation framework based on geodesic graph cuts. After a texture-based classifier assigns to each voxel a probability of belonging to a SEAD, this information is used to compute minimal geodesic paths to a set of identified points belonging to background and SEAD. The distances are then combined with intensity and edge-based information, and solved optimally with graph cut optimization, to precisely localize the boundaries. This stands in contrast to the recent, more common approach, where the retina is first over-segmented and then the superfluous regions discarded using machine learning. In addition to achieving state of the art results, we believe the approach reported here is better suited for later interactive segmentation refinement since the graph cut segmentation is used at the end, not at the beginning of the processing pipeline.

## 2 Methods

Our approach to SEAD segmentation is based on the premise that SEADs form well connected regions, dark in intensity and homogeneous in texture. We expect to learn both characteristics from a training set of labeled images. We then combine those properties within a geodesic graph cut framework [7].

### 2.1 Preprocessing

OCT images have low SNR due to the strong presence of speckle noise (Fig. 1). Denoising is hence the first step, where we process a 3D volume with a median filter. This is followed by a piece-wise linear intensity transformation [8], designed to enhance contrast and keep retinal tissues having similar intensities across images to help the later classification and segmentation steps. To denote retina and its principal layers, a graph-theoretic surface segmentation is performed

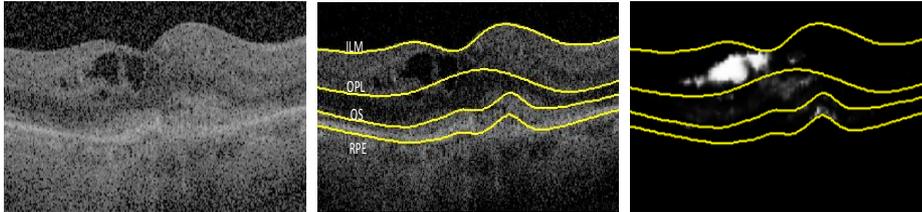


Fig. 1: Left: OCT cross-section (B-scan). Middle: After denoising, intensity normalization and retinal layer segmentation. Right: ANN SEAD probability map  $P_{\mathcal{F}}^{\text{ann}}$ .

based on [9]. Four retinal surfaces are found (Fig. 1): inner limiting membrane (ILM), outer plexiform layer (OPL), inner-outer segment surface (IS-OS), and retinal pigment epithelium (RPE).

## 2.2 Location and Texture Based Voxel Classification

To obtain a local probability of SEAD presence, a recently proposed voxel classification technique was employed [8]. It leverages the fact that retinal fluid is known to exhibit layer dependent properties. For example, fluid in the form of intraretinal cystoid-like cavities appears mainly above the OPL. An ellipsoid-like subretinal fluid occurs below IS-OS and above RPE. The outer nuclear layer (in-between OPL and IS-OS) has very low reflectivity and similar texture to SEADs, hence tends to be a frequent site of false positive detections. Thus, features computed for each voxel include distances to the above-segmented surfaces. In addition, various three-dimensional features, based on textural characteristics, are extracted. Specifically, eigenvalues of Hessian across five scales ( $\sigma = 1, 3, 6, 9, 14$  voxels), zero, first, and second order Gaussian derivatives at three scales ( $\sigma = 2, 4, 8$  voxels), mean intensity, co-occurrence matrix entropy and inertia are all computed. This yields a 52-dimensional feature vector.

Training using a layer-dependent sampling model was performed with an artificial neural network (ANN) with two hidden layers consisting of 10 and 3 nodes, respectively. An advantage of ANN, over e.g. k-NN classifier, is their fast classification at run time, which is of special importance when it is performed on a per-voxel basis of high-resolution images. The resulting voxel probabilities (Fig. 1 right) then serve as soft segmentation constraints as explained next.

## 2.3 Geodesic Graph Cut Segmentation

Given a 3D OCT image  $I$ , the segmentation is posed as searching for binary voxel labeling  $L = \{1, \dots, l_{|I|}\}$ ,  $l \in \{\mathcal{B}, \mathcal{F}\}$ , where the two labels correspond to background (retina) and foreground (SEAD). We then specify the segmentation as the minimization of the following energy functional:

$$E(L) = \sum_{p \in I} R_p(l_p) + \lambda \sum_{(p,q) \in \mathcal{N}} B_{p,q}(l_p, l_q), \quad (1)$$

consisting of unary regional cost  $R_p$  at each voxel  $p$ , and a pairwise boundary cost  $B_{p,q}$  between neighboring voxels  $p, q$ . The two costs are mutually weighted by  $\lambda$ .

Regional probability based on locational and textural properties are represented by the output of the ANN (Fig. 1 right), denoted  $P_l^{\text{ann}}(p)$ , and used in the following way. We first globally threshold the  $P_l^{\text{ann}}$  to obtain for each label  $l$  a set of seed points  $\Omega_l$ , i.e., points that we consider to be part of the foreground  $\mathcal{F}$  and those to be part of the background  $\mathcal{B}$ . Then, from all the voxels  $p \in I$  within the retina we compute a geodesic distance to the nearest foreground and background seeds  $D_l(p) = \min d_l(p, s_l), s_l \in \Omega_l$ . The geodesic distance is the cost of the shortest weighted path

$$d_l(x_0, x_1) = \min_{\gamma_{x_0, x_1}} \int_0^1 |W_l(\gamma_{x_0, x_1}(t)) \cdot \dot{\gamma}_{x_0, x_1}(t)| dt, \quad (2)$$

where  $W_l$  encodes the weights of a particular label and  $\gamma$  is a parameterized path connecting two points  $x_0$  and  $x_1$ . The emphasis is then on designing the weights  $W_l$  such that they are small inside the region corresponding to the label  $l$ , and large otherwise. This will lead to small geodesic distances to voxels that are part of the same region as the seeds [10]. Here, we obtain these weights from the probabilistic output of the ANN by setting  $W_l = 1 - P_l^{\text{ann}}$ . To compute the minimal geodesic distances, a fast marching algorithm is employed solving the Eikonal equation with the speed function  $F_l = 1/W_l$ , where an initial position of the front is defined by the seed points  $\Omega_l$ .

We then define (Eq. 3) normalized relative regional cost  $G_p(l_p)$  for assigning a label  $l_p$  to a voxel  $p$ , which is based on the above-computed geodesic distances. In addition, we add normalized relative intensity cost  $C_p(l_p)$ , based on the voxel intensity  $i_p$ , where the intensity probability distribution  $P_l^{\text{int}}$  of each label is modeled with a histogram of voxels obtained from a labeled training set (Fig. 2).

$$G_p(l_p) = \frac{D_{l_p}(p)}{D_{\mathcal{F}}(p) + D_{\mathcal{B}}(p)}; \quad C_p(l_p) = 1 - \frac{P_{l_p}^{\text{int}}(i_p)}{P_{\mathcal{F}}^{\text{int}}(i_p) + P_{\mathcal{B}}^{\text{int}}(i_p)}. \quad (3)$$

Now, we can define the unary regional term as:

$$R_p(l_p) = G_p(l_p) + C_p(l_p) + S_p(l_p), \quad (4)$$

where the seed points serve as segmentation hard constraints and are encoded with the following cost

$$S_p(l_p) = \begin{cases} 0 & \text{if } p \in \Omega_{l_p} \\ \infty & \text{otherwise} \end{cases}. \quad (5)$$

Finally, the pairwise boundary term is based on the difference of neighboring intensity probabilities when their labels differ.

$$B_{p,q}(l_p, l_q) = \exp \left( - \frac{\left( P_{l_p}^{\text{int}}(i_p) - P_{l_q}^{\text{int}}(i_q) \right)^2}{2\sigma^2} \right) |l_p - l_q|, \quad (6)$$

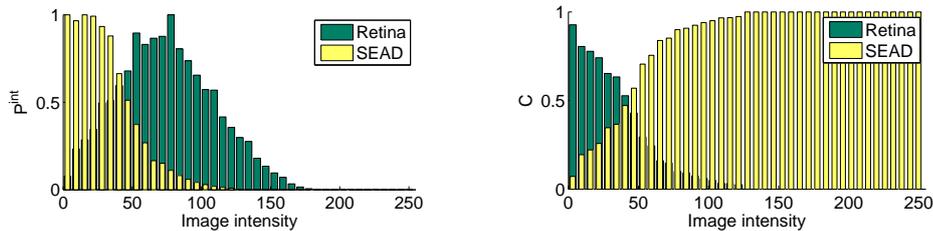


Fig. 2: Left: Normalized intensity probability  $P^{\text{int}}$ . Right: Resulting normalized relative intensity cost  $C$ . Forty bins spanning the intensity range were used.

where  $\mathcal{N}$  is a set of pairs of neighboring voxels (6-neighborhood used), and  $\sigma$  is a parameter. The optimal solution of the submodular functional (Eq. 1) is found exactly and efficiently (low-order polynomial complexity) with min-cut/max-flow graph-partitioning optimization algorithm [11].

### 3 Results

The proposed method is evaluated on a clinical longitudinal dataset consisting of 30 OCT volumes from 10 AMD patients taken at 3 different stages of treatment with anti-VEGF injections. The baseline scan, 2-month, and 6-month follow-up scans were used for evaluation. Image acquisitions were performed with the commercially available swept-source SS-OCT scanner (Topcon Inc., Pyramid, NJ) with 1050 nm central wavelength, centered on the macula. The device acquires anisotropic 3D images having  $512 \times 128 \times 885$  voxels with the size of  $11.7 \times 46.9 \times 2.6 \mu\text{m}^3$ , covering the volume of  $6 \times 6 \times 2.3 \text{mm}^3$ . Two retinal specialists manually annotated all the intraretinal and subretinal SEADs, which were taken as gold standard segmentations for evaluation.

Leave-one-patient-out cross-validation was performed to obtain voxel probabilities. The weighting term  $\lambda$  (Eq. 1) and the boundary cost parameter  $\sigma$  (Eq. 6) were experimentally set to 10 and 0.1, respectively, by observing the performance on an independent set of images. The performance was not sensitive to the parameter choice within the same order of magnitude. The thresholds for obtaining region seeds were set to 0.6 for the SEAD (twice the optimal value for voxel classification, as shown later), and 0.98 (a high value) for the background.

*Qualitative evaluation* For visual comparison we show a series of examples (each from a different patient) of segmentation and gold standard annotations (Fig. 3). It can be observed that the method delineates the SEAD boundaries well, something a voxel-classification-only approach is not designed to do.

*Quantitative evaluation* Manual annotation of SEADs is a very difficult and tedious task and we found that the inter-observer variability was large, with the two observers having a mean Dice overlap index of only 0.55. In particular,

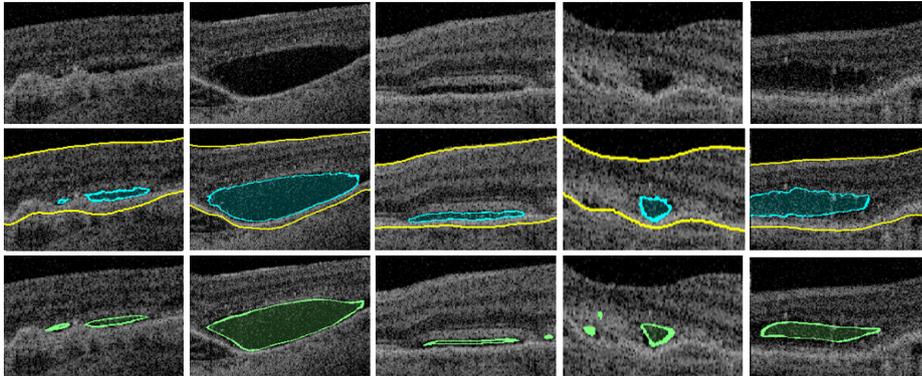


Fig. 3: Top row: OCT intensity image. Middle row: Results of SEAD segmentation (in blue) with segmented surfaces denoting retina (in yellow). Bottom row: Gold standard annotations (in green).

on the follow-up scans, where the SEAD volume decreases substantially due to treatment response, in  $\approx 70\%$  of cases the two manual annotations had no overlap. Thus, we evaluate separately the baseline scans where SEAD presence is abundant from the follow-up scans where SEAD appears less often.

Evaluation of baseline scans is based on measuring *precision* and *recall*, which due to large inter-observer variability we measure in the following way. For measuring the *recall* ( $TP/(TP + FN)$ ), we considered the positive area to be the intersection of annotations, i.e., we measure what percentage of SEAD volume was recovered that both observers agreed on. For computing the *precision* ( $TP/(TP + FP)$ ), we considered the positive area to be the union of annotations, i.e., we measure what percentage of our segmentation has been marked as SEAD by at least one observer. Finally the precision and recall are combined with an  $F_\beta$  score

$$F_\beta = (1 + \beta^2) \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}},$$

where  $\beta$  was set to 1 to weigh the precision and recall, equally.

The results on the baseline scans are given in Table 1, where we furthermore directly compare the proposed method with the voxel classification only approach [8], where the segmentations were obtained by thresholding at 0.3 level, which maximized the mean  $F_1$  score. It can be observed that the inclusion of geodesic graph cuts improves the performance, with the statistically significant difference in  $F_1$  scores (paired t-test,  $p < 0.05$ ).

Evaluation of follow-up scans focused on the total SEAD volume recovered. The gold standard volume was taken to be the mean volume from the annotations of the two observers. We then compute volume error and the inter-observer variability as the difference of the segmented volumes from the gold standard ones. Comparison across the treatment stages is shown in Fig. 4. The segmented volume errors are shown to be within the inter-observer variability.

Table 1: Performance across segmentation methods: Voxel classification (VC), and the geodesic graph cut (VC+GeoGC). The values are expressed as mean (std).

	VC~[8]	VC+GeoGC (proposed)
Precision	0.79 (0.18)	0.88 (0.07)
Recall	0.81 (0.08)	0.83 (0.11)
$F_1$	0.79 (0.13)	0.85 (0.09)

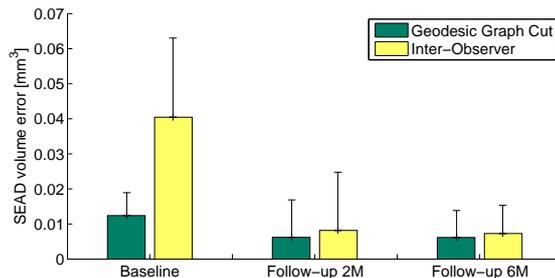


Fig. 4: Volume errors of our geodesic graph cut approach compared to the inter-observer variability.

Finally, due to a very high resolution of OCT images ( $\approx 60M$  voxels), computational efficiency is of high importance. Feature computation and ANN classification required  $\approx 60$  sec, and the geodesic graph cut segmentation  $\approx 40$  sec, on a Xeon CPU with 3.6 GHz.

## 4 Conclusion

We proposed a robust geodesic graph cut framework for SEAD segmentation from OCT images. As opposed to majority of related work that segment a scan slice-by-slice (i.e., per B-scan) ours operates fully in 3D, taking advantage that OCT is a true 3D imaging modality. Graph cut based segmentations are especially suitable for the task as they are able to find the solution of a submodular energy optimally and very efficiently.

The main contribution is in combining the state-of-the-art voxel classification with the improved graph cut approach. An added value of using geodesic graph cut is that the geodesic term decreases the boundary-length bias of a standard graph cut formulation. We believe the geodesic component would be even more valuable in the subsequent segmentation refinement setting [11], where user-provided scribbles would add or remove the identified region seed points, as it would force updates to be local with respect to the provided scribbles. Although the proposed method clearly adds value to the voxel classification, ultimately its performance is limited by the capabilities of the underlying classifier.

The evaluation was done in a very realistic clinical setting on images taken from patients undergoing AMD treatment. However, it was challenging due to highly variable manual annotations. The manual annotations are not only tedious but there is a substantial disagreement and different clinical interpretations between retinal specialists. This further motivates the need for developing automated, reproducible segmentation methods. Nevertheless, the results demonstrate that the proposed method is a promising step towards the accurate quantification of SEADs, much needed for the successful clinical management of AMD patients.

**Acknowledgments.** This work was supported by NIH grants R01 EY019112, R01 EY018853 and R01 EB004640; the Department of Veterans Affairs.

## References

1. Fernández, D.C.: Delineating fluid-filled region boundaries in optical coherence tomography images of the retina. *IEEE Trans. Med. Imag.* **24**(8) (2005) 929–945
2. Chen, X., Niemeijer, M., Zhang, L., Lee, K., Abramoff, M.D., Sonka, M.: Three-dimensional segmentation of fluid-associated abnormalities in retinal OCT: probability constrained graph-search-graph-cut. *IEEE Trans. Med. Imag.* **31**(8) (2012) 1521–31
3. Wilkins, G.R., Houghton, O.M., Oldenburg, A.L.: Automated segmentation of intraretinal cystoid fluid in optical coherence tomography. *IEEE Trans. Biomed. Eng.* **59**(4) (2012) 1109–14
4. Ding, W., Young, M., Bourgault, S., Lee, S., Albiani, D.a., Kirker, A.W., Forooghian, F., Sarunic, M.V., Merkur, A.B., Beg, M.F.: Automatic detection of subretinal fluid and sub-retinal pigment epithelium fluid in optical coherence tomography images. In: *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*. (2013) 7388–91
5. Gonzalez, A., Remeseiro, B., Ortega, M., Penedo, M.G., Charlon, P.: Automatic cyst detection in OCT retinal images combining region flooding and texture analysis. In: *Proc. IEEE Int. Symp. Comput. Med. Syst. (CBMS)*. (2013) 397–400
6. Shi, F., Chen, X., Zhao, H., Zhu, W., Xiang, D., Gao, E., Sonka, M., Chen, H.: Automated 3-D retinal layer segmentation of macular optical coherence tomography images with serous pigment epithelial detachments. *IEEE Trans. Med. Imag.* **34**(2) (2015) 441–52
7. Price, B., Morse, B., Cohen, S.: Geodesic graph cut for interactive image segmentation. In: *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recogn. (CVPR)*. (2010) 3161–68
8. Xu, X., Lee, K., Zhang, L., Sonka, M., Abramoff, M.D.: Stratified sampling voxel classification for segmentation of intraretinal and subretinal fluid in longitudinal clinical OCT data. *IEEE Trans. Med. Imag.* **34**(7) (2015) 1616–1623
9. Quéllec, G., Lee, K., Dolejsi, M., Garvin, M.K., Abramoff, M.D., Sonka, M.: Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula. *IEEE Trans. Med. Imag.* **29**(6) (2010) 1321–30
10. Protiere, A., Sapiro, G.: Interactive image segmentation via adaptive weighted distances. *IEEE Trans. Image Process.* **16**(4) (2007) 1046–57
11. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*. (2001) 105–112