Published in final edited form as:

IEEE Trans Med Imaging. 2015 September ; 34(9): 1854–1866. doi:10.1109/TMI.2015.2412881.

Multimodal Segmentation of Optic Disc and Cup from SD-OCT and Color Fundus Photographs Using a Machine-Learning Graph-Based Approach

Mohammad Saleh Miri,

Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA, 52242 (mohammadsalehmiri@uiowa.edu).

Michael D. Abramoff [Senior Member, IEEE],

Department of Ophthalmology and Visual Sciences and the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA, 52242. He is also with the Iowa City VA Health Care System, Iowa City, IA, 52246 (michael-abramoff@uiowa.edu).

Kyungmoo Lee,

Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA, 52242 (kyungmoo-lee@uiowa.edu).

Meindert Niemeijer,

IDx, LLC, Iowa City, IA, 52244 (meindertn@gmail.com).

Jui-Kai Wang,

Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA, 52242 (jui-kaiwang@uiowa.edu).

Young H. Kwon, and

Department of Ophthalmology and Visual Sciences, The University of Iowa, Iowa City, IA, 52242 (youngkwon@uiowa.edu).

Mona K. Garvin [Member, IEEE]

Iowa City VA Health Care System, Iowa City, IA, 52246 and the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 (mona-garvin@uiowa.edu).

Abstract

In this work, a multimodal approach is proposed to use the complementary information from fundus photographs and spectral domain optical coherence tomography (SD-OCT) volumes in order to segment the optic disc and cup boundaries. The problem is formulated as an optimization problem where the optimal solution is obtained using a machine-learning theoretical graph-based method. In particular, first the fundus photograph is registered to the 2D projection of the SD-OCT volume. Three in-region cost functions are designed using a random forest classifier corresponding to three regions of cup, rim, and background. Next, the volumes are resampled to

create radial scans in which the Bruch's Membrane Opening (BMO) endpoints are easier to detect. Similar to in-region cost function design, the disc-boundary cost function is designed using a random forest classifier for which the features are created by applying the Haar Stationary Wavelet Transform (SWT) to the radial projection image. A multisurface graph-based approach utilizes the in-region and disc-boundary cost images to segment the boundaries of optic disc and cup under feasibility constraints. The approach is evaluated on 25 multimodal image pairs from 25 subjects in a leave-one-out fashion (by subject). The performances of the graph-theoretic approach using three sets of cost functions are compared: 1) using unimodal (OCT only) in-region costs, 2) using multimodal in-region costs, and 3) using multimodal in-region and disc-boundary costs. Results show that the multimodal approaches outperform the unimodal approach in segmenting the optic disc and cup.

Index Terms

SD-OCT; ophthalmology; retina; multimodal; segmentation; optic disc; Bruch's Membrane Opening

I. Introduction

GLAUCOMA is among the common causes of blindness worldwide. While visual field tests are the clinical standard for monitoring functional deficits, structural imaging modalities, such as color (stereo) fundus photography and, more recently, spectral-domain optical coherence tomography (SD-OCT) [1], are also important to monitor the characteristic optic nerve cupping and other structural changes that are characteristic of this progressive disease (Fig. 1). For example, one structural parameter of interest that is obtainable from both color fundus photographs and SD-OCT volumes is the cup-to-disc ratio (CDR), which is defined as the cup area over the optic disc area. A larger CDR is associated with glaucomatous damage because of the presence of fewer remaining nerve fiber bundles in the neuroretinal rim.

With the inherent subjectivity and time required for expert optic disc and cup segmentations, there has been great interest in automated approaches. Most prior optic disc and cup segmentation approaches have focused on segmenting color fundus images alone [2]–[10], with few approaches having been presented for the segmentation within SD-OCT volumes alone [11]–[13]. More specifically, the prior fundus-only approaches for segmenting the disc and cup include pixel-based classification methods [2], [3], model-based approaches [4]–[7], and graph-based approaches [8]–[10]. Initial SD-OCT-only approaches include the work of Lee et al. [11] and Abramoff et al. [12] where each A-scan (i.e., projected pixel location) was classified as cup, rim, or background from SD-OCT features within the A-scan. While the final approach was an SD-OCT-only approach, the pixel-classification-approach was trained and tested using expert-marked color fundus images.

Another type of SD-OCT approach has focused on directly segmenting the termination of Bruch's membrane at the optic nerve head – Bruch's membrane opening (BMO, Fig. 2), also referred to as the neural canal opening (NCO) – within SD-OCT volumes using a graph-based approach [13]. Anatomically, it was originally assumed that the projected location of

the BMO would coincide with that of the clinical disc margin visible in a color fundus photograph. (In fact, the automated BMO segmentation work [13] was evaluated via comparisons with the clinical disc margin from color fundus photographs.) While it has been found that the projected location of the BMO often does indeed closely coincide with the clinical disc margin, differences do exist, such as in the presence of externally oblique border tissue [13]–[16], challenging the original anatomical assumptions behind the fundus-based clinical disc margin. Correspondingly, when available, the BMO (or, more precisely, the projected BMO for computation of 2D parameters) is becoming recognized as an appropriate disc margin definition.

While SD-OCT-only approaches have been shown to outperform fundus-only approaches [2], [17] (in part due to their ability to provide 3D information of anatomic landmarks such as the BMO) for the segmentation of the disc and cup, the color information and higher resolution in the enface (i.e., projected plane) in color fundus photographs (6–10 microns in fundus photographs versus 30–100 microns spacing between A-scans of SD-OCT volumes) provides complementary information to help localize such structures. Given that both fundus and SD-OCT images are often acquired for the assessment of glaucoma and with existence of fundus-OCT image-registration algorithms [18]–[21], it makes sense to combine the complementary information from both sources for the segmentation of the disc and cup. Use of multimodal information is further justified via prior work in the multimodal segmentation of the retinal blood vessels by Hu et al. [21], where it was shown that approaches that combined the use of fundus and SD-OCT information outperformed SD-OCT-only approaches.

Correspondingly, the purpose of this work is to develop a multimodal approach that utilizes information from both SD-OCT volumes and color fundus photographs to segment the disc and cup. In particular, after fundus-to-OCT image registration, our approach uses a machine-learning strategy to compute the likelihood of each projected pixel belonging to either cup or rim or background classes (from multimodal features) and the likelihood of belonging to the projected BMO boundary locations (from SD-OCT features). These in-region and disc-boundary likelihoods are then used as part of the cost functions in an optimal graph-based approach to simultaneously find the disc and cup boundaries. A preliminary version of our multimodal region-based-classification-only results were presented in [17], but we did not use a graph-based strategy to find the final boundaries and did not compute the BMO boundary likelihoods. Furthermore, to reflect our updated understanding of the best reference to use for the disc margin, the projected location of the BMO is used as the disc margin reference standard for training/testing in this work rather than the disc margin marked from color fundus photographs. (The expert-marked cup boundary from fundus photographs is still used as the reference standard for the cup.)

II. Methods

An overall flowchart for the proposed method is shown in Fig. 3. The four major components are as follows: 1) a preparation step including SD-OCT intraretinal layer segmentation, fundus-to-OCT registration, and radial transformation (Section II-A), 2) in-region cost function computation (Section II-B), 3) disc-boundary cost function computation

(Section II-C), and 4) optic disc and cup boundaries segmentation using a theoretical graph-based approach (Section II-D). The in-region and disc boundary cost functions contain the information regarding the likelihood of a pixel belonging to each region (cup, rim, and background) and disc boundary, respectively. Both the in-region and disc-boundary cost functions are created using a machine-learning approach. Features used for in-region cost function computation come from both modalities (a multimodal feature set) and features for the disc-boundary cost function computation are extracted from SD-OCT volumes only. The likelihood maps that are computed, as described in more detail in Sections II-B and II-C, are then utilized in the theoretical graph-based approach described in Section II-D to find the globally optimal (with respect to the cost functions) boundaries of the optic disc and cup.

A. Intraretinal Layer Segmentation, Registration, and Radial Transformation

The first part of the proposed method includes several steps in preparation for the extraction of features from two modalities for use in the machine-learning cost-function design steps.

1) Intraretinal Layer Segmentation—The intraretinal surfaces are first segmented within the 3D SD-OCT volumes (Fig. 4) using a theoretical multi-resolution graph-based approach [11], [22] to: (1) enable the creation of a 2D SD-OCT projection image to be used for fundus-to-OCT image registration, and (2) to enable layer-based features to be extracted. In particular, the surface segmentation problem is formulated as an optimization problem with the goal of simultaneously finding a set of feasible surfaces with the minimum cost. The problem is transformed as a minimum-cost closed set graph problem that is solved through computation of a minimum s - t cut on an additionally transformed graph [22]. Performing the process in a multi-resolution fashion speeds up the segmentation [11]. Three surfaces are used for further computations (although eleven surfaces are originally segmented). The first surface corresponds to the Internal Limiting Membrane (ILM). Surface two is the junction of the inner and outer segments of photoreceptors (IS/OS), and surface three is the outer boundary of the Retinal Pigment Epithelium (RPE), also called the Bruch's membrane surface. Using the method described in [22] a thin-plate spline is fitted to the third surface in order to flatten the OCT images and to enable a consistent optic nerve head shape across patients. Afterwards, an SD-OCT projection image is created by averaging the voxel intensities in the z -direction between the IS/OS junction and Bruch's membrane surfaces.

2) Registration—In order to benefit from the complementary information of both modalities at a pixel/A-scan level, registration needs to be performed. The 2D projection image obtained from the 3D SD-OCT volume (Section II-A1) is utilized to register the 2D fundus photograph to the SD-OCT projection image (Fig. 5). The registration is performed on the fundus and the SD-OCT projection vessel maps. A pixel classification based vessel segmentation algorithm is used to obtain the vessel probability maps of fundus photographs [23] and SD-OCT projection images [24]. The fundus image vessel map is subsampled using linear interpolation so that the optic disc is about the same size in both images. Next, a skeletonization algorithm is applied to the binary vessel segmentations to obtain the vessel centerlines. All centerline pixels with more than two neighbors are removed to obtain a set of vessel segments. An Iterative Closest Point (ICP) algorithm [25] is then applied in two

phases: a first rough initial alignment and a second fine alignment. The rough initial alignment is performed using the simulated annealing optimization algorithm [26]. The finer alignment is performed using the Powell optimization algorithm [27]. The cost function, F , in (1) below for both these steps is the same and consists of the distance between closest points, d (as in a traditional ICP approach), and additionally the difference between the local vessel orientation in radians between the target and the moving image, d_θ :

$$F = (1+d) \times (1+d_\theta). \quad (1)$$

The additional term, d_θ , causes the cost function to have a lower value when the centerline pixels are matched and the local orientation of vessels are similar [28].

3) Radial transformation and layer segmentation—As a preliminary step in the creation of the disc-boundary costs (Section II-C), each original SD-OCT volume in the Cartesian domain (xyz) is resampled (with θ -precision = 1°) using bilinear interpolation to create a radial volume ($r\theta z$). With such a transformation, each slice contains two BMO endpoints (whereas in the original volume, BMO points become less obvious in slices approximately tangent to the superior/inferior disc boundary, Fig. 6a, and are additionally not present in slices that do not intersect with the disc).

Once the radial scans are created, the same approach as explained in [11] is used to segment the ILM surface, IS/OS junction and RPE-complex lower bound. Since the RPE-complex ends at the optic disc margin, the RPE layer segmentation inside the optic disc region is not valid. Therefore, an optic disc margin (1.73 mm from the center) which is larger than the typical optic disc size, is assigned to every volume, and within this margin, the second and third surfaces are interpolated (Fig. 6b).

A radial projection image (Fig. 6c) is then created to reflect the (maximum) intensities of the interpolated RPE-complex surfaces. In particular, for each A-scan, the maximum intensity value in the z -direction 15 voxels above ($29.3\mu\text{m}$) the second surface and 15 voxels below the third surface is selected and the projection image is reformatted such that radial values $(0, 100]$ appear on the vertical axis and θ values $[0^\circ, 360^\circ)$ appear on horizontal axis. The radial projection image is used in Section II-C.

B. In-region Cost Function Computation

The flowchart of the algorithm for creating in-region probability maps is depicted in Fig. 7. The in-region likelihood maps contain the regional costs associated with each pixel. A lower cost indicates that under feasibility constraints the pixel has a higher chance to be assigned to the corresponding region. Three in-region cost functions corresponding to three classes of cup, rim, and background are computed by training a random forest classifier [29] using a multimodal feature set.

1) SD-OCT Features—There are eight features extracted for each projected x - y location from SD-OCT volumes (similar to features in [11]). The first one is the intensity of the same projection image used for registering the fundus image to SD-OCT volume. The next features are the average intensities of four subvolumes above the reference spline fitted to

the outer boundary of RPE (90–120, 60–90, 30–60, 0–30 voxels above) and the average intensities of two subvolumes below the reference spline (0–30 and 30–60 voxels below). The last feature is the optic disc depth information which is extracted by measuring the distance between the ILM surface and the reference spline. An example set of SD-OCT feature is shown in Fig. 8.

2) Fundus Features—As in [2] three intrinsic color channels of red, green, and blue as well as three color-opponent channels of dark-bright, blue-yellow, and red-green are extracted from fundus photographs. The color-opponent channels are computed as in the equations below:

$$I_{\text{dark-bright}} = \frac{I_{\text{red}} + I_{\text{green}} + I_{\text{blue}}}{3} \quad (2)$$

$$I_{\text{blue-yellow}} = I_{\text{red}} + I_{\text{green}} - 2 \times I_{\text{blue}} \quad (3)$$

$$I_{\text{red-green}} = I_{\text{red}} - I_{\text{green}} \quad (4)$$

Note that the original use of the color-opponent channels as in [2] is motivated by the theory of color vision [30] where three signals are produced in the retina by three different cone types in response to a light stimulation. These signals are transformed to three opponents before passing to the brain. The first opponent contains illumination information and is achromatic (the “dark-bright” opponent). The other two opponents (the “red-green” and “blue-yellow” opponents) contains color information.

A zero-order Gaussian filter bank ($\sigma=4, 6, 8$) is applied to all color channels to extract features in different scales as shown in Fig. 9.

3) Spatial features—In the registered fundus-OCT images, it is more likely that the pixel in the center of the image belongs to optic cup than to rim area or background. Therefore, in order to incorporate the *a priori* expectation of the class of a pixel based on its location in the image, a principal component analysis is performed on all reference standard images in the training set and only the first principal component is kept. The *a priori* expectation of the class maps of the cup and disc are similar to the shapes of 2D Gaussians. An *a priori* map describing the rim area is created by subtracting the *a priori* map of cup from the *a priori* map of optic disc. Three *a priori* maps are shown in Fig. 10. In addition to the three *a priori* maps, three optic disc center-based features are extracted. In particular, after defining the lowest point (in the *z*-direction) of the first intraretinal surface as the center of the optic disc, the distance of the *x* and *y* positions, as well as the radial distance of each pixel, are measured with respect to this disc center.

4) Classification—Once the features are extracted from both modalities, a random forest classifier [29] is used to classify each pixel into optic cup, rim, or background. A summary of the 31 multimodal features that are used for training the random forest classifier are listed in Table I.

Random forests are categorized as ensemble classifiers. Once a feature vector is entered to a random forest classifier, all trees (N) in the forest classify the input feature vector based on the specific number (m) of randomly selected features from the input feature vector at each decision split. The final classification is determined by taking the majority vote over the entire forest. In addition to robustness, the other advantage of this classifier is the very low number of parameters to be tuned. These are the number of trees in the forest N and the number of features to be randomly selected at each decision split m . Here, we used $N = 500$ trees (larger numbers increased the training time without improving the accuracy) and $m = 10$.

The likelihood map of each class at a pixel is computed by dividing the number of trees voted for each class by the total number of the trees in the forest. In order to convert the likelihood maps to the in-region cost functions, they are inverted (1–probability map) such that a higher probability value in the likelihood map corresponds to a lower cost in the in-region cost function.

C. Disc-boundary Cost Function Computation

The flowchart of the disc-boundary cost function design is shown in Fig. 11. Similar to the in-region cost function design, a machine-learning approach is used to create the disc-boundary cost function. The features come from processing the radial projection image using a stationary wavelet transform (Section II-C1) and spatial features (Section II-C2). The reference standard is the projected location of the BMO endpoints so that, after the classification step (Section II-C3), the disc-boundary cost function will have a low cost at expected boundary positions.

1) Stationary wavelet transform—The Haar stationary wavelet transform [31] is a translation-invariant type of digital wavelet transform and can be used to suppress the effects of the shadows of blood vessels that make the boundary of the BMO difficult to detect in projection images. For example, the multi-scale attribute of SWT makes it capable of capturing different sizes of blood vessels in different decomposition levels. Furthermore, since the blood vessels tend to be vertically placed in the radial SD-OCT projection image (Fig. 6c) the directionality of the SWT enables it to separate the optic disc boundary (appears in horizontal coefficients) from the blood vessel (appears mostly in vertical and diagonal coefficients).

An example of projection image decomposition using the Haar SWT is shown in Fig. 12. The vessels mostly appear in vertical coefficient images whereas the optic disc boundary is easily recognizable in the horizontal coefficient images. Therefore, the intensity of horizontal coefficients of the 1st, 2nd, 3rd and 4th decomposition levels are added to the feature set. The horizontal coefficients in level 5 and 6 are not considered because at higher decomposition levels (coarser levels), the boundary of interest becomes blurrier. Note that the output of each decomposition level in SWT has the same number of samples (coefficients) as the input which makes the SWT an inherently redundant procedure.

Furthermore, a vessel-free projection image is created from the wavelet coefficients by suppressing all the vertical and diagonal coefficients (setting them to zero) at all levels and

keeping the approximation and horizontal components at each level. Applying the Inverse SWT (ISWT) to the modified wavelet coefficients creates a vessel-free projection image (Fig. 13e). The intensity of the vessel-free projection image along with its derivative (Fig. 13f), computed using an averaging derivative of Gaussian filter, are added to the feature set as well.

2) Spatial features—In addition to the above SWT-based features, there are two spatial features created to learn the position of the optic disc boundary in the radial projection images. The first spatial feature intuitively reflects the radial distance from a rough circular approximation of the optic disc boundary. More specifically, as any circle centered at the optic disc center appears as a horizontal line in the radial vessel-free projection image, P_{radial} , we first locate the first “dark” horizontal line (from the top) $r = r_{\text{dark}}$ in the projection image as follows:

$$r_{\text{dark}} = \min\{r \mid I_{\text{row}}(r) \leq \mu_{I_{\text{row}}} - \sigma_{I_{\text{row}}}\}, \quad (5)$$

where

$$I_{\text{row}}(r) = \frac{1}{360} \sum_{\theta=0}^{359} P_{\text{radial}}(r, \theta), \quad (6)$$

$$\mu_{I_{\text{row}}} = \frac{1}{R} \sum_{r=1}^R I_{\text{row}}(r), \text{ and} \quad (7)$$

$$\sigma_{I_{\text{row}}} = \sqrt{\frac{1}{R-1} \sum_{r=1}^R (I_{\text{row}}(r) - \mu_{I_{\text{row}}})^2}. \quad (8)$$

Once the first dark row is detected using (5), the signed radial distance from this row is considered as the spatial feature. This information helps classifier to narrow the spatial search region for detecting the optic disc boundary (Fig. 13g).

The second spatial feature contains information regarding the *a priori* knowledge of the BM-ILM anatomical relationship that dictates that Bruch’s membrane (and correspondingly the BMO), even in the severe cases of glaucoma where there is no rim tissue left, may touch the ILM surface, but not cross with it. In order to include that information in the feature set, the radial distance to the crossing location of ILM at the level of the interpolated surface three is also considered as a spatial feature (Fig. 13h).

3) Classification—The summary of the features used for disc-boundary classification is as follows:

1. Features 1–4: The intensity of SWT horizontal coefficients of the 1st, 2nd, 3rd and 4th decomposition levels.

2. Feature 5: The intensity of the vessel-free projection image.
3. Feature 6: The intensity of the derivative of the vessel-free projection image.
4. Features 7–8: The spatial features carrying the information regarding the anatomy and shape of the optic disc boundary.

Once the feature set is completed, a random forest classifier (with $N = 500$ trees and $m = 3$ randomly selected features at each decision split) is trained to learn the position of BMO end points using the extracted features to classify the pixels into two classes of boundary and non-boundary (background). The classes are highly imbalanced as for each radial image (100×360), the size of the background class (99×360) is much larger than the size of the boundary class (360). Therefore, the class priors are presented to the random forest classifier as well.

A probability map with low values at the expected boundary locations needs to be used as the disc-boundary cost function for the graph-theoretic step. Hence, similar to in-region cost function computation, the probability map of optic disc boundary is inverted (1 – probability map).

D. Theoretical Graph Based Boundary Detection

The underlying graph-based method that is used for simultaneously segmenting the optic disc and cup boundaries is similar to what is described in [22]. The problem of multiple boundary segmentation can be considered an optimization problem with the goal of finding the feasible optic disc and cup boundaries with the minimum cost. The feasibility constraints include a *boundary smoothness* constraint to ensure that each individual boundary is smooth and a *boundary interaction* constraint to specify the minimum and maximum allowed distances between the boundaries. A schematic representation of the problem of segmenting optic disc and cup boundaries is shown in Fig. 14. The following equations are written in the radial domain. Consider a 2D image in the radial domain $I(r, \theta)$ of size $R \times \Theta$ and assume that the cup and disc boundaries can be defined as $f_{cup}(\theta)$ and $f_{disc}(\theta)$ functions (mapping θ values to r values), respectively. Also, assume the functions use a two-neighbor relationship so that for each column, each boundary intersects with each column (each θ) once (at r). The cup boundary smoothness constraint for neighboring columns $\{\theta_1, \theta_2\}$ in the θ -direction is defined as below

$$-\Delta_\theta \leq f_{cup}(\theta_1) - f_{cup}(\theta_2) \leq \Delta_\theta. \quad (9)$$

The smoothness constraint of the optic disc boundary is similar to that of the cup boundary. For the boundary interaction constraints, the minimum allowed distance between the cup boundary and the disc boundary is defined as $\delta_{cup-disc}^l$ and similarly the maximum allowed distance is defined as $\delta_{cup-disc}^u$.

There are three regions: the optic cup region, the rim region (the region between optic disc and optic cup), and the background. The in-region cost functions associated with the three regions are $C_{cup-reg}(r, \theta)$, $C_{rim-reg}(r, \theta)$, and $C_{background-reg}(r, \theta)$, respectively. $C_{disc-bound}(r, \theta)$ is the disc-boundary cost function. The total cost of a set of boundaries, $C_{\{f_{disc}(\theta), f_{cup}(\theta)\}}$,

can be written as the weighted sum of in-region costs and the disc-boundary cost as given below:

$$C_{\{f_{\text{disc}}(\theta), f_{\text{cup}}(\theta)\}} = \alpha C_{B_{\text{disc}}} + (1 - \alpha)[C_{R_{\text{cup}}} + C_{R_{\text{rim}}} + C_{R_{\text{background}}}], \quad (10)$$

where

$$C_{B_{\text{disc}}} = \sum_{\{(r, \theta) \mid r = f_{\text{disc}}(\theta)\}} C_{\text{disc-bound}}(r, \theta) \quad (11)$$

and

$$C_{R_{\text{cup}}} = \sum_{(r, \theta) \in R_{\text{cup}}} C_{\text{cup-reg}}(r, \theta)$$

$$C_{R_{\text{rim}}} = \sum_{(r, \theta) \in R_{\text{rim}}} C_{\text{rim-reg}}(r, \theta) \quad (12)$$

$$C_{R_{\text{background}}} = \sum_{(r, \theta) \in R_{\text{background}}} C_{\text{background-reg}}(r, \theta).$$

Note that $C_{B_{\text{disc}}}$ represents the total cost corresponding to pixels on the disc boundary and $C_{R_{\text{cup}}}$, $C_{R_{\text{rim}}}$, and $C_{R_{\text{background}}}$ represents the cost related to pixels belonging to cup, rim, background regions, respectively. In (10) $\alpha = 1$ implies using boundary information only and $\alpha = 0$ implies using region information only.

The in-region cost functions and disc-boundary cost function are computed using a machine-learning approach as described in Sections II-B4 and II-C3 and transferred to the radial domain (Fig. 15). The two globally optimal feasible boundaries are obtained by finding a minimum closure on the vertex-weighted graph. In order to do that, the constructed graph is transferred to a closely related edge-weighted graph and finding a minimum $s - t$ cut gives us the optimal solution [22]. The α value in (10) changes over the range of 0 to 1 in increments of 0.1 and the value that produces the lowest segmentation error is selected (here $\alpha = 0.9$). The process is similar to what is outlined in [22]. Moreover, the continuity condition of the boundaries is imposed in constructing the graph by enforcing the constraint that the first and last columns of the image are neighbors.

Once the optimal boundaries are segmented, a postprocessing step is applied to remove any possible remaining errors at the blood vessel locations by fitting a smoothing spline to the optic disc and cup boundaries. The last step is to transfer the segmentation results back to Cartesian coordinates.

III. Experimental Methods

A. Data

The dataset includes 25 SD-OCT scans of glaucoma patients centered at optic nerve head that were acquired using a Cirrus HD-OCT device (Carl Zeiss Meditec, Inc., Dublin, CA) at the University of Iowa. The size of each scan was $200 \times 200 \times 1024$ voxels (in the x - y - z direction, respectively) which corresponds to a voxel size of $30 \times 30 \times 2 \mu\text{m}$, and the voxel depth was 8 bits in grayscale. The stereo color photograph pairs of the optic disc corresponding to each SD-OCT scan were taken from each patient the same day using a stereo-base Nidek 3-Dx stereo retinal camera (Nidek, Newark, NJ). The size of the stereo color photographs was 768×1019 pixels, and the pixel depth was 3 8-bit red, green and blue channels.

B. Reference Standard

The reference standard for the optic cup came from the stereo color photographs. Computer-aided planimetry was performed by three fellowship-trained glaucoma experts on stereo color photographs of the optic disc [2] and the reference standard was obtained based on consensus, meaning that the pixel that has the majority of votes (two of three) for cup region was assigned the optic cup label. The boundary of the region labeled as cup was used as the reference standard of the cup boundary in the evaluation.

To obtain the disc-boundary reference standard, one expert first traced the BMO endpoints on the 3D SD-OCT volume with two additional experts providing corrections resulting in final tracing that was the result of the consensus of three experts through a discussion. The 2D projection of the BMO endpoints produced the optic disc boundary reference standard.

C. Experiments

Three methods were compared based on the optic disc and cup segmentation performance: 1) optic disc and cup segmentation using region information only for which the in-region cost functions were produced by a classifier trained using a unimodal feature set (SD-OCT features as in [11]); 2) optic disc and cup segmentation using region information only for which the in-region cost functions were produced by a classifier trained using a multimodal feature set (fundus and SD-OCT features); and 3) optic disc and cup segmentation using region and disc-boundary information for which the in-region cost functions were produced by a classifier trained using a multimodal feature set. Note that all three approaches combine use of a machine-learning approach for designing the cost functions for use in a final graph-based step to obtain the final disc/cup boundaries simultaneously. The first two approaches only involve in-region cost function terms (with the first approach using only the SD-OCT features proposed in this work and the second approach using the region-based fundus features in addition to the SD-OCT features), whereas the last approach builds upon the second approach by also incorporating a machine-learned disc-boundary cost function term.

To evaluate the three methods for optic disc and cup segmentation, a leave-one-subject-out experiment (including the training of the classifiers) was carried out on the 25 subjects and the results were compared with the reference standard. Dice similarity coefficient (DSC), as

well as unsigned and signed border positioning errors, were used to assess the accuracy of optic disc and cup segmentation. More specifically, for two regions, A and B , the DSC was given by:

$$DSC(A, B) = \frac{2(A \cap B)}{(A + B)}. \quad (13)$$

The signed and unsigned border positioning errors of the optic disc and cup were calculated in the radial domain. The unsigned border positioning error was calculated by averaging the distances between all boundary points from the segmentation result and the ones from the reference standard. The signed border positioning error was calculated similarly to the unsigned border positioning error, but the signs of the distances were retained. The sign was considered positive if the algorithm's boundary point was farther away from the optic disc center than the boundary point of the reference standard. Paired t -tests were performed to compare the segmentation results (p -values < 0.05 were considered significant).

The proposed methods were also evaluated based on the CDR values. The mean and standard deviation of absolute difference with the reference standard as well as the Pearson correlation with the reference standard were computed.

IV. Results

An example segmentation result is shown in Fig. 16. The DSC values are shown in Table II. Based on DSC values for the optic disc segmentation, the multimodal approaches outperformed the unimodal approach ($p < 0.05$). Also, the multimodal approach with inclusion of the disc-boundary information outperformed the region-only multimodal approach for the optic disc segmentation ($p < 0.05$). For the optic cup boundary, the multimodal approach outperformed the unimodal approach ($p < 0.05$). As expected, since no cup-boundary cost function was used, both multimodal approaches resulted in the same DSC values for the cup.

The unsigned border positioning errors are reported in Table III. The unsigned border positioning error of the optic cup segmentation of the multimodal methods was significantly lower than the unsigned border positioning error of the optic cup segmentation of the unimodal method ($p < 0.05$). As with the DSC values, the unsigned border positioning errors for the cup were the same for both multimodal approaches. In addition, the unsigned border positioning error of the optic disc segmentation of the multimodal methods were significantly lower than the unsigned border positioning error of the optic disc segmentation of the unimodal method ($p < 0.05$). Moreover, the multimodal method that used the disc-boundary information outperformed the region-only multimodal method as it had a significantly lower unsigned border positioning error of the optic disc segmentation ($p < 0.05$).

The signed border positioning errors are reported in Table IV. The multimodal approaches had a significantly smaller positive bias for the optic disc segmentation than the unimodal approach ($p < 0.05$). Also, the positive bias of optic disc segmentation for the multimodal approach that used the disc-boundary information was significantly less than the region-only

multimodal method ($p < 0.05$). However, the positive bias of the optic cup segmentation for the multimodal method and the unimodal method weren't significantly different ($p > 0.1$).

The CDR correlations (Table V) of the first and second methods with the reference standard were not significantly different ($p = 0.27$). However, the correlation of the third method with the reference standard was significantly better than the correlations of the other two approaches with the reference standard ($p < 0.05$). Similarly, the absolute CDR differences (with the reference standard) of the first and second methods were not significantly different ($p = 0.15$), whereas, the third method had significantly smaller absolute CDR differences with the reference standard than that of the first and second methods.

V. Discussion and Conclusion

In this paper, we presented a unimodal (from SD-OCT volumes only) and two multimodal (from SD-OCT volumes and color fundus photographs) machine-learning graph-based approaches for automated segmentation of the optic disc and cup. Our results showed that both multimodal approaches outperformed the unimodal approach and that the multimodal approach that incorporated an additional disc-boundary cost term outperformed the multimodal approach that only incorporated region-based terms. Our multimodal approaches differ from prior work that has focused on segmenting the optic disc/cup in color fundus images alone [2]–[10] or SD-OCT images alone [11]–[13]. Another difference from prior work is our combined use of a machine-learning approach with a graph-theoretic approach to simultaneously find the globally optimal disc- and cup-boundary pair (with respect to the cost function). Thus, in addition to the novelty of the two multimodal approaches for optic disc/cup segmentation, the unimodal (SD-OCT only) approach was novel as well as prior SD-OCT-only approaches were either classification-based without a graph-based step [11], [12] or focused on only finding the projected BMO using a graph-based approach without any incorporation of machine-learned costs [13]. An additional, more subtle, difference from prior work is that our reference standard for the disc boundary was defined based on the projective location of the BMO from SD-OCT volumes rather than the fundus-based disc boundary to reflect the current understanding of the most appropriate boundary to use for SD-OCT-based definitions of optic-nerve-head structures.

While our results demonstrated the improved performance using a combined machine-learning and graph-based multimodal approach over that using a combined machine-learning and graph-based SD-OCT-only approach, our results did not explicitly demonstrate the extent the multimodal approach would provide an improvement over a fundus-only approach. However, this comparison is perhaps somewhat less interesting given our prior experience in the improved segmentation results one can obtain using SD-OCT-only information from that of fundus-only information [2], [17]. For example, in the dataset used in this work, at the intermediate classification stage (i.e., before use of the graph-based step), using the fundus-only and SD-OCT-only defined features in this work, the fundus-only approach has a significantly smaller ($p < 0.05$) pixel-based classification accuracy (overall: 86%; cup: 78%; rim: 72%; background: 89%) compared to the SD-OCT-only approach (overall: 95%; cup: 85%; rim: 83%; background: 96%). While it is possible that use of a different approach or a different set of fundus-based features could result in higher fundus-

only accuracies than those resulting from the fundus-only features used in this work, the fact that the additional inclusion of the fundus features provides an improvement over the SD-OCT-only approach is important.

Hu *et al.* previously described a multimodal approach for the segmentation of retinal vessels in color-fundus photographs and SD-OCT volumes and also demonstrated that multimodal approaches outperform SD-OCT-only approaches [21]. In that work, multimodal information was most beneficial for the segmentation of the vessels in the optic-nerve-head region. This is not surprising as the visibility of retinal vessels in SD-OCT projection images is often drastically reduced inside the disc region because of the lack of RPE tissue to offer contrast with the vessel shadows. Correspondingly, the addition of complementary color fundus information to the SD-OCT information was particularly beneficial in the disc region. Our present work is consistent with this prior work as we again find that the complementary color information from fundus photographs to be particularly helpful in the segmentation of structures (rim/cup in this case) in the disc region.

As in the multimodal work of Hu *et al.* for vessel segmentation [21], this work includes a registration step to register the fundus photograph to the SD-OCT projection image so that multimodal features can be computed at each projected location. While a modified ICP algorithm was used for registration in this work, it can be expected that other successful fundus-OCT registration approaches could be used as well. However, one potential limitation of the registration step is that any registration errors (e.g., due to motion artifacts) may limit the local accuracy of the multimodal features. While registration errors were not a problem in our dataset, it may be useful in future work to consider developing an approach to automatically assess the quality of the registration to help provide an estimate of the reliability of the multimodal features. In cases where the image registration is deemed unreliable, the approach could depend more on SD-OCT-only features.

The presence of blood vessels near the optic nerve head traditionally presents a challenge for disc/cup segmentation approaches [13], [21], [32]. The vessels in this area are wider, and their location near the optic disc/cup boundaries could potentially cause a disturbance of the local geometry and depth measurements (causing misclassification of the A-scans on the vessels used in the cost-function terms or affecting the layer segmentation locally). In this work, the use of smoothed fundus photographs in which the blood vessels are less pronounced in addition to our use of a vessel-free radial projection as a feature for computing the disc-boundary cost function, helped to minimize the impact of the blood vessels. However, development of alternative strategies for dealing with blood vessels may prove beneficial in future work.

Because of the ability of random forests to more naturally deal with low-quality or redundant features during training (compared to other candidate approaches, such as k -nearest-neighbor approaches), a feature selection stage was not incorporated for the design of the in-region and disc-boundary cost terms in this work. However, there of course is still potential value in using a smaller subset of features (e.g., a lower computation time). Thus, it may useful in future work to evaluate the accuracy with smaller subsets of features. As a start in this direction, the seven most discriminative features for the classifier used for

generating the multimodal in-region costs included the optic disc depth information, the outputs of the Gaussian filter ($\sigma = 6$) on dark-bright and ($\sigma = 8$) on red-green channels, the *a priori* probability of pixel being cup and disc, the OCT projection image, and the average of the intensities of the first SD-OCT sub-volume above the fitted BM spline surface. The three least discriminative features included the average of the intensities of the second SD-OCT sub-volume below the fitted BM spline surface and the outputs of the Gaussian filter ($\sigma = 4$) on the red and blue channels.

While we have only applied our approach to the SD-OCT volumes from a single manufacturer, it is expected that the general scheme of the proposed method (i.e., use of multimodal information and combining a machine-learning with a graph-theoretic approach) would be applicable to other SD-OCT machines as long as the B-scans in the y-direction are dense enough to allow the registration of the fundus photograph to the SD-OCT projection image. Moreover, if the B-scans are too sparse, the SD-OCT features proposed in this work may not be meaningful or be possible to extract.

In summary, we showed that we can benefit from use of complementary information from both modalities instead of unimodal information for segmenting the optic disc and cup boundaries. Different features were extracted from registered fundus photographs and SD-OCT volumes to create a multimodal feature set. A random forest classifier was trained using the multimodal feature set to produce three in-region cost functions associated with cup, rim, and background regions. A set of SWT-based features were extracted from SD-OCT volumes to be utilized in training a random forest classifier for producing the disc-boundary cost function. The graph-theoretic approach was able to well-define the boundaries of interest using the cost functions and enabled us to impose the desired constraints to the boundaries. Use of a similar multimodal graph-based strategy is expected to be beneficial in other application areas as well.

Acknowledgments

This work was supported, in part, by the Department of Veterans Affairs Rehabilitation Research and Development Division (Career Development Award IK2RX000728); the National Institutes of Health (R01 EY018853 and R01 EY023279); and the Marlene S. and Leonard A. Hadley Glaucoma Research Fund.

References

1. Abramoff MD, Garvin MK, Sonka M. Retinal imaging and image analysis. *IEEE Rev. Biomed. Eng.* 2010; 3:169–208. [PubMed: 22275207]
2. Abramoff MD, Alward WLM, Greenlee EC, Shuba L, Kim CY, Fingert JH, Kwon YH. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. *Invest. Ophthalmol. Vis. Sci.* 2007 Apr;48(4):1665–1673. [PubMed: 17389498]
3. Cheng J, Liu J, Xu Y, Yin F, Wong DWK, Tan N-M, Tao D, Cheng C-Y, Aung T, Wong TY. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Trans. Med. Imag.* 2013 Jun; 32(6):1019–1032.
4. Xu J, Chutatape O, Sung E, Zheng C, Kuan PCT. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. *Patt. Rec.* 2007 Jul; 40(7):2063–2076.
5. Yin, F.; Liu, J.; Ong, SH.; Sun, Y.; Wong, DWK.; Tan, NM.; Cheung, C.; Baskaran, M.; Aung, T.; Wong, TY. Model-based optic nerve head segmentation on retinal fundus images; 33rd Ann. Int. Conf. IEEE EMBS; Aug. 2011; Boston, Massachusetts USA. p. 2626-2629.

6. Lowell J, Hunter A, Steel D, Basu A, Ryder R, Fletcher E, Kennedy L. Optic nerve head segmentation. *IEEE Trans. Med. Imag.* 2008 Feb.23(2):256–264.
7. Chràsteka R, Wolfa M, Donatha K, Niemann H, Paulusb D, Hothorne T, Lausenc B, Lmmerd R, Mardind C, Michelsond G. Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Med. Imag. Anal.* 2005 Aug.9(4):297–314.
8. M MB Jr, Wu X, Sonka M, Abràmov MD. Optimal segmentation of the optic nerve head from stereo retinal images. in *Proc. SPIE, Med. Imag. 2006: Phys. Func. Struc. Med. Imag. 2006; 6143:61433B.* (8 pages).
9. M MB Jr, Abràmov MD, Sonka M, Wu X. Segmentation of the optic nerve head combining pixel classification and graph search. *Proc. SPIE, Med. Imag. 2007: Imag. Proc. 2007; 6512:651215.* (10 pages).
10. Zheng, Y.; Stambolian, D.; O'Brien, J.; Gee, JC. *Proc. MICCAI 2013, LNCS, Part II. Vol. 8150. Nagoya, Japan: 2013 Sept.. Optic disc and cup segmentation from color fundus photograph using graph cut with priors; p. 75-82.*
11. Lee K, Niemeijer M, Garvin MK, Kwon YH, Sonka M, Abràmov MD. Segmentation of the optic disc in 3-D OCT scans of the optic nerve head. *IEEE Trans. Med. Imag.* 2010 Jan.29(1):159–168.
12. Abràmov MD, Lee K, Niemeijer M, Alward WLM, Greenlee EC, Garvin MK, Sonka M, Kwon YH. Automated segmentation of the cup and rim from spectral domain oct of the optic nerve head. *Invest. Ophthalmol. Vis. Sci.* 2009 Dec.50(12):5778–5784. [PubMed: 19608531]
13. Hu Z, Abràmov MD, Kwon YH, Lee K, Garvin MK. Automated segmentation of neural canal opening and optic cup in 3D spectral optical coherence tomography volumes of the optic nerve head. *Invest. Ophthalmol. Vis. Sci.* 2010 Nov.51(11):5708–5717. [PubMed: 20554616]
14. Chauhana BC, Burgoyne CF. From clinical examination of the optic disc to clinical assessment of the optic nerve head: A paradigm change. *Am. J. Ophthalmol.* 2013 Aug.156(2):218–227. [PubMed: 23768651]
15. Chauhan BC, O'Leary N, AlMobarak FA, Reis AS, Yang H, Sharpe GP, Hutchison DM, Nicolela MT, Burgoyne CF. Enhanced detection of open-angle glaucoma with an anatomically accurate optical coherence tomography derived neuroretinal rim parameter. *Ophth.* 2013 Mar.120(3):535–543.
16. Reis AS, Sharpe GP, Yang H, Nicolela MT, Burgoyne CF, Chauhan BC. Optic disc margin anatomy in patients with glaucoma and normal controls with spectral domain optical coherence tomography. *Ophth.* 2012 Apr.119(4):738–747.
17. Miri, MS.; Lee, K.; Niemeijer, M.; Abràmov, MD.; Kwon, YH.; Garvin, MK. *Proc. SPIE, Med. Imag. 2013: Imag. Proc. Vol. 8669. Orlando, Florida: 2013 Mar.. Multimodal segmentation of optic disc and cup from stereo fundus and SD-OCT images; p. 866900(9pages).*
18. Kolar R, Tasevsky P. Registration of 3D retinal optical coherence tomography data and 2D fundus images. in *Proc. WBIR 2010, LNCS. 2010; 6204:72–82.*
19. Li Y, Gregori G, Knighton RW, Lujan BJ, Rosenfeld PJ. Registration of oct fundus images with color fundus photographs based on blood vessel ridges. *Opt. Express.* 2011 Jan.19(1):7–16. [PubMed: 21263537]
20. Golabbakhsh M, Rabbani H. Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model. *IET Imag. Proc.* 2013; 7(8):768–776.
21. Hu Z, Niemeijer M, Abràmov MD, Garvin MK. Multimodal retinal vessel segmentation from spectral-domain optical coherence tomography and fundus photography. *IEEE Trans. Med. Imag.* 2012 Oct.31(10):1900–1911.
22. Garvin MK, Abràmov MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans. Med. Imag.* 2009 Sept.28(9):1436–1447.
23. Niemeijer M, Staal J, van Ginneken B, Loog M, Abràmov MD. Comparative study of retinal vessel segmentation methods on a new publicly available database. *Proc. SPIE, Med. Imag. 2004: Imag. Proc. 2004; 5370:648–656.*

24. Niemeijer M, Garvin MK, van Ginneken B, Sonka M, Abràmoff MD. Vessel segmentation in 3-D spectral OCT scans of the retina. *Proc. SPIE, Med. Imag. 2008: Imag. Proc. 2008; 6914:69141R.* (8 pages).
25. Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans. Patt. Anal. Mach. Intell.* 1992 Feb.12(14):239–256.
26. Corana A, Marchesi M, Martini C, Ridella S. Minimizing multimodal functions of continuous variables with the “simulated annealing” algorithm. *ACM Trans. Math. Software.* 1987; 13(3): 262–280.
27. Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Comput. J.* 1964; 7(2):155–162.
28. Niemeijer M, Garvin MK, L K, Abràmoff MD, Sonka M. Registration of 3-D spectral OCT volumes combining ICP with a graph-based approach. *Proc. SPIE, Med. Imag. 2012: Imag. Proc. 2012; 8314:83141A.* (9 pages).
29. Breiman L. Random forests. *Mach. Learn.* 2001; 45(1):5–32.
30. Schwarz MW, Cowan WB, Beatty JC. An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Trans. Graphics.* 1987 Apr.6(2):123–158.
31. Fowler JE. The redundant discrete wavelet transform and additive noise. *IEEE Sig. Proc. Let.* 2005 Sept.12(9):629–632.
32. Hu Z, Niemeijer M, Abràmoff MD, Lee K, Garvin MK. Automated segmentation of 3-D spectral oct retinal blood vessels by neural canal opening false positive suppression. *Proc. MICCAI 2010, LNCS, Part III.* 2010; 6363:33–40.

Figure 1. Cupping causes an increase in cup-to-disc ratio (CDR). (a) CDR = 0.15, (b) CDR = 0.46, (c) CDR = 0.84.

(a)

(b)

(c)

(d)

Figure 2.

Bruch's membrane opening (BMO) within an SD-OCT volume. (a) An SD-OCT B-scan with BMO points marked with two filled circles. RPE = retinal pigment epithelium. (b) 3D view of all BMO points for the entire SD-OCT volume. (c) SD-OCT projection image. (d) Projected view of BMO points on SD-OCT projection image.

Figure 3.
Flowchart of overall method.

Figure 4.

An example of a central B-scan of an SD-OCT volume and intraretinal layer segmentation. (a) The original OCT B-scan. (b) Segmented three layers: the first surface is the ILM (red), the second surface (yellow) is the IS/OS junction, and the third surface (blue) is the lower bound of RPE complex. The pink surface indicates the thin-plate spline fitted to the third surface. (c) Flattened OCT B-scan along with surfaces. (d) 3D view of the three surfaces.

Figure 5. Registration of fundus photograph to projection image of SD-OCT volume. (a) Original fundus photograph. (b) Corresponding projection image of SD-OCT volume. (c) Alignment of fundus photograph to OCT projection image. (d) Registered fundus photograph.

Figure 6.

Radial slice segmentation. (a) Example (non-central) scan from original SD-OCT volume demonstrating how the BMO points may appear close together. (b) Example radial scan (with BMO points appearing as they would in a central scan from the original volume) with segmented surfaces. Interpolation is used to define the second and third surfaces in the neural canal region. (c) The radial projection image.

Figure 7.
In-region cost function design flowchart.

Figure 8. OCT features. (a)–(f) Average intensity of subvolumes in z-direction. (g) Distance of first surface to the thin-plate spline fitted to the third surface. (h) The SD-OCT projection image.

Figure 9.

Fundus pixel features. From left to right are the filtered image using Gaussian filter bank having sizes $\sigma=4$, 6, 8 respectively. (a)–(c) Red channel. (d)–(f) Green channel. (g)–(i) Blue channel. (j)–(l) Dark-bright channel. (m)–(o) Blue-yellow channel. (p)–(r) Red-green channel.

Figure 10.

Spatial features. (a)–(c) Three *a priori* maps corresponding to cup, rim, and optic disc regions derived from PCA. (d) Distance of the x position with respect to the optic disc center. (e) Distance of the y position with respect to the optic disc center. (f) Radial distance with respect to the optic disc center.

Figure 11.
Disc-boundary cost function design flowchart.

Figure 12.

SWT decomposition. (a)–(d) SWT 6-level decomposition. In each image, upper left is the approximation, upper right is the horizontal, lower left is the vertical, and lower right is the diagonal SWT coefficient.

Figure 13.

On-boundary cost function feature set. (a)–(d) Horizontal coefficients of level 1, 2, 3 and 4.

(e) Vessel-free projection image. (f) Result of averaging derivative of Gaussian in the vertical direction. (g) Spatial feature imposing the shape of the optic disc boundary. (h) Spatial feature that has the anatomic information of BMO points.

Figure 14.

A schematic representation of segmenting the optic disc and cup boundaries using a theoretical graph based approach. (a) Left is the original image and right is the resampled image in radial domain. (b) Example cost of two boundaries for the multiple boundary segmentation problem. The two boundaries divide the images into three regions.

Figure 15.

An example of cost functions. (a) The in-region cost function for the background. (b) The in-region cost function for the rim. (c) The in-region cost function for the optic cup. (d) The optic-disc-boundary cost function. Note that there is no cup-boundary cost function.

Figure 16.

An example segmentation result. The first row contains the (a) registered fundus photograph with (b) the reference standard boundaries, (c) the boundaries of the first method (using unimodal region costs), (c) the boundaries of the second method (using multimodal region costs), and (d) the boundaries of the third method (using multimodal region plus disc-boundary costs). The blue boundary corresponds to the optic disc boundary and the green boundary corresponds to the cup boundary. The second row contains the boundaries of the methods shown on the SD-OCT projection image. The third row contains a central B-scan of the SD-OCT with green indicating the rim region and red indicating the cup region from the different methods. The last row contains the region-based segmentation results (black = background; gray = rim; white = cup). It is especially noticeable on the inferior (I) and temporal (T) sides of optic disc boundary that the third method has the closest boundary to

the reference standard. In addition, the unimodal approach has a relatively smaller optic cup than the multimodal approach in comparison with the reference standard.

TABLE I

Multimodal feature set used for producing in-region cost functions

Features	Fundus	SD-OCT	Spatial
Features 1–18: Filtered color channels (red, green, blue, dark-bright, yellow-green, and red-blue) using Gaussian zero-order filter bank ($\sigma=4,6,8$).	✓		
Feature 19: The pixel intensity in the SD-OCT projection image. The SD-OCT projection image is produced by averaging the voxel intensities between surface two (IS/OS junction) and three (the outer boundary of the RPE) in the z -direction.		✓	
Features 20–25: Average intensities of six subvolumes in the z -direction (120 pixels above and 60 pixels below the fitted reference spline to surface three), each of them representing the average voxel intensity of 30 voxels within the voxel column.		✓	
Features 26–28: Pixel <i>a priori</i> expectation map of each class based on pixel location.			✓
Features 29–31: Distance of x and y positions as well as the radial distance of each pixel with respect to the approximate center.			✓

TABLE IIDice Similarity Coefficients (Mean $\pm$ SD)

Methods	Cup	Disc
in-region (OCT)	0.851 $\pm$ 0.031	0.735 $\pm$ 0.102
in-region (OCT+Fundus)	0.855 $\pm$ 0.009	0.781 $\pm$ 0.095
in-region + disc-boundary	0.855 $\pm$ 0.009	0.824 $\pm$ 0.086

TABLE IIIUnsigned Border Positioning Error (Mean $\pm$ SD)

Methods	Cup (pixel)	Cup (mm)	Disc (pixel)	Disc (mm)
in-region (OCT)	2.520 $\pm$ 0.723	0.075 $\pm$ 0.021	2.053 $\pm$ 0.731	0.061 $\pm$ 0.021
in-region (OCT+Fundus)	2.262 $\pm$ 0.621	0.067 $\pm$ 0.018	1.971 $\pm$ 0.694	0.059 $\pm$ 0.020
in-region + disc-boundary	2.262 $\pm$ 0.621	0.067 $\pm$ 0.018	1.602 $\pm$ 0.587	0.048 $\pm$ 0.017

TABLE IVSigned Border Positioning Error (Mean $\pm$ SD)

Methods	Cup (pixel)	Cup (mm)	Disc (pixel)	Disc (mm)
in-region (OCT)	0.015 $\pm$ 1.08	0.000 $\pm$ 0.032	0.351 $\pm$ 1.03	0.010 $\pm$ 0.030
in-region (OCT+Fundus)	0.017 $\pm$ 1.01	0.000 $\pm$ 0.029	0.171 $\pm$ 0.87	0.005 $\pm$ 0.025
in-region + disc-boundary	0.017 $\pm$ 1.01	0.000 $\pm$ 0.029	0.019 $\pm$ 0.61	0.001 $\pm$ 0.017

TABLE V

CDR analysis. Pearson correlation (r) with reference standard and the absolute difference from the reference standard (Mean $\pm$ SD)

Methods	Correlation	Difference
in-region (OCT)	0.890	0.101 $\pm$ 0.082
in-region (OCT+Fundus)	0.899	0.097 $\pm$ 0.076
in-region + disc-boundary	0.928	0.082 $\pm$ 0.048