Automatic Optic Disc Detection in OCT Slices via Low-Rank Reconstruction

Huazhu Fu, Dong Xu, Stephen Lin, Damon Wing Kee Wong, Jiang Liu

Abstract—Optic disc measurements provide useful diagnostic information as they have correlations with certain eye diseases. In this paper, we provide an automatic method for detecting the optic disc in a single OCT slice. Our method is developed from the observation that the retinal pigment epithelium (RPE) which bounds the optic disc has a low-rank appearance structure that differs from areas within the disc. To detect the disc, our method acquires from the OCT image an RPE appearance model that is specific to the patient and imaging conditions, by learning a low-rank dictionary from image areas known to be part of the RPE according to priors on ocular anatomy. The edge of the RPE, where the optic disc is located, is then found by traversing the retinal layer containing the RPE, reconstructing local appearance with the low-rank model, and detecting the point at which appearance starts to deviate (i.e., increased reconstruction error). To aid in this detection, we also introduce a geometrical constraint called the distance bias that accounts for the smooth shape of the RPE. Experiments demonstrate that our method outperforms other OCT techniques in localizing the optic disc and estimating disc width. Moreover, we also show the potential usage of our method on optic disc area detection in 3D OCT volumes.

Index Terms—optic disc detection, layer segmentation, optical coherence tomography.

I. INTRODUCTION

The location where ganglion cell axons exit the eye to form the optic nerve is called the optic disc [1], [2]. Since the measurement of the optic disc is important for many medical applications such as glaucoma screening [3] and large exudative lesion analysis [4], there has been much recent effort on automatically detecting the optic disc in ocular images. Many existing optic disc detection methods focus on segmenting the optic disc region in fundus images [5]–[8]. For example, Xu et al. [9] employed the deformable model technique through minimization of an energy function to detect the disc. Cheng et al. [8] considered optic disc detection as a superpixel classification problem based on center-surround statistics. The method of Morales et al. [10] extracts the optic disc contour based on mathematical morphology and principal components analysis. However, a major problem of these methods is that they easily fail when the optic disc does not have a distinct color in the fundus image.

This work is supported by the Singapore A*STAR SERC Grant (112-148-0003). H. Fu and D. Xu are with School of Computer Engineering, Nanyang Technological University, Singapore. (E-mail: {hzu, dongxu}@ntu.edu.sg). S. Lin is with Microsoft Research, Beijing, China. (E-mail: stevelin@microsoft.com). D. W. K. Wong and J. Liu are with iMED Ocular Imaging Programme, Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore. (e-mail: {wkong, jliu}@i2r.a-star.edu.sg).

A relatively new imaging technique called Optical Coherence Tomography (OCT) provides a clearer view of intraretinal morphology and enables non-invasive depth-resolved functional imaging of the retina [11]. In an OCT slice, the boundary of the optic disc appears at the end of the Retinal Pigment Epithelium (RPE) [12]–[14], as shown in Fig. 1. Some OCT-based optic disc detection methods operate on 3D OCT volumes [15]–[19]. For example, Lee et al. [18] extracts intraretinal surfaces from 3D OCT volumes, and classifies the optic disc, cup and neuroretinal rim using a k-NN classifier. Mori et al. [19] provided a multimodal pixel-classification approach to segment the optic disc by combining information from stereo fundus images and an OCT volume. However, acquisition of numerous OCT slices for 3D reconstruction is susceptible to misalignment error due to eye movement. Although 3D volume scanning acquisition is becoming more common with advances in OCT technology, examination of a single B-scan for optic disc detection remains an important problem in practice for routine clinical applications. In the state-of-the-art method for optic disc detection from a single OCT slice, Boyer et al. [20] extracted Hadamard transform features along a retinal layer containing the RPE and grouped the features into two clusters to identify the optic disc endpoints. However, feature-based clustering can be sensitive to appearance variations due to noise and blood vessel shadows in the OCT image. Moreover, the method in [20] does not take advantage of the smooth geometric structure of the RPE in localizing the optic disc endpoints.

In this paper, we present a general technique for optic disc detection in a single OCT slice via low-rank reconstruction. Based on retinal structure priors, we divide the pixels along the RPE layer boundary into training data and optic disc edge
Our approach first segments the retinal layers and divides the region into a training region and a candidate region. The red and cyan curves denote the boundaries of ILM layer and the segmented RPE layer, respectively. Then the low-rank dictionary is learned from the training data. We next reconstruct the layer appearance at points along the candidate region, and output the reconstruction errors, including intensity, LBP, and distance bias features. Finally, the optic disc boundary points are detected by combining these three error curves.

The OCT slices in our work are obtained by using a TOP-CON DRI OCT-1 Atlantis swept source OCT imaging device, which has a depth resolution of 8μm and lateral resolution of 10μm. A slice consists of 1024 A-scans (columns) each with 992 pixels.

A. Retinal layer boundary segmentation

While there exist methods for extracting multiple retinal layers from OCT slices [21]–[23], only the RPE and ILM layer boundaries are needed in our work. These two layers have specific characteristics that distinguish them from the surrounding areas and layers. We can thus employ a fast and simple method to extract the ILM and RPE layer boundaries from an OCT slice.

ILM layer boundary segmentation: The ILM is defined as the boundary between the retina and the vitreous body, which is the first boundary in the OCT slice as shown in Fig. 1. After reducing the speckle noise of the OCT slice with a 2D Gaussian filter, we threshold the OCT slice by using Otsu’s
method [24], as shown in Fig. 3 (b). Then, the topmost binary edge is selected as the ILM layer, as indicated by the red boundary in Fig. 3 (d). The lowest point on the ILM layer boundary is used to approximately localize the ONH center.

RPE layer boundary segmentation: In general, the two strongest gradient edges in the vertical direction of each column of the OCT slice corresponds to the ILM (upper) and RPE (lower) layer boundaries, as shown in Fig. 3 (c). After extracting the ILM layer boundary, we can take the strongest gradient edge below the ILM layer as the RPE layer boundary. To elevate the robustness of the RPE layer boundary segmentation, our method employs three steps. First, a 2D Gaussian filter is used to reduce speckle noise in the OCT slice. Then we select one point in each column of the OCT slice that is below the ILM layer and has the strongest positive vertical gradient. Finally, we connect these points as the RPE layer boundary, and employ a one-dimensional median filter [25] to smooth its vertical position in each column. Note that the extracted RPE layer extends across the entire OCT slice, including the non-RPE part, as shown by the cyan curve in Fig. 3 (d). Thus, we will use the extracted RPE layer boundary as a starting point along which we will identify the true endpoint of the RPE, and hence locate the boundary of the optic disc.

ONH localization: In our method, the ONH is employed as a landmark to define the training region and the candidate region. Generally, the lowest point on the ILM layer boundary could be used to locate the ONH. However, some OCT slices may exhibit serious sloping of the retina, such that the lowest ILM point is not in the ONH, as shown in Fig. 4 (a). For these OCT slices, we employ a preprocessing step to correct this sloping. First, we compute the inclination $\theta$ of the retina in the OCT slice from the edge points of the ILM layer, e.g. points A and B in Fig. 4 (a). Then the corrected coordinate $(\hat{x}, \hat{y})$ for the point at $(x, y)$ is obtained through a rotation with a fixed $y$ value:

$$\hat{x} = y \tan(\theta) + x,$$

$$\hat{y} = y.$$

Fig. 4 (b) shows the corrected OCT slice, where the lowest ILM point is located in the ONH.

B. Low-rank dictionary learning

The optic disc boundary points are defined as the left/right endpoints of the RPE, which are each detected separately in our work. It has been observed that the regions along the RPE layer boundary, including the OPL and HRC as shown in Fig. 1, have a consistent retinal structure that includes multiple surrounding bands, while the regions beyond the RPE endpoints are different [26]. This observation motivates our use of low-rank reconstruction for detecting RPE endpoints. Toward this end, we construct a low-rank dictionary from the training region to model RPE appearance, and then use the low-rank dictionary to detect the optic disc point in the candidate region based on changes in low-rank reconstruction error.

To divide the segmented layer into the training and candidate regions, we make use of prior knowledge that the optic disc is approximately centered on the ONH and has a vertical and horizontal diameter of about 1.92 ± 0.29 mm and 1.76 ± 0.31 mm [27], [28]. Based on this, we define a loose candidate region as having a 360-pixel width (corresponding to about 2.8 mm) centered horizontally on the lowest ILM point. The remaining part of the segmented layer is taken as the training region. We denote the width of the training region as $M$. As the training data is derived from the input image itself, the learned model is specifically tailored to the individual and the imaging apparatus used to capture the OCT image. A set of labeled training images is not required.

We extract appearance features along each pixel of the training region, and arrange them into a feature matrix $X_t$. Because of the consistent retinal structure along the RPE layer boundary, the feature matrix $X_t$ can be decomposed into a low-rank matrix $\hat{D}$ and sparse errors $E_1$ by using robust principal components analysis (RPCA) [29]:

$$\hat{D} = \arg \min_D \| D \|_*, \lambda_1 \| E_1 \|_1,$$

$$s.t. \ X_t = \hat{D} + E_1,$$

where $\| \cdot \|_*$ denotes the nuclear norm, $\| \cdot \|_1$ denotes the $\ell_1$ norm, and $X_t$ is the feature matrix which is decomposed into the low-rank dictionary $\hat{D}$ and sparse error matrix $E_1$.

C. Low-rank reconstruction

With the learned low-rank dictionary $\hat{D}$, the feature matrix of the candidate region will be reconstructed to obtain the reconstruction error. Similar to low-rank dictionary learning, we also employ the low rank property to constrain the reconstruction based on low-rank representation (LRR) [30]:

$$\hat{E}_2 = \arg \min_{Z, E_2} \| Z \|_* + \lambda_2 \| E_2 \|_{21},$$

$$s.t. \ X_c = \hat{D}Z + E_2,$$

where $X_c$ denotes the feature matrix of the candidate region, which is decomposed into the low rank dictionary $\hat{D}$, low-rank representation $Z$, and sparse error matrix $E_2$. $\| \cdot \|_{21}$ is
We normalize the feature data not reconstructed in the low-rank representation, algorithm [31] is employed to solve Eqs. (2) and (3).

vessel shadows. The Augmented Lagrange Multiplier (ALM) are sample-specific, i.e., some data vectors are corrupted by 

\( \ell_2,1 \) norm, which encourages the columns of \( \mathbf{E}_2 \) to be zero. The underlying assumption here is that image corruptions are sample-specific, i.e., some data vectors are corrupted by vessel shadows. The Augmented Lagrange Multiplier (ALM) algorithm [31] is employed to solve Eqs. (2) and (3).

The error matrix \( \mathbf{E}_2 \) corresponds to components of the feature data not reconstructed in the low-rank representation, and has the same dimensions as the candidate feature matrix. We normalize the \( \mathbf{E}_2 \) by using the Frobenius Norm [32], and then sum each column of \( \mathbf{E}_2 \) to generate the low-rank reconstruction error curve. In this error curve, points belonging to the RPE will be reconstructed well and have low values, while points outside the RPE will have high errors.

**D. Features and distance bias**

Technically, any feature could be employed in our method to describe pixels in the RPE layer. However, features whose range can cover the OPL and HRC are preferable, since the OPL and HRC are stable and consistent along the RPE layer and are thus helpful for determining the endpoint of the RPE layer. In our method, we employ two types of features: intensity values and local binary patterns (LBP) [33]. The intensity values are specifically a vector of pixel intensities in the OCT image from 20 rows below each point on the segmented boundary to 60 rows above it. Together, the intensity vectors form an intensity feature matrix, illustrated in Fig. 5 (b). The second feature, LBP, is widely-used for many applications. In this work, we use a \( 10 \times 5 \) block within a \( 80 \times 10 \) window around each extracted RPE layer boundary point, with the point located 25% from the bottom as shown within the red box in Fig. 5 (a). The extracted LBP feature vectors are collected into an LBP feature matrix.

Moreover, we also introduce a geometrical constraint called distance bias in our method. The RPE is identified as a smooth, convex surface composed of a single layer of hexagonal cells that help to maintain the integrity of the barrier between the choroid and the retina [34]. We empirically found that the smooth shape of the RPE can be well approximated by a quadratic curve. A quadratic curve is thus fitted to the RPE layer boundary by linear least squares in the training region and used to constrain the endpoint position in the candidate region. Based on this constraint, we define distance bias as the vertical distance between the fitted curve and the RPE layer boundary in the candidate region, as illustrated in Fig. 5 (a) by yellow bars. From the distance bias, we obtain a geometrical error curve, as shown in Fig. 5 (c).

**E. Optic disc localization**

The error curves for intensity, LBP, and distance bias are combined by normalizing the curves and computing their sum. By treating the three factors separately until now, the issue of normalizing different features within a single feature matrix has been avoided. In the aggregated error curve, points that belong to the RPE layer should have low error, while other points within the optic disc will have higher errors. Between the corresponding two parts of the segmented layer the optic disc point should lie. To identify this point, we fit a sigmoid function to the error curve and use its midpoint to locate the optic disc boundary as shown in Fig. 2 (d). We have found this approach to give an accurate estimate of the optic disc position.

III. EXPERIMENTS

For testing our method, we collected OCT slices centered at the ONH from 20 normal persons, four of whom were selected randomly for re-capture of their OCT slices after a long time interval (more than six months). This dataset thus consists of 48 OCT slices in total. These slices were captured by using the 2D imaging protocol, at a 1024\( \times \)992 image resolution and with a depth resolution of \( 8 \mu \text{m} \)/pixel and lateral resolution of \( 10 \mu \text{m} \)/pixel. A trained labeler marked the ground-truth optic disc points manually in each of the images, and two experts examined the ground-truth labelings for quality control. We set parameter \( \lambda_1 = 0.35 \) in RPCA, parameter \( \lambda_2 = 0.45 \) in LRR, and dictionary size \( M = 140 \) as the default parameters for all the experiments.

Our evaluation employs two error metrics. The first is the distance error in terms of image columns between the detected optic disc and the ground-truth: \( m_d = |C_d - C_{gt}| \), where \( C_d \) and \( C_{gt} \) denote the column coordinates of the detected and
A. Optic disc detection performance

We evaluated the optic disc detection performance of our method through comparisons to simplified versions of our method with only a subset of the features and to other techniques: (1) each of the three features alone (‘Intensity’, ‘LBP’, and ‘Distance bias’), where the final localization curve is generated by using only one error curve; (2) a combination of only two features (‘Intensity + LBP’, ‘Intensity + Distance’, and ‘LBP + Distance’); (3) our full method without dictionary learning (‘w/o Dictionary’), i.e., directly using the candidate data itself as a self-expressive dictionary [30] for the low-rank representation in Eq. 3; (4) the existing method for optic disc detection in a single OCT slice (‘Boyer [20]’); (5) a baseline method that detects the optic disc based on the average intensity of the RPE layer boundary (‘Baseline 2D’), which takes advantage of the fairly large contrast that appears at the endpoint of the RPE layer. Table I shows the performance (average ± standard deviations) of the various optic disc detection methods.

In Table I, it can be seen that the combination of intensity and LBP (‘Intensity+LBP’) generally outperforms each of the features individually. Adding the distance bias feature to LBP leads to an improvement of about 12% in the width error ratio $m_w$, which indicates the benefit of accounting for distance bias. In the table, the results of our method without dictionary learning (‘w/o Dictionary’) are also reported, in which the candidate data itself serves as the dictionary for the low-rank representation. Since the candidate data contains regions both with and without the RPE layer, the dictionary learned from it does not have the low-rank property. As a result, reconstruction error with the dictionary provides only a weak indicator of the endpoint of the RPE layer. The low-rank dictionary learning step in the proposed method makes the reconstruction-based approach more robust and effective by helping to remove spurious outliers such as heavy noise and vessel shadows from the RPE model. This is different from the baseline techniques ‘Boyer [20]’ and ‘Baseline 2D’ which directly cluster pixels via extracted features that can be distorted by outlier elements.

Fig. 6 displays some detection results, where the red, green and yellow lines indicate our results, those of ‘Boyer [20]’ and the ground-truth, respectively. Our method generally outperforms ‘Boyer [20]’. In the figure, the blue dashed lines mark the RPE layer segmented according to Gaussian filter responses. The segmentations include the RPE as well as non-RPE extensions into the optic disc. In many cases, the geometrical constraint from distance bias provides a useful detection cue. However, for some cases, non-RPE points in the layer may satisfy the distance bias constraint. Our detection relies more on intensity and LBP features in these cases.

We performed our experiments using a PC with a 3.2GHz CPU and 16GB RAM. The code is implemented in Matlab without optimization. Fig. 7 shows the computation time (in seconds) of our entire method (red curve) and the main steps: feature extraction (green), dictionary learning via RPCA (blue), and error reconstruction via LRR (black). Feature extraction and dictionary learning require the most computation in our method. The computation time, especially the dictionary learning component, increases consistently with dictionary size $M$. Typically, our method takes about 50 seconds for an OCT slice of image resolution $1024 \times 992$.

B. Optic disc area detection in the OCT volume

Although we focus on optic disc detection from a single OCT slice, our method can be extended to handle an OCT volume. For a 3D OCT volume, we first find the OCT slices that cross the ONH, determined by the height of the ILM layer within the training region as shown in Fig. 8 (a). We employ a height threshold $t$ ($t = 100$ in our experiment) to select the OCT slices, and then detect the optic disc in each single OCT slice by using our method. Finally, a fitted ellipse [35] is computed as the disc boundary on the OCT fundus image, as shown in Fig. 8 (c).

For testing our method on OCT volumes, we collect seven OCT volumes (each of $992 \times 512 \times 512$ resolution) by using the 3D imaging protocol. The ground-truth optic disc boundary of a 3D OCT volume is obtained by first manually labeling the optic disc points in each ONH-centered slice in the same manner as that for 2D OCT slice labeling (with a trained labeler and two experts for quality control). These labeled points are then fit with an ellipse to generate the ground-truth optic disc boundary.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>THE PERFORMANCES (AVERAGE ± STANDARD DEVIATIONS) OF VARIOUS OPTIC DISC DETECTION METHODS.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$m_d$ (pixel)</td>
</tr>
<tr>
<td>Baseline 2D</td>
<td>25.8 ± 23.1</td>
</tr>
<tr>
<td>Boyer [20]</td>
<td>20.7 ± 19.8</td>
</tr>
<tr>
<td>Intensity</td>
<td>22.2 ± 20.8</td>
</tr>
<tr>
<td>LBP</td>
<td>24.4 ± 17.7</td>
</tr>
<tr>
<td>Distance bias</td>
<td>21.4 ± 18.4</td>
</tr>
<tr>
<td>Intensity + LBP</td>
<td>18.6 ± 17.7</td>
</tr>
<tr>
<td>Intensity + Distance</td>
<td>16.1 ± 14.5</td>
</tr>
<tr>
<td>LBP + Distance</td>
<td>14.8 ± 12.9</td>
</tr>
<tr>
<td>w/o Dictionary</td>
<td>18.1 ± 15.9</td>
</tr>
<tr>
<td>Our method</td>
<td>12.4 ± 12.1</td>
</tr>
</tbody>
</table>
Fig. 6. Optic disc detection results of our method (only a rectangular region around the ONH is shown). The cyan dashed curve denotes the segmented RPE layer boundary, and the red, green and yellow lines indicate the detection locations for our method, Boyer [20], and ground-truth, respectively.

We compare our method with four other techniques. The first two, namely ‘Boyer [20]’ and ‘Baseline 2D’, use detection methods for single OCT slices and then fit an ellipse to the detected points to find the detected optic disc area in the 3D OCT volume, similar to our method. The third method is based on Lee et al. [18], which operates on 3D OCT volumes. Note that Lee et al. [18] involves kNN classification, which is not applicable to our unsupervised setting without labeled training data. Thus we modify the last step in Lee et al. [18] to make it unsupervised. We extract the RPE surface from the 3D OCT volume and flatten the surface. Then a binary clustering method is applied based on the features (e.g. intensity and gradient) extracted from the RPE surface to determine the optic disc region. We refer to this modification of Lee et al. [18] as ‘Baseline 3D’. The fourth comparison method is our implementation of Ishikawa et al. [17], which detects the optic disc area in a 3D OCT volume by using the modified active contour model [5].

For OCT volumes, we employ two evaluation criteria to measure disc region detection accuracy. The first measure is the non-overlap ratio:

$$m_1 = 1 - \frac{\text{Area}(R \cap R_{GT})}{\text{Area}(R \cup R_{GT})},$$

where $R$ and $R_{GT}$ denote the detected optic disc region and the ground-truth ellipse, respectively. The second measure is the relative absolute area difference:

$$m_2 = \frac{|\text{Area}(R) - \text{Area}(R_{GT})|}{\text{Area}(R_{GT})}.$$
Fig. 9. Optic disc detection results on OCT volumes (only a rectangular region around the ONH is shown). The top row shows disc area detection results on OCT fundus images, and the bottom row displays the corresponding OCT slice results on the ONH. The red, green, blue and yellow lines indicate the detection results of our method, Ishikawa et al. [17], Baseline 3D, and the ground-truth, respectively.

poor performance on single OCT slices leads to distorted disc regions. ‘Boyer [20]’ and ‘Ishikawa [17]’ obtain similar performance. ‘Baseline 3D’ generates stable and regular disc regions, but it depends on accurate OCT surface flattening to reduce the influence of fore-aft eye movement. Furthermore, it tends to produce disc sizes smaller than the ground-truth, such as the blue curves in the third column of Fig. 9. In contrast, by inheriting the advantages of processing single OCT slices, our method detects the endpoints of the RPE layer more accurately through low-rank reconstruction, and outperforms the other methods. Moreover, in contrast to a single OCT slice, an OCT volume with a fitted ellipse essentially provides a smooth neighborhood constraint, which is beneficial for removing outliers introduced by disc detection errors on a small number of OCT slices.

C. Discussion

In this work, we present a technique for optic disc detection without also addressing the problem of optic cup detection, though it is also needed for computing the cup-to-disc ratio (CDR), the most commonly used clinical feature in glaucoma diagnosis. While investigating optic cup detection in a similar fashion would be an interesting and important study, optic disc detection nevertheless remains an important problem itself as it is often used to support other detection and assessment tasks. Fully/semi-automated quantitative disc assessment using ocular imaging devices (for fundus images and OCT volumes) usually starts with detecting the optic disc margin [13]. The optic disc margin also provides a fundamental landmark for detecting other retinal parts. For example, optic cup detection is generally performed based on an assumption that cupping occurs only within the disc area [7], [8]. Typically in an OCT slice, the optic cup diameter is defined as the length of the line that connects the outermost borders of the cup at the level of 150 µm above the optic disc reference line [1], [12]. Thus the accuracy of optic disc detection is essential for accurate optic cup detection.

A limitation of our work is that experiments on glaucoma patients were not included. We note, however, that for some retinal diseases such as glaucoma, the early changes in the optic disc are subtle [36], [37], such that the structure prior for dividing the training and candidate regions in our method remain valid. Though we did not provide an evaluation of glaucoma diagnosis, our disc detection method outperforms the other methods by nearly 10% in width accuracy for 2D OCT slices, and by 5% in area accuracy for 3D OCT volumes.

Another limitation is in dealing with peripapillary atrophy. Because of the surrounding structure/tissue changes of the layer representing the RPE-choriocapillaris complex [38], [39], the RPE endpoints may be detected erroneously at the margin of the peripapillary atrophy with our technique. This challenging problem would also lead to failure of other optic disc detection methods. How to deal with peripapillary atrophy is an important direction for future work.

IV. CONCLUSION

In this paper, we have proposed a method for detecting the optic disc in a single OCT slice. Our method takes advantage of the low-rank appearance structure and smooth shape variation along the RPE to identify the RPE endpoints that bound the optic disc. The low-rank dictionary discovers the intrinsic appearance structure of the RPE layer from training data that may contain outlier elements such as heavy noise and vessel shadows. Through a combination of low-rank reconstruction errors and a prior on RPE shape, the transition from RPE to optic disc can be detected with high accuracy in comparison to the current state-of-the-art methods. Moreover, our approach can also be employed to handle 3D OCT volumes, for which promising results are also achieved.

REFERENCES

IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING


