



HAL
open science

Automated segmentation of macular layers in OCT images and quantitative evaluation of performances

Itebeddine Ghorbel, Florence Rossant, Isabelle Bloch, Sarah Tick, Michel Paques

► **To cite this version:**

Itebeddine Ghorbel, Florence Rossant, Isabelle Bloch, Sarah Tick, Michel Paques. Automated segmentation of macular layers in OCT images and quantitative evaluation of performances. *Pattern Recognition*, 2011, 44 (8), pp.1590-1603. 10.1016/j.patcog.2011.01.012 . hal-02559230

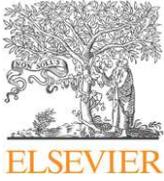
HAL Id: hal-02559230

<https://hal.science/hal-02559230>

Submitted on 30 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Automated segmentation of macular layers in OCT images and quantitative evaluation of performances

Itebeddine Ghorbel^{a,b,c,*}, Florence Rossant^a, Isabelle Bloch^b, Sarah Tick^d, Michel Paques^d

^a Institut Supérieur d'Electronique de Paris (ISEP), Paris, France

^b Télécom ParisTech, CNRS UMR 5141 LTCI, Paris, France

^c FOVEA Pharmaceuticals Paris, France

^d Clinical Investigation Center 503, Centre Hospitalier National des Quinze-Vingts Paris, France

ARTICLE INFO

Keywords:

Optical coherence tomography
Retinal imaging
Automated segmentation
Quantitative evaluation

ABSTRACT

Optical coherence tomography (OCT) allows high-resolution and noninvasive imaging of the structure of the retina in humans. This technique revolutionized the diagnosis of retinal diseases in routine clinical practice. Nevertheless, quantitative analysis of OCT scans is yet limited to retinal thickness measurements. We propose a novel automated method for the segmentation of eight retinal layers in these images. Our approach is based on global segmentation algorithms, such as active contours and Markov random fields. Moreover, a Kalman filter is designed in order to model the approximate parallelism between the photoreceptor segments and detect them. The performance of the algorithm was tested on a set of retinal images acquired in-vivo from healthy subjects. Results have been compared with manual segmentations performed by five different experts, and intra and inter-physician variability has been evaluated as well. These comparisons have been carried out directly via the computation of the root mean squared error between the segmented interfaces, region-oriented analysis, and retrospectively on the thickness measures derived from the segmentations. This study was performed on a large database including more than seven hundred images acquired from more than one hundred healthy subjects.

1. Introduction

Recently, optical coherence tomography (OCT) has revolutionized the clinical imaging of the retina, the eye fundus membrane including the photosensitive tissues of the eye. OCT is a recent real-time imaging modality that allows the acquisition of high-resolution cross-sectional images of thin layers of biological tissues [1]. OCT is based on the principle of the Michelson interferometer and can be seen as the optical equivalent of echography [2].

The spectral domain OCT technology now reaches a 5–10 μm resolution while the high acquisition speed reduces motion artifacts, allowing reliable image averaging [3,4]. The introduction of the tracking laser tomography, a technology which enables real-time tracking of eye movements, provides precise location information and improves the OCT image quality [3]. Innovative segmentation approaches can now be proposed for automatically analyzing the retina OCT images, while reaching a better accuracy and a better reliability. We refer to [5] for an overview of general segmentation approaches. Then, measures derived from the

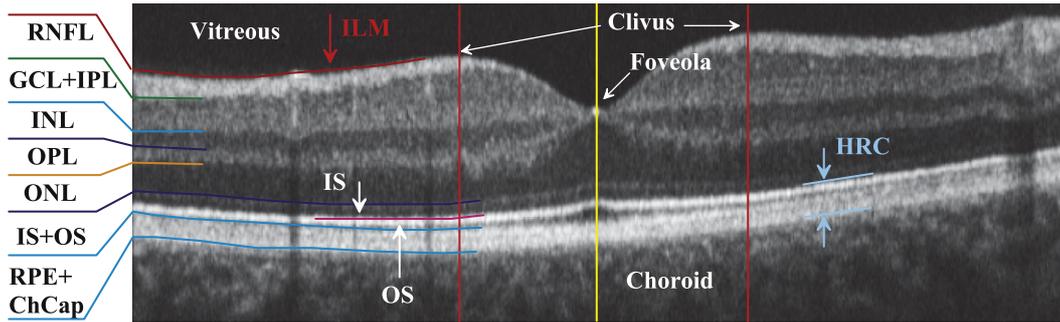
segmented retinal layers (Fig. 1) will help to better understand the anatomy of the human fovea.

Given the difficulty and the tedious nature of performing manual segmentations (by physicians), a number of automated approaches [6–15] have been proposed in order to segment the different retinal layers. Although there are differences among existing methods, most of them rely on three main steps: pre-processing, interface detection and post-processing.

The pre-processing step aims at improving the quality of the OCT image. We can distinguish three main methods, used independently or in combination:

- attenuation of the noise by applying mean filters [6], median filter [7,8], Gaussian filter [8,9] or nonlinear diffusion filters [10–12];
- OCT image enhancement, by performing image equalization [9], applying directional filters [9] or coherence-enhanced diffusion filtering [10];
- A-scans (columns) alignment with regard to a reference retina layer, determined by cross correlation of adjacent columns [6] or with regard to an already detected interface [9,11].

The filtered OCT images are then passed to the interface detection process. This second step is generally based on the search for peak



ILM: Inner Limiting Membrane; RNFL: Retina Nerve Fiber Layer; GCL+IPL: Ganglion Cell Layer+Inner Plexiform Layer; INL: Inner Nuclear Layer; OPL: Outer Plexiform Layer; ONL: Outer Nuclear Layer; IS: Inner Segments of the photoreceptors; OS: Outer Segments of the photoreceptors; RPE: Retinal Pigment Epithelium; HRC: Hyper-Reflective Complex.

Fig. 1. Cross-sectional OCT image of the retina, acquired with the 3D OCT-1000, Topcon [4], and retinal layers definition.

intensity in each pre-processed A-scan, each one considered separately. Thus, authors make extensive use of local boundary detection methods, as local profile analysis or local gradient analysis [6–11,13]. The most used filter is the first derivative of a 1D Gaussian kernel, whose standard deviation is tuned in order to obtain a compromise between robustness and accuracy. These methods exploit knowledge about the relative position and relative intensity of the retinal layers. For example, the authors in [7] use the fact that GCL+IPL (*Ganglion Cell Layer + Inner Plexiform Layer*) and OPL (*Outer Plexiform Layer*) appear brighter than INL (*Inner Nuclear Layer*) and ONL (*Outer Nuclear Layer*). However, these methods are not very robust to noise and varying contrast, due to their local nature, and lack of global consistency. Moreover, they generally require a lot of parameters empirically tuned [6,11].

For this reason, further processing steps are often required to select and label the significant edges. Several approaches have been proposed in order to achieve this third stage, namely the post-processing stage. A first idea is to check the continuity of the detected points [6,7,10,11]. The missing points are then obtained by interpolation (linear interpolation [10], elliptic Fourier approximation [8]), or model fitting imposing that the searched curve passes through one point at the foveola [10]. In other methods, the positions of the detected edges are compared with retina models [7,9]. For example, Bagci et al. use a normalized retina model giving the average location of each boundary [9]. In the same way, a mathematical model of the retinal structure is defined as a Markov process in [7], to identify, complete, and characterize the interfaces. These methods require a prior training phase which may be a limitation for clinical routine.

Contrary to the approaches described previously, some authors propose methods based on global optimizations, performed by maximizing a cost function that integrates local information on contours or regions. These algorithms allow one to better take into account the regularity of the retinal layers. However, the relevant pixels have to be preselected, either by local methods previously cited [8], or by checking the regularity and the relative position of the possible interfaces, which needs many parameters [11,12].

Very recently, a new global method was presented to deal with the presence of noise and realize the simultaneous segmentation of all retinal layers [14]. It uses an active contour with an external force following the gradient of the image through a vector fitness function kernel. This algorithm produces good results on retinal regions where the contours are approximately parallel. Nevertheless, it seems impractical around the foveal depression (in the center of the macula), since all interfaces almost coincide at the foveola.

Finally, it is worth mentioning other kinds of approaches that are mostly region oriented. Baroni et al. characterize the retina

tissues through indices derived from co-occurrence matrices and train a neural network in order to classify the pixels into retinal layers [15]. The main advantage is that no a priori about the relative position of the interfaces is needed, allowing to apply the algorithm on pathological images. Nevertheless, the obtained results show a lack of accuracy. Fuzzy membership degrees have also been defined and incorporated into a global optimization process, in order to extract the inner retinal layers [12]. This method still requires some anatomical prior knowledge to initialize the parameters defining the membership functions, in order to cope with the image varying contrast.

Despite the noticeable level now reached by the research in this domain, all retinal layers are not yet detected. In particular, only two interfaces are detected in [7], three in [8], five in [6], six in [11], seven in [10], and eight in [9]. Moreover, the segmentation is not always made around the foveola [10,11,14].

The specific purpose of this work is to develop a completely automated and reliable segmentation method, that is accurate and robust to the variability of images and morphologies. In addition it should be applicable to images acquired from different devices and on a large data base including hundreds of images. For this, we propose to avoid very local analysis method and to apply more global segmentation algorithms, such as active contours, k-means and Markov random fields, combining local and global information. Another originality of our approach is the modeling of the approximate parallelism between the layers, based on a Kalman filter. As a result, nine interfaces of the retina can be determined in the foveal region, including the interface between the inner segments of photoreceptors (IS: *Inner Segments of the photoreceptors*) and the outer nuclear layer (ONL), which, to our knowledge, had not yet been addressed until now.

The remainder of the article is organized as follows: the segmentation method is presented in Section 2, experiments, including quantitative measurements, are discussed in Section 3. This paper extends our preliminary work presented in [16], in particular by improving the proposed segmentation method mainly for the detection of HRC layers. In addition, results have been improved and are obtained on a much larger database of images, coming from different acquisition devices.

2. Segmentation

The images were acquired with the 3D OCT-1000 (Topcon) [4] and the Spectralis HRA+OCT (Heidelberg) [3] imaging systems, to provide vertical and horizontal retina cross-sections. Examples are displayed in Fig. 2.

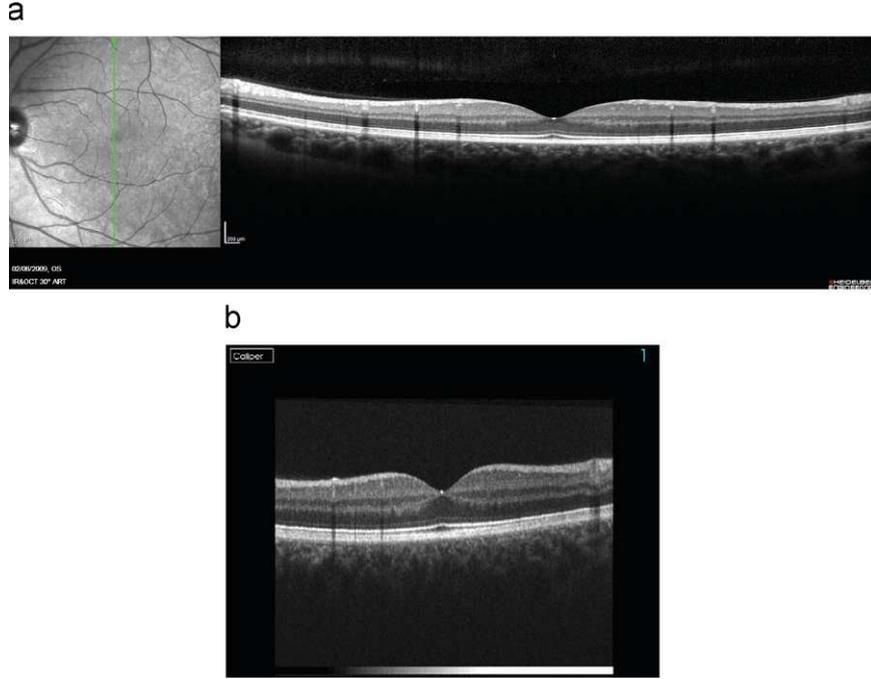


Fig. 2. Examples of vertical cross-sections provided (a) by the Spectralis [3] and (b) by the 3D OCT-1000 [4]; both images are resized with the same scale factor.

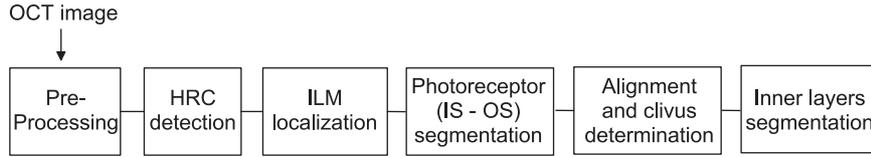


Fig. 3. Flowchart of the proposed segmentation method.

Table 1
The characteristics of both types of images.

	Topcon	Spectralis
Typical sizes (pixels), $H \times W$	450×600	496×1537
Section width (mm)	6	9
Horizontal resolution ($\mu\text{m}/\text{pixel}$)	10	6.0
Vertical resolution ($\mu\text{m}/\text{pixel}$)	2.6	4.0

Fig. 3 depicts the flowchart of the proposed analysis process. The images are first pre-processed, in order to extract the OCT data and remove noise. Then, we use anatomical knowledge about the relative position of the retinal layers and their appearance in the OCT images, to detect and localize sequentially the retinal layers. The two first steps consist of the detection of the interfaces presenting the highest contrasts, namely the contours of the hyper-reflective complex (HRC) and the inner limiting membrane (ILM: *Inner Limiting Membrane*) (see Fig. 1). At this stage, the retina area is completely delineated. The analysis of the inner boundary of the HRC is then refined so that the inner (IS) and outer (OS: *Outer Segments of the photoreceptors*) photoreceptor segments are accurately localized. The last two steps aim at aligning the images and segmenting the inner retinal layers (RNFL: *Retina Nerve Fiber Layer*, GCL+IPL, INL, OPL, ONL).

2.1. Pre-processing

The images are first cropped, in order to extract the OCT data, and normalized between 0 and 1. Table 1 summarizes the typical

sizes, section width, and resolution of the obtained images depending on the OCT device.

Note that the actual values may vary around these typical values among the cases, and are not always precisely known.

A nonlinear diffusion filter [17] is applied to the image (Fig. 4). This filter performs better than median or Gaussian filters since edges are better preserved.

In what follows, we denote by W and H the image width and height. The coordinates are defined by the origin at the top left corner, the vertical x -axis and the horizontal y -axis.

2.2. Detection of the hyper-reflective complex (HRC)

The hyper-reflective complex includes the ChCap, RPE (*Retinal Pigment Epithelium*) and OS layers, and appears as a high intensity image strip above the choroid. The segmentation proceeds in three sequentially applied steps. First, a pixel (x_0, y_0) belonging to the HRC is selected as the maximum response of an average filter applied on the central column ($y_0 = W/2$). Then, the profile analysis allows estimating roughly the HRC thickness T_{HRC} . The second step consists in detecting the HRC median line. A Gaussian 1-D filter, whose standard deviation is set to $\sigma = T_{HRC}/2$, is applied column-wise in order to provide a vertically smoothed image $S(x, y)$. Starting from the inner point (x_0, y_0) , the HRC median line is iteratively deduced, column by column, by looking for the maximum output of the following recursive low-pass filter:

$$C(x, y) = (1 - \alpha)S(x, y) + \alpha C(x, y \pm 1) \quad (-\text{from left to right, } +\text{from right to left}) \quad (1)$$

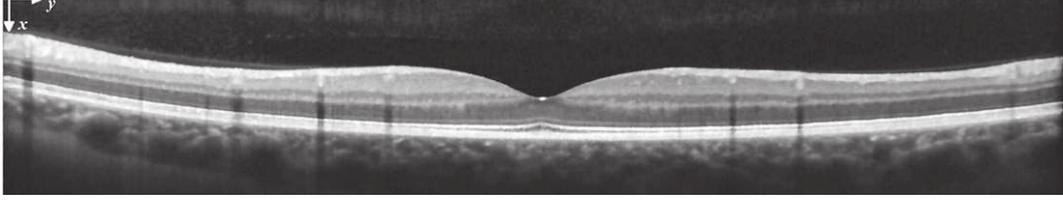


Fig. 4. Result of a non-linear diffusion filter [17], applied on an image acquired with the Spectralis device.

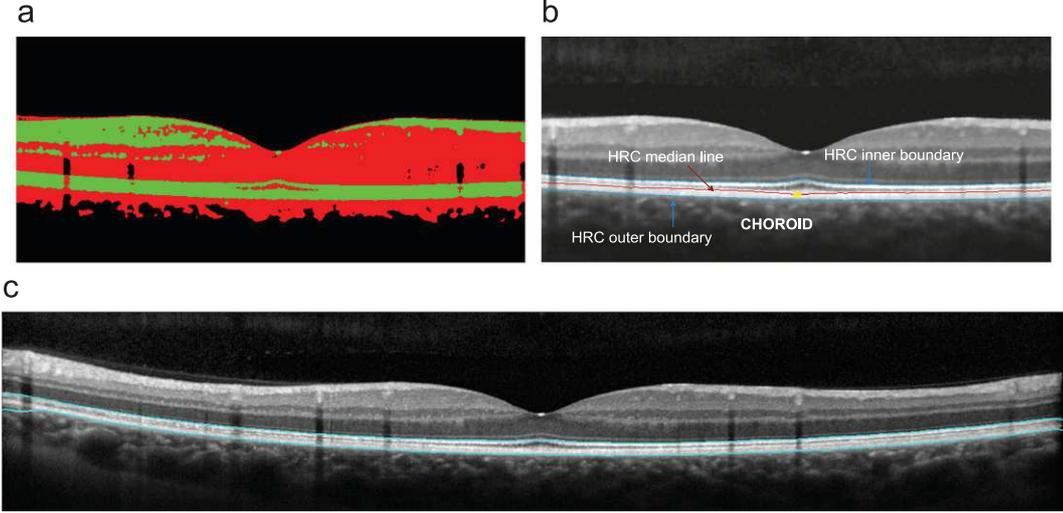


Fig. 5. (a) k-means segmentation (label 1 in black, label 2 in green, label 3 in red); (b) median line (red) deduced from a first inner point (yellow) and a first approximation of the HRC contours (cyan) used as initialization of an active contour; (c) final HRC segmentation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In this method, the $S(x,y)$ coefficients are continuously integrated before making the decision, in order to model the continuity of the searched layer. Setting $\alpha = 0.9$ leads to a regular median line, despite the residual noise and the retinal blood vessel shades (Fig. 5b). The third step aims at localizing the HRC contours. We first apply a k-means classification ($k=3$ classes) (Fig. 5a). The regions containing the median line and labeled ($k=2$) are selected and their upper boundary defines the inner contour of the HRC. The cumulative profile of the intensity image, computed along the median line and centered on it, allows us to determine the approximate distance between the median line and the HRC outer boundary, providing a good and regular estimation of this one, whatever the contrast between the HRC and the choroid (Fig. 5b). Thus, it is possible to initialize an active contour [18], which will converge accurately towards the boundaries (Fig. 5c). The active contour consists of a curve $X(s) = [x(s), y(s)]$, $s \in [0, 1]$, that moves through the spatial domain of an image to minimize the following energy functional:

$$E(X) = \int_0^1 \frac{1}{2} (\alpha |X'(s)|^2 + \beta |X''(s)|^2 + E_{Image}(X(s))) ds \quad (2)$$

where α and β are controlling the rigidity and the tension of the contour, respectively, hence the regularization. The image term (data fidelity) is computed as a spatial diffusion of the gradient of an edge map derived from the image.

The active contour is used for regularization purpose. Searching for a local minimum of the energy functional ensures the global consistency of the segmentation results.

2.3. Localization of the inner limiting membrane (ILM)

The k-means classification image (Fig. 5a) is again used in order to provide a first localization of the ILM. Given that the

vitreous pixels are predominantly classified in class 1 (corresponding to the darkest regions), while the inner retina pixels are labeled 3 or 2, the following algorithm can be applied. Each image column is scanned from top to bottom. The pixel boundary is identified as the first pixel classified in class 3, if it is found above the HRC, or as the first pixel classified in class 2, otherwise. However, this method does not provide a continuous interface, due to the presence of noise in the vitreous and inner layers. Consequently, the obtained results are only used to select two pixels presenting the highest gradients on the left and right parts of the image. Then, an edge-tracking algorithm, based on the maximization of the local mean gradient, is applied, starting from the two selected pixels. Both obtained curves are merged by choosing the solution that maximizes the mean gradient on the non-coincident parts. The ILM boundary such defined is then refined and regularized by applying the active contour algorithm. Finally, the foveola $F(x_F, y_F)$ is defined as the ILM curve pixel whose x-coordinate is maximal in the central area (Fig. 6).

2.4. Segmentation of the photoreceptor segments (IS, OS)

The junction between the inner (IS) and the outer (OS) segments of the photoreceptors appears as a bright narrow band just above the RPE+ChCap layer, with a larger separation around the foveola y-coordinate. We apply a peak detector on an image area determined with respect to the HRC. The maxima are detected in each column and labeled in order to form peak lines (Fig. 7a). Then, the IS/OS junction extraction process is initialized by selecting the peak line that minimizes the distance to the inner side of the hyper-reflective complex. Other peak lines are iteratively selected in order to complete the detection of the IS/OS junction, based on two criteria defined with respect to the current detected curve: similar mean distance with respect to the HRC, no

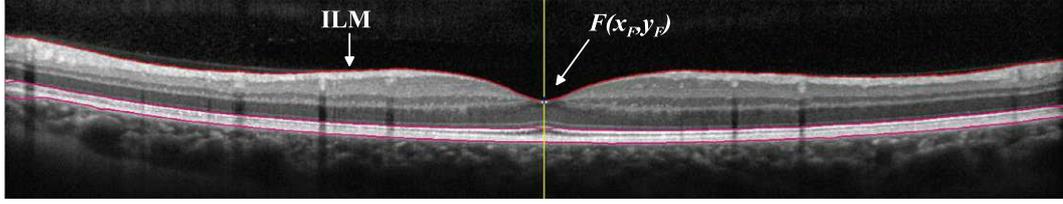


Fig. 6. Detection of the ILM, and first localization of the foveola.

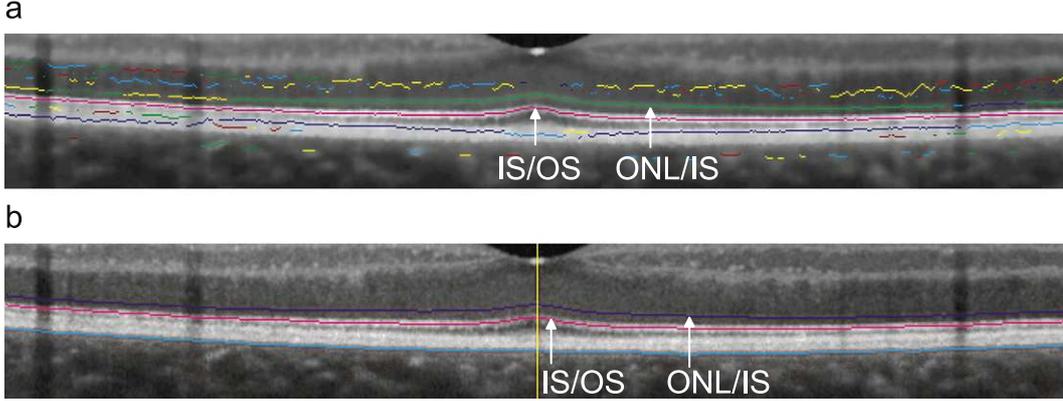


Fig. 7. ONL/IS and IS/OS localization: (a) detection and labeling of the intensity maxima; (b) final result obtained after concatenation for the IS/OS interface and with the Kalman filter for the ONL/IS interface.

horizontal overlap. The detected curve is used as initialization of an active contour algorithm to provide a continuous and regular curve (Fig. 7b). The interface between the ONL and the IS layers corresponds to a second fine line, parallel to the IS/OS junction, but less bright and more noisy. We propose to apply a Kalman filter [19] to track column by column the ONL/IS boundary. The Kalman filter is defined by a state vector X , that characterizes the tracked pixels, and a theoretical model that allows predicting the evolution of X . In our application, the state vector includes the intensity value of the pixel and its distance to the IS/OS line, and both values are supposed to be constant along the searched curve. Note that the distance provides an information about the parallelism between lines, as an original feature of our approach. State noise and measurement noise variances complete the model. All the parameters are dynamically initialized from the peak lines found just above the IS/OS junction, except the state noise which has been set for all images.

The Kalman equations (3) formalize the principle of prediction/update, realized in three steps.

$$X = \begin{pmatrix} \text{Intensity} \\ \text{Distance} \end{pmatrix}$$

Prediction:

$$\hat{X}(n | n-1) = F\hat{X}(n-1 | n-1)$$

Update:

$$\hat{X}(n | n) = \hat{X}(n | n-1) + G(n)(Y(n) - \hat{X}(n | n-1)) \quad (3)$$

Firstly, the state vector $\hat{X}(n | n-1)$ at column n is predicted from the previous estimate $\hat{X}(n-1 | n-1)$, according to the evolution model (F is the identity matrix in our application). Secondly, a measure $Y(n)$ is made in the image, by seeking the maximum intensity pixel around the predicted position. Thirdly, the error between the prediction and the measure is computed in order to correct the prediction and provide the new state vector estimate $\hat{X}(n | n)$. The correction is weighted by the filter gain $G(n)$, the optimal Kalman gain, which is also adjusted at each iteration.

Nevertheless, when no image pixel fits the prediction, the predicted value is kept and no measure is integrated in the filter. This allows overcoming the problem of short interruptions. The process stops after several iterations without acceptable measures. The detected curve is then regularized by applying an active contour. The proposed method leads to an accurate localization of the ONL/IS boundary, despite the low contrast and the noise, as illustrated in Fig. 7b.

A second Kalman filter could be applied on the gradient image, to deduce the inner side of the RPE layer (i.e OS/RPE boundary) from the outer side of the RPE+ChCap layer (i.e. the outer side of the HRC), since both curves are almost parallel as can be observed in Fig. 8d. The method works very well for images of good quality but fails on a number of configurations for which the contrast and the parallelism are simultaneously lost. For this reason, a less accurate but more robust approach is proposed.

The thickness $T_{RPE+ChCap}$ of the RPE+ChCap layer is first estimated around the foveola, based on the k-means segmentation (Fig. 8a) realized between the IS/OS junction and the HRC outer boundary, with $k=2$ classes. This provides a first estimation of the RPE+ChCap inner boundary, by simply shifting the outer boundary by $T_{RPE+ChCap}$ pixels upwards (Fig. 8b). Then, a local analysis is performed around this curve on each image column. Given an y -coordinate, the OS/RPE edge pixel is defined as the first gradient maximum found just under the intensity minimum, this one being itself located below the IS/OS junction. The detected points form a curve (Fig. 8c) that is then regularized by applying an active contour (Fig. 8d).

2.5. Alignment and clivus determination

The image is aligned by performing vertical column translation against the outer side of the RPE+ChCap layer. The clivus is defined by the two highest points of the ILM found on both sides of the foveola, denoted by (x_{Cl}, y_{Cl}) and (x_{Cr}, y_{Cr}) , as illustrated in Fig. 9.

The foveola position is refined by maximizing the distance between the ONL/IS and the OS/RPE interface.

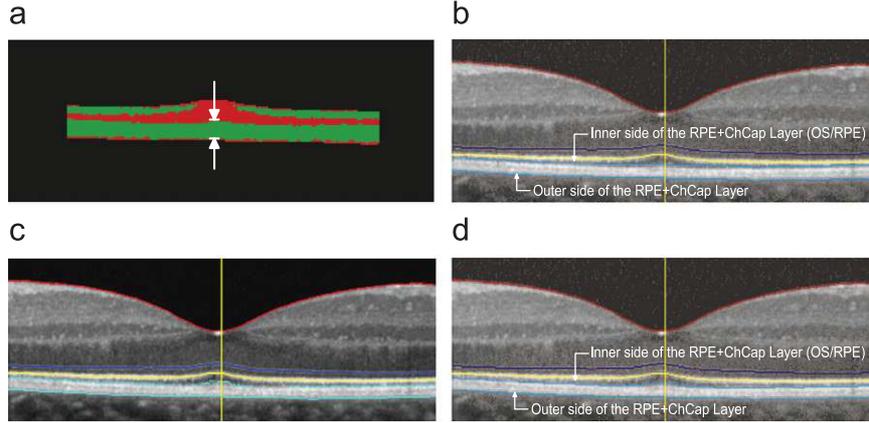


Fig. 8. Localization of the OS/RPE boundary: (a) k-means for the RPE+ChCap thickness estimation; (b) first estimation of the RPE+ChCap inner boundary; (c) result obtained with the local analysis; (d) final result obtained after regularization.

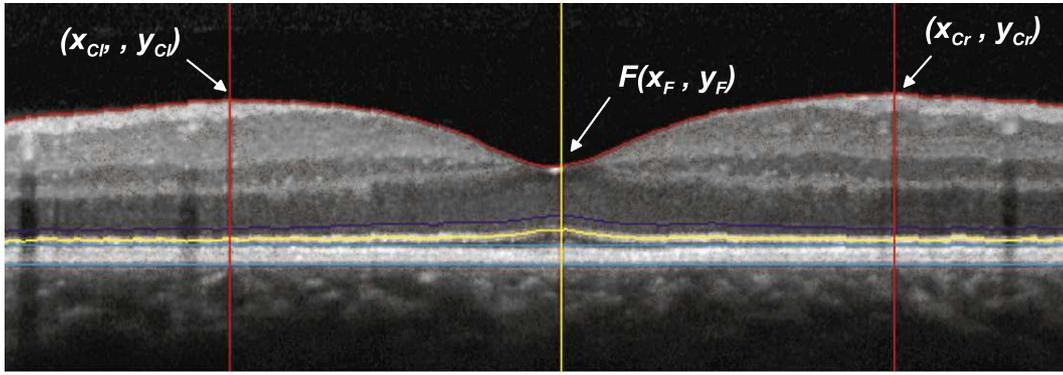


Fig. 9. Alignment and clivus determination.

2.6. Inner layer segmentation (RNFL, GCL+IPL, INL, OPL, ONL)

The inner layers are very noisy and low contrasted. This observation led us to choose a region-oriented segmentation approach. Especially, the Markov Random Field (MRF) [20] segmentation method allows us to model the noise and to take into account spatial interactions between connected pixels. In our application, the classification is performed according to the Bayesian maximum a posteriori (MAP) criterion, i.e. by looking for the label configuration that maximizes the probability of the class field (the labels) conditionally to the observation field (the intensity image). This optimal configuration corresponds to a minimum state of an energy function, defined as follows.

Let us denote by $P(f_s | w_s = i)$ the probability distribution of the pixel intensities (f_s), conditional to the class i . The analysis of the image intensity within inner retinal regions showed empirically that the image noise is Gaussian. The probability distribution is then defined by

$$P(f_s | w_s = i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(f_s - \mu_i)^2}{2\sigma_i^2}} \quad (4)$$

where μ_i and σ_i are, respectively, the mean and the standard deviation. We use the Potts model [21] $\varphi(w_s, w_t)$ to express the interactions between the 8-connected sites s and t . Under these assumptions, the energy function is defined by

$$U(w | f) = \sum_s \left[\frac{(f_s - \mu_{w_s})^2}{2\sigma_{w_s}^2} + \ln(\sqrt{2\pi}\sigma_{w_s}) \right] + \beta \sum_{(s,t)} \varphi(w_s, w_t) \quad (5)$$

$\mu_{w_s} = \mu_i$ and $\sigma_{w_s} = \sigma_i$ given $w_s = i$.

The first term is related to the image data, while the second one is a regularization term. The parameter β , empirically set, weights the relative influence of both. We first apply the k-means algorithm to initialize the labels and estimate automatically the noise parameters (μ_i, σ_i) from the classification results. So, the segmentation process is fully unsupervised. Then, the energy function is minimized by running the simulated annealing (SA) algorithm [22]. This segmentation method is applied several times, on the nonpre-processed aligned image, the region of interest being adjusted each time. The contours are deduced from the labeled image, and regularized by the active contour. More details are given in what follows for each layer. The final results are shown in Fig. 13.

2.6.1. RNFL

The retinal nerve fiber layer is segmented by applying the k-means algorithm on the top part of the aligned image, around the ILM membrane, with $k=2$ classes. As the contrast is very high, the classification results are very homogeneous (Fig. 10a) and the MRF classification is not required. The clusters corresponding to the higher mean intensity and connected to the ILM membrane are selected. From this first result, it is easy to deduce the contours of the RNFL, on both sides of the foveola (Fig. 10b), and to regularize them by applying the active contour algorithm (Fig. 13).

2.6.2. ONL

The k-means/MRF segmentation process is applied on the foveal region (Fig. 11a), defined horizontally by the clivus, limited vertically by the ILM or the RNFL and by the ONL/IS boundary,

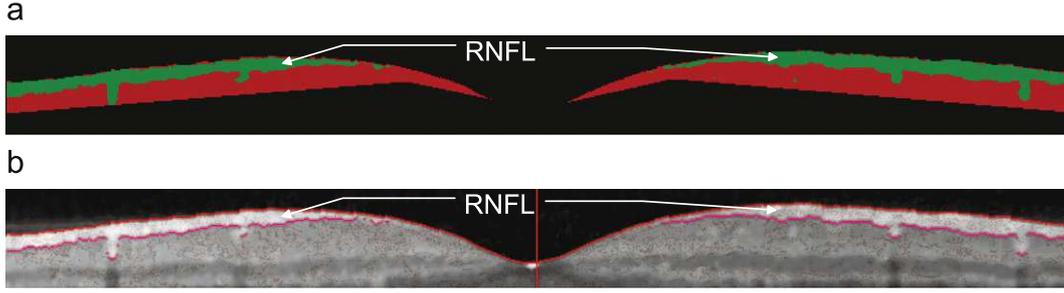


Fig. 10. Segmentation of the RNFL layer: (a) k-means applied on an area deduced from the ILM; (b) initialization of the active contour.

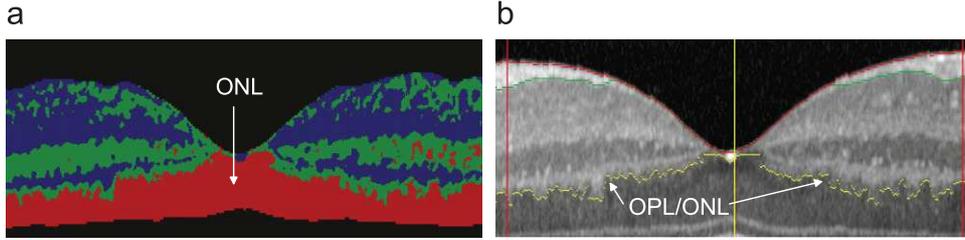


Fig. 11. Segmentation of the ONL layer: (a) k-means/MRF segmentation outputs (label 1 in red), (b) OPL/ONL localization. (For a better visualization, the image resolution is increased by a factor 2 in the vertical direction). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

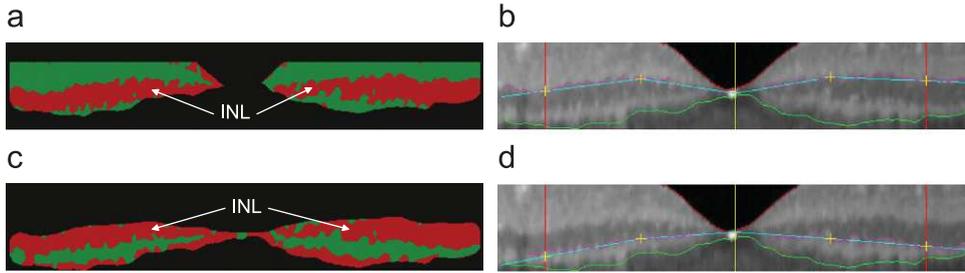


Fig. 12. Localization of the INL layer: (a) k-means/MRF segmentation used for the detection of the GCL+IPL/INL boundary (label 1 in red, label 2 in green); (b) first estimate of the GCL+IPL/INL interface (in cyan) and transition points detected around the previous estimate (in purple); (c) k-means/MRF segmentation used for the detection of the INL/OPL boundary; (d) first estimate of the INL/OPL interface. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with $k=3$ and $\beta=2$. The ONL appears as the cluster corresponding to the lowest mean intensity (label 1) just above the ONL/IS boundary. So the OPL/ONL boundary can be easily deduced from the labeled image (Fig. 11b). It is then regularized using the active contour algorithm (Fig. 13).

2.6.3. INL

This layer appears less bright than its neighbors. The region of interest is now more restricted, knowing the location of the OPL/ONL interface (Fig. 12a). The method is applied with $k=2$ classes and $\beta=5$, on the left and right parts independently, since the average contrast may be slightly different on both sides. Clusters of low mean intensity (label 1) and connected to the upper boundary of the region of interest are fused with the surrounding cluster (label 2). The location of the searched boundary is then estimated at four y -coordinates ($y_{cl}, y_{cl}/2, y_{cr}/2, y_{cr}$) and forced at the foveola y_f to pass between the ILM and the OPL/ONL interface. These five points are linearly interconnected to provide a first estimate of the searched boundary (Fig. 12b).

Finally, the labeled image is explored column by column from top to down. Pixels corresponding to a transition between labels 2 and 1 and close to the first estimate are marked as boundary points for the upper part of INL (GCL+IPL/INL interface) (Fig. 12b). These points serve as initialization of an active contour, for the purpose of filling in the gaps and regularizing the curve (Fig. 13).

A similar segmentation process is applied on the region limited by the curve found previously and the OPL/ONL interface. In this case, the classification image (Fig. 12c) is explored from down to top in order to detect the transition between labels 2 and 1, corresponding to the INL/OPL interface (Fig. 12d). The obtained INL segmentation is shown in Fig. 13.

3. Experiments and results

We detail in this section the results we obtained experimentally, as well as the evaluation protocol of these results. This protocol relies on a comparison of the segmentations realized automatically by the proposed method and manually by the experienced physicians, and a study of the intra and inter-physician variability. Derived clinically relevant measurements are compared as well. All these measures are analyzed to estimate the precision and the accuracy of the proposed method.

3.1. Database

The database includes more than 700 images, from more than 100 healthy subjects. For the majority of the patients, two sections of the retina were acquired, along a horizontal and a vertical line for each eye. All subjects had a best corrected visual acuity of at least 80 points on the ETDRS (Early Trial Diabetic

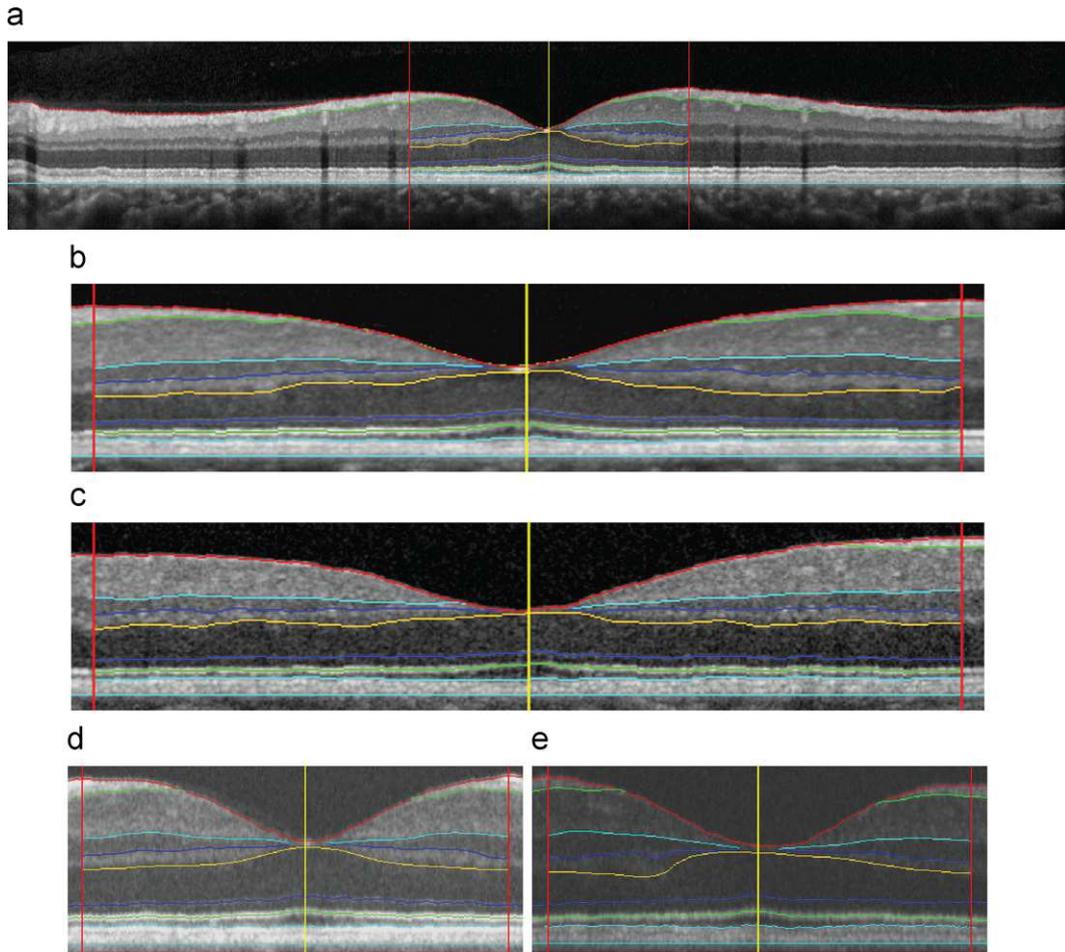


Fig. 13. (a) Final segmentation; (b) and (c) zoom on the foveal region between the clivus from two different Spectralis images; (d) and (e) zoom on the foveal region between the clivus from two different Topcon images.

Retinopathy Study) charts [23], and a normal fundus. This study was conducted in accordance with French ethics regulation and all subjects gave informed consent to participate. These images were acquired by two different devices from the same model, 3D OCT-1000 of Topcon [4], in different intervals of time, and from the Spectralis Heidelberg OCT [3]. The segmentation method was applied to all the images, and allowed determining more than 99% of all layer interfaces. These results were visually validated by an expert, whose experience in the field of OCT image interpretation overtakes 12 years. The 1% failure detection occurs mostly for the OPL/ONL boundary in some especially blurred images due to the lack of strong gradients. The continuation of the evaluation concerns the 99% of well-detected interfaces.

3.2. Evaluation protocol

We performed a quantitative evaluation of our method by comparing the measures obtained automatically with those derived from manual segmentations performed by five physicians, on 14 images of the Topcon database and 16 images of the Spectralis database. These images were selected to be representative of the quality and noise levels that occur throughout the image databases.

The manual segmentation was performed using a software developed under Matlab. Each physician has only to click with the mouse on a finite number of points which belong to the chosen

interface and a spline interpolation is then computed. The physician can repeat as many times as he wants each interface segmentation until he is satisfied. The physicians have different experiences in the field of the OCT images, varying from 4 years to more than 12 years. The evaluation protocol is decomposed in several steps:

- direct evaluation of the precision of interface locations by calculating the mean square error (MSE), for each one, between two different segmentations (Section 3.3.1);
- direct evaluation of regions (i.e. retinal layers) by calculating similarity (S), specificity (SPEC) and sensitivity (SENS) indices between two different segmentations (Section 3.3.2);
- evaluation of the derived measures of thickness of the main retinal layers (retrospective evaluation of the segmentation) to provide data for physicians, in order to study morphological variations within a healthy population or provide some indications about a potential pathology (Section 3.4).

This protocol will allow us studying the inter-physicians variability and the difference between automated results and physician's ones, by choosing as reference physician the most experienced one. We will also calculate the intra-physician variability of the MSE and derived measures, based on two segmentations performed by the same physician. Note that this last evaluation will be performed on only one image of the database segmented by the same doctor at two different times.

3.3. Evaluation of the segmentation

In this section we will proceed to a first evaluation of our method by direct comparison of the segmentation results. For this, we will start with an evaluation of the interface segmentation, and then we will evaluate the segmented regions.

3.3.1. Evaluation of interfaces detection and localization

We calculate for each interface the mean square error (MSE) between two different segmentations. Let us denote by $Seg^{(i,j)}(y)$ the abscissa of the interface i at the ordinate y for the segmentation performed by the physician j . We note y_L and y_R the ordinates of the left and right clivus. The MSE measured between two physicians j_1 and j_2 , expressed in pixels, is calculated by

$$MSE(Seg^{(i,j_1)}, Seg^{(i,j_2)}) = \frac{\sqrt{\sum_{y=y_L}^{y_R} \min_{y_C} \{d^2((Seg^{(i,j_1)}(y), y), (Seg^{(i,j_2)}(y_C), y_C))\}}}{y_R - y_L + 1} \quad (6)$$

where d is the Euclidean distance.

Tables 2 and 3 summarize the MSE obtained when comparing the segmentations provided by each physician (Ph_i) with the ones realized by the most experienced one (Ph_{ref}). The indicated values are the mean and standard deviation calculated on the evaluation sub-database. We also compare, in the same way, the interfaces provided by our automatic method with those of the most experienced physician.

As shown in Tables 2 and 3, our segmentation method gets, for all detected interfaces, a MSE in the range of the inter-physician ones. This proves the accuracy of the position of the interfaces obtained by the proposed method.

Moreover, for several interfaces such as GCL/INL and IS/OS, the MSE between our method and the most experienced physician is lower than the inter-physicians variability. This is especially noticeable for the interface ONL/IS that had not been detected in previously published works (Fig. 15). Conversely, for the RNFL, the mean MSE between our method and the most experienced

Table 2
Inter-physicians and automatic/the most experienced physician mean squared error (MSE) for the Topcon database.

	Ph_1/Ph_{ref}	Ph_2/Ph_{ref}	Ph_3/Ph_{ref}	Ph_4/Ph_{ref}	Auto/ Ph_{ref}
ILM	1.44 ± 0.39	1.56 ± 0.69	1.67 ± 0.47	1.73 ± 0.32	1.84 ± 0.59
RNFL	2.55 ± 0.92	2.07 ± 0.78	2.02 ± 0.46	1.84 ± 0.44	2.46 ± 0.53
GCL/INL	1.65 ± 0.75	1.94 ± 0.71	2.00 ± 0.80	2.32 ± 0.98	1.49 ± 0.49
INL/OPL	1.74 ± 0.92	1.77 ± 1.03	1.87 ± 0.76	1.74 ± 1.04	1.83 ± 0.82
OPL/ONL	2.73 ± 1.15	3.18 ± 1.72	3.59 ± 1.36	3.25 ± 1.40	3.03 ± 1.29
ONL/IS	1.47 ± 0.36	1.43 ± 0.42	1.42 ± 0.38	1.55 ± 0.63	1.33 ± 0.33
IS/OS	1.23 ± 0.34	1.32 ± 0.33	1.27 ± 0.34	1.53 ± 0.52	0.99 ± 0.20
OS/RPE	1.98 ± 1.12	1.99 ± 0.87	2.47 ± 1.57	3.04 ± 1.54	1.82 ± 0.77
RPE/Ch	1.83 ± 0.68	2.53 ± 1.43	1.70 ± 0.65	1.83 ± 0.85	1.87 ± 0.56

Table 3
Inter-physicians and automatic/the most experienced physician mean squared error (MSE) for the Spectralis database.

	Ph_6/Ph_{ref}	Ph_7/Ph_{ref}	Ph_8/Ph_{ref}	Ph_9/Ph_{ref}	Auto/ Ph_{ref}
ILM	1.87 ± 0.33	1.75 ± 0.28	1.45 ± 0.35	1.37 ± 0.42	1.89 ± 0.48
RNFL	1.96 ± 1.67	2.62 ± 1.82	1.97 ± 1.83	1.65 ± 1.81	1.30 ± 0.83
GCL/INL	1.21 ± 0.92	1.55 ± 0.89	1.59 ± 0.64	1.29 ± 0.52	1.19 ± 0.54
INL/OPL	0.97 ± 0.42	1.71 ± 0.72	1.20 ± 0.57	1.29 ± 0.59	1.65 ± 0.75
OPL/ONL	2.17 ± 0.52	3.64 ± 2.05	3.09 ± 1.02	2.47 ± 1.18	2.85 ± 1.17
ONL/IS	0.76 ± 0.18	1.36 ± 0.39	1.22 ± 0.40	1.20 ± 0.58	1.08 ± 0.38
IS/OS	1.06 ± 0.48	1.28 ± 0.44	1.15 ± 0.22	1.08 ± 0.15	0.94 ± 0.36
OS/RPE	0.91 ± 0.43	2.25 ± 1.00	2.97 ± 0.69	2.78 ± 0.99	1.48 ± 0.62
RPE/Ch	2.56 ± 0.91	1.53 ± 0.38	1.44 ± 0.48	1.45 ± 0.53	1.52 ± 0.51

physician is greater than the inter-physicians MSE, for the Topcon database. As shown in Fig. 14, physicians tend to distinguish the RNFL/GCL+IPL interface until the foveola, while this interface converges earlier to the ILM with our method. These differences explain the lower performances observed in this case. However, the same algorithm provides very accurate results on the Spectralis database.

As shown in Table 3, there is a high similarity in the segmentations made by the most experienced physician (Phref) and the physician Ph6. This is probably because Ph6 and Phref work together in the same team.

We also study the intra-physician variability, by calculating for each physician the MSE between two different segmentations of the same image. The results are presented in Table 4.

Table 4 shows the important intra-physician variability, with a mean square error exceeding 3 pixels in some cases. Averaged across the physicians, it varies between 1.33 pixels and 2.57 pixels. Although the results are not directly comparable, since the number of involved images is different, the intra-physician errors are on average of the same order of magnitude as the inter-physician errors, explaining the observed variability in Tables 2 and 3. This is due to the lack of strong gradients between the interfaces, leading automatically to some indetermination on their position, so that the same physician cannot reproduce identical segmentations. This is particularly obvious when considering the OPL/ONL interface (Fig. 15), which leads to the highest variability, for both inter and intra studies.

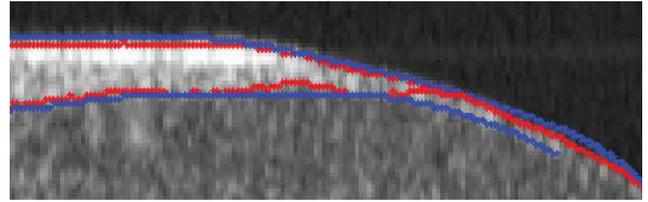


Fig. 14. Automatic (red) and manual (blue) segmentation of RNFL layer. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Intra-physicians mean squared error (MSE).

	Ph_{ref}	Ph_1	Ph_2	Ph_3	Ph_4	Mean
ILM	2.31	2.8	1.83	2.94	2.08	2.39 ± 0.42
RNFL	0	1.68	1.85	1.55	1.59	1.33 ± 0.67
GCL/INL	1.87	1.91	1.03	2.36	1.88	1.81 ± 0.43
INL/OPL	2.06	3.27	1.71	1.98	1.56	2.12 ± 0.60
OPL/ONL	2.31	3.02	2.04	2.99	2.51	2.57 ± 0.38
ONL/IS	2.9	2.26	1.23	2.59	1.76	2.15 ± 0.59
IS/OS	2.56	2.47	1.03	2.58	2.07	2.14 ± 0.59
OS/RPE	2.61	1.77	3.78	2.22	2.02	2.48 ± 0.71
RPE/Ch	2.11	1.34	2.28	1.93	1.87	1.91 ± 0.32

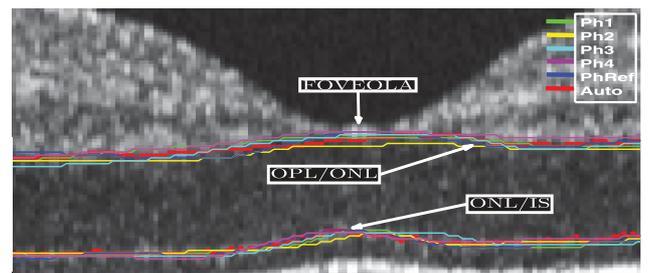


Fig. 15. Inter-physicians variability for interfaces OPL/ONL and IS/OS.

Despite these observed differences between the automatic segmentations and the ones provided by the most experienced physician, this quantitative analysis shows that the proposed segmentation method gives reproducible interfaces with enough accuracy for clinical practice.

3.3.2. Evaluation of region segmentation

In the previous section, we have evaluated the accuracy of our interface detection method. This study is now completed by a region-oriented analysis, in order to provide more information about the classification errors. Especially, we would like to quantify for each delineated retinal layer whether it is more prone to under-segmentation or over-segmentation, with regard to the manual segmentation realized by the physicians. In addition, and in the clinical context, ophthalmologists need to quantify some regions in order to diagnose some pathology. For example, a decrease of the ONL surface, compared to normal cases, can indicate a retinitis pigmentosa disease in its early stage [24].

Thus, we compared each automatically segmented region with the one segmented by the most experienced physician, based on similarity (Eq. 7), sensitivity (Eq. 8) and specificity (Eq. 9) indices.

In the following equations, $|Reg^{(i,j)}|$ represents the cardinality of the set of pixels of the region i segmented by the physician j .

$$S(Reg^{(i,j)}, Reg^{(i,Ref)}) = \frac{2|Reg^{(i,j)} \cap Reg^{(i,Ref)}|}{|Reg^{(i,j)}| + |Reg^{(i,Ref)}|} \quad (7)$$

$$SENS(Reg^{(i,j)}, Reg^{(i,Ref)}) = \frac{|Reg^{(i,j)} \cap Reg^{(i,Ref)}|}{|Reg^{(i,Ref)}|} \quad (8)$$

$$SPEC(Reg^{(i,j)}, Reg^{(i,Ref)}) = \frac{|Reg^{(i,j)} \cap Reg^{(i,Ref)}|}{|Reg^{(i,j)}|} \quad (9)$$

The similarity is sensitive to variations in shape, size and position. The sensitivity (*SENS*) and specificity (*SPEC*) measures give complementary information about how the overlap of both structures occurs. For example, if we have a low sensitivity with a high level of specificity, this means that the result only covers a part of the reference and does not overlap with the outside (under-segmentation) and vice versa. Both criteria are equal to 1 if a perfect overlap is achieved.

In this evaluation the two regions ONL and OPL are merged together, since we observed a large inter-physician and intra-physician variability in the determination of the OPL/ONL interface, revealed by the MSE previously measured (Tables 2–4).

Tables 5–8 show the obtained measures. It can be observed that results are satisfactory as *SPEC* and *SENS* are close to 1 in all cases, and almost equal to the inter-physician ones.

We obtain slightly worse results for the RNFL layer on the Topcon database, indicating an overall under-segmentation, which can be explained by the differences between the automated and manual detections of the RNFL interface, as underlined

Table 5
Sensitivity for the Topcon database.

	Ph_1/Ph_{ref}	Ph_2/Ph_{ref}	Ph_3/Ph_{ref}	Ph_4/Ph_{ref}	$Auto/Ph_{ref}$
RNFL	0.87 ± 0.06	0.91 ± 0.05	0.91 ± 0.06	0.88 ± 0.05	0.82 ± 0.08
GCL+IPL	0.97 ± 0.03	0.93 ± 0.05	0.93 ± 0.04	0.92 ± 0.04	0.97 ± 0.03
INL	0.91 ± 0.05	0.88 ± 0.08	0.92 ± 0.06	0.92 ± 0.06	0.88 ± 0.06
OPL+ONL	0.97 ± 0.02	0.97 ± 0.02	0.97 ± 0.03	0.97 ± 0.03	0.97 ± 0.02
IS	0.93 ± 0.03	0.94 ± 0.03	0.92 ± 0.03	0.94 ± 0.03	0.95 ± 0.02
OS	0.86 ± 0.10	0.84 ± 0.10	0.76 ± 0.16	0.9 ± 0.07	0.89 ± 0.10
RPE	0.92 ± 0.04	0.83 ± 0.08	0.96 ± 0.02	0.77 ± 0.06	0.86 ± 0.04

Table 6
Sensitivity for the Spectralis database.

	Ph_6/Ph_{ref}	Ph_7/Ph_{ref}	Ph_3/Ph_{ref}	Ph_8/Ph_{ref}	$Auto/Ph_{ref}$
RNFL	0.98 ± 0.03	0.97 ± 0.03	0.97 ± 0.02	0.95 ± 0.03	0.96 ± 0.03
GCL+IPL	0.91 ± 0.02	0.90 ± 0.04	0.92 ± 0.04	0.93 ± 0.03	0.92 ± 0.04
INL	0.92 ± 0.02	0.94 ± 0.03	0.95 ± 0.04	0.96 ± 0.03	0.97 ± 0.02
OPL+ONL	0.80 ± 0.01	0.95 ± 0.02	0.96 ± 0.02	0.96 ± 0.02	0.96 ± 0.02
IS	0.86 ± 0.01	0.98 ± 0.02	0.98 ± 0.01	0.98 ± 0.02	0.95 ± 0.02
OS	0.75 ± 0.07	0.73 ± 0.08	0.76 ± 0.07	0.77 ± 0.08	0.78 ± 0.14
RPE	0.99 ± 0.04	0.78 ± 0.05	0.79 ± 0.04	0.80 ± 0.05	0.81 ± 0.10

Table 7
Specificity for the Topcon database.

	Ph_1/Ph_{ref}	Ph_2/Ph_{ref}	Ph_3/Ph_{ref}	Ph_4/Ph_{ref}	$Auto/Ph_{ref}$
RNFL	0.94 ± 0.04	0.93 ± 0.03	0.94 ± 0.03	0.96 ± 0.02	0.97 ± 0.03
GCL+IPL	0.94 ± 0.04	0.97 ± 0.03	0.97 ± 0.03	0.97 ± 0.03	0.94 ± 0.04
INL	0.91 ± 0.08	0.91 ± 0.08	0.87 ± 0.08	0.86 ± 0.08	0.94 ± 0.06
OPL+ONL	0.97 ± 0.02	0.98 ± 0.03	0.98 ± 0.02	0.97 ± 0.02	0.97 ± 0.02
IS	0.89 ± 0.03	0.89 ± 0.03	0.9 ± 0.04	0.87 ± 0.05	0.89 ± 0.03
OS	0.88 ± 0.09	0.88 ± 0.09	0.91 ± 0.05	0.74 ± 0.14	0.88 ± 0.08
RPE	0.89 ± 0.07	0.93 ± 0.07	0.86 ± 0.08	0.97 ± 0.03	0.94 ± 0.07

Table 8
Specificity for the Spectralis database.

	Ph_6/Ph_{ref}	Ph_7/Ph_{ref}	Ph_3/Ph_{ref}	Ph_8/Ph_{ref}	$Auto/Ph_{ref}$
RNFL	0.89 ± 0.03	0.87 ± 0.03	0.88 ± 0.03	0.89 ± 0.03	0.89 ± 0.07
GCL+IPL	0.86 ± 0.06	0.84 ± 0.07	0.86 ± 0.05	0.87 ± 0.06	0.89 ± 0.08
INL	0.78 ± 0.09	0.74 ± 0.07	0.71 ± 0.05	0.73 ± 0.07	0.81 ± 0.09
OPL+ONL	0.72 ± 0.08	0.75 ± 0.08	0.69 ± 0.06	0.71 ± 0.09	0.79 ± 0.06
IS	0.84 ± 0.06	0.86 ± 0.06	0.82 ± 0.04	0.83 ± 0.06	0.88 ± 0.04
OS	0.73 ± 0.09	0.81 ± 0.10	0.72 ± 0.08	0.76 ± 0.09	0.81 ± 0.05
RPE	0.99 ± 0.02	0.92 ± 0.06	0.95 ± 0.05	0.92 ± 0.06	0.97 ± 0.05

in Section 3.3.1. On the other hand, the highest indices are obtained for the OPL+ONL region, which reaches very high sensitivity and specificity (0.97) on the Topcon images. Indeed, the result of the automatic segmentation is very close to the manual one. We also observe a very good robustness of the automatic segmentation process, since the standard deviations of the indices measured between the automatic and the reference segmentations do not exceed 0.14.

Finally, we verified for all regions the index of similarity (S). The index S is sensitive to variations in shape, size and position and a value of $S > 0.7$ indicates a strong agreement [25]. We found that S is in all cases greater than 0.7 and that it ranges from 0.87 for the RNFL layer to 0.97 for OPL+ONL layer.

3.3.3. Discussion of results of segmentation

This direct evaluation of the retinal layer segmentation confirms that our approach allows us to get accurate and robust segmentations, which are comparable to the physician's ones. Moreover, neither under nor over-segmentation was observed excepted for the RNFL layer on the Topcon images. This is important to avoid potential bias in the diagnosis. In addition, our method shows a good agreement with physicians for the localization of photoreceptor segments that had not been detected before.

3.4. Evaluation of retinal thickness measures

In this section we will evaluate the measures derived from the segmentation. The physicians defined ten thickness measures (Fig. 16), by taking the clivus as reference. These measures are:

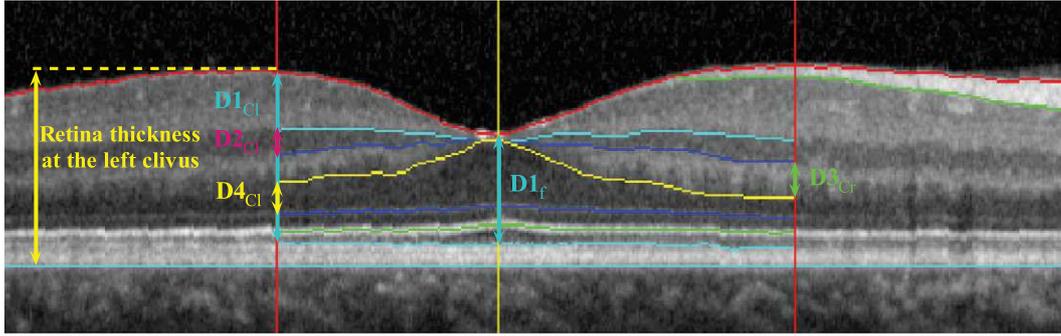


Fig. 16. Segmentation of the retinal layers and derived measures.

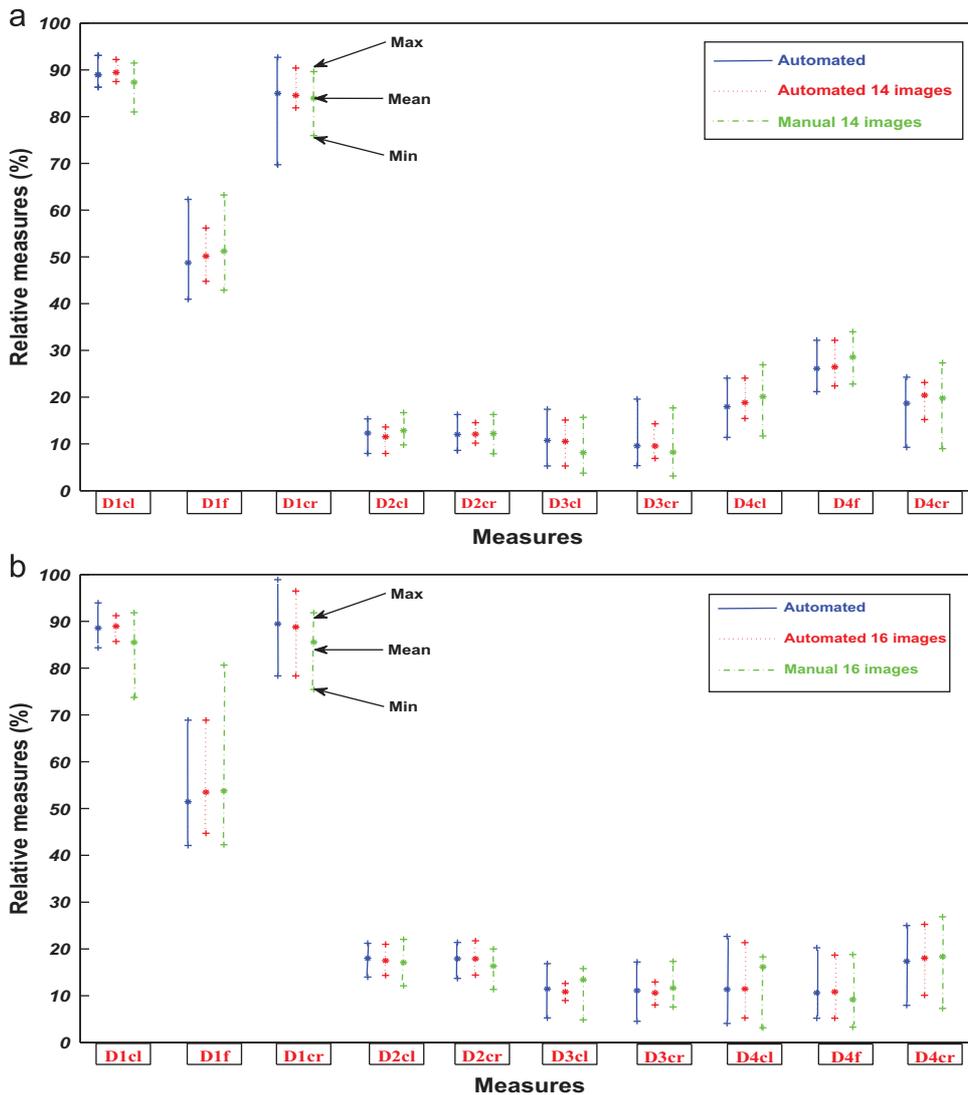


Fig. 17. Ranges of measures derived from the automatic and manual segmentations (Topcon database (a) and Spectralis database (b)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

foveal thickness ($D1_f$), total thickness at both clivus ($D1_{cl}$ and $D1_{cr}$), thickness of the INL at both clivus ($D2_{cl}$ and $D2_{cr}$), thickness of the OPL at both clivus ($D3_{cl}$ and $D3_{cr}$), foveal thickness of the ONL ($D4_f$) and thickness of the ONL at both clivus ($D4_{cl}$ and $D4_{cr}$). The spatial resolution of the images is not exactly known for the images acquired from the Topcon device and varies from an acquisition to another one, which does not allow having an exact

equivalence between number of pixels and microns. We have therefore agreed with the physicians to use relative measures, normalized by the total thickness of the retina evaluated at the left clivus.

The graph in Fig. 17 shows the thickness measurements derived from manual (green) and automatic segmentations (red) of the sub-base as well as the segmentation of the whole database

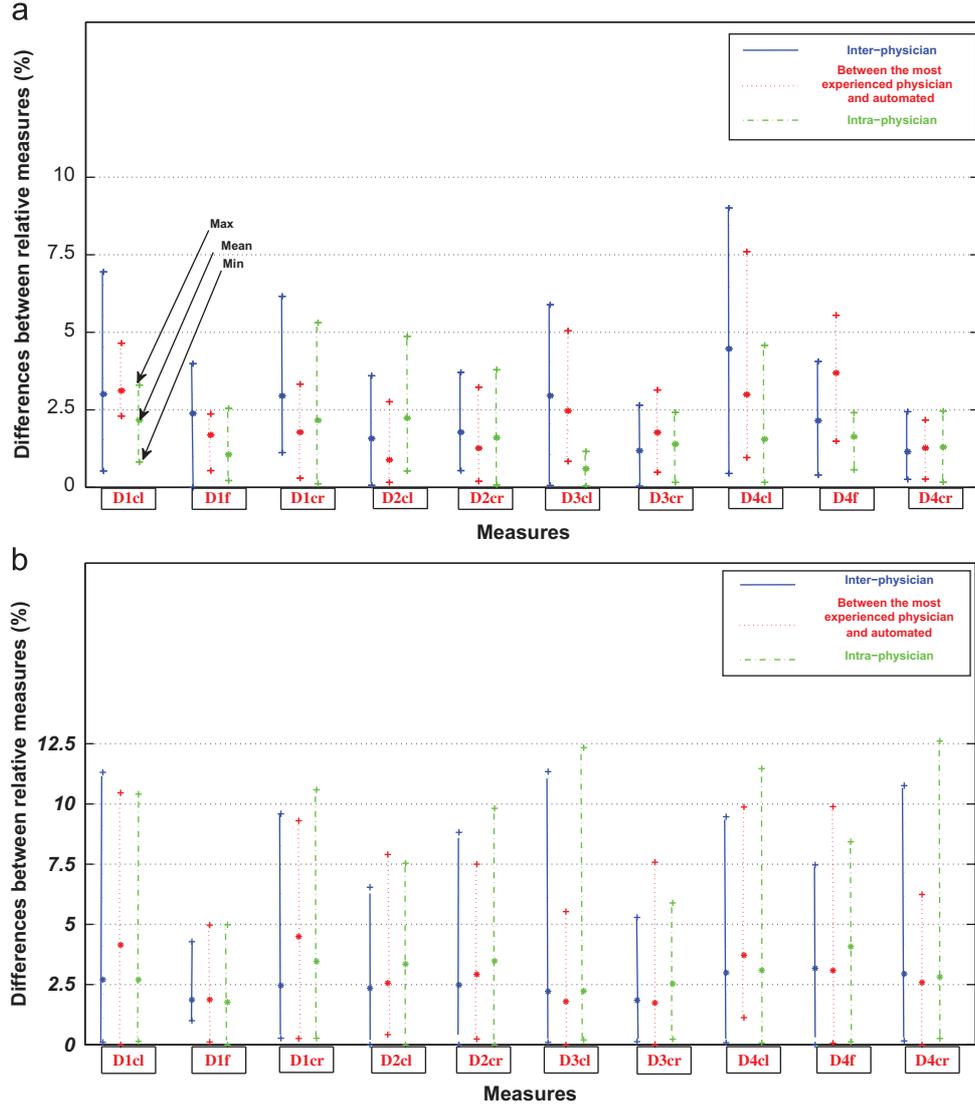


Fig. 18. Thickness measure differences: intra-physician, inter-physician and automated/reference physician (Topcon database (a) and Spectralis database (b)).

(blue). We represent for every measure the interval of the values (between two '+') and their average (*').

Fig. 17 indicates a very good agreement between measurements performed manually and those obtained by the automatic segmentation. It is interesting to note that the measures calculated on the entire database provide average values similar to those obtained on the whole validation base, with standard deviations slightly larger, which illustrates the representativeness of the chosen examples for the quantitative evaluation.

Let us denote by $D^{(i,j)}$ a thickness measure computed on the image i by the physician j or by our automatic method. We calculate for each measure the difference between the measurement made by the most experienced physician ($j=Ref$) and our automatic method ($|D^{(i,Ref)} - D^{(i,Auto)}|$, in red), inter-physician difference ($|D^{(i,Ref)} - D^{(i,j)}|$, in blue) and intra-physician difference ($|D_{t_1}^{(i,j)} - D_{t_2}^{(i,j)}|$, in green).

All measures are normalized by the total thickness, as explained before. Fig. 18 represents the interval of variation of the measure errors, as well as the mean values.

We notice that the differences between our results and those of the most experienced physician do not exceed the values of inter-physician variability for all measures except for $D3_{cr}$

(thickness of the OPL at right clivus) and $D4_f$ (foveal thickness of the ONL) for both databases, and for $D2_{cl}$ (thickness of the INL at left clivus) for the Spectralis database, however, with small differences. The intra-physician variability, although generally weaker than the inter-physician variability, is nevertheless important, and stronger than the automatic-physicians variability in some cases (INL thickness $D2$, foveal thickness $D1_f$ and $D1_{cr}$ for the Topcon database and foveal thickness $D1_{cr}$, INL thickness $D2_{cr}$, OPL thickness $D3_{cl}$ and ONL thickness $D4_{cr}$, $D4_{cl}$ at the clivus for the Spectralis database). This shows that our segmentation method provides sufficiently accurate measurements.

The thicknesses are measured at three points (two clivus and the fovea) and thus derive directly from the interface detection. It is therefore not surprising to observe a strong correlation between the thickness precision and the mean square errors (Tables 2 and 3). For instance, the INL thickness ($D2$) measure, which is based on GCL/INL and INL/OPL interfaces, has a greater inter and intra-physician variability than our method compared to the most experienced physician. This is confirmed by the MSE values in Tables 2–4. It is worth noting that results could be different if the physicians had made direct measurements of thickness on the images without preliminary delineation of the interfaces.

3.5. Discussion

Our algorithm has been applied on images coming from two different OCT devices and from more than 100 healthy subjects. Thus, there is a great variability in the database, in terms of image contrast, signal to noise ratio and retina anatomy. In spite of this, our method leads to satisfactory segmentation results and thickness measurements, even for the most difficult interfaces. In particular, the ONL/IS interface, which to our knowledge had not yet been addressed, is very accurately located, when compared to the physician delineations. The evaluation proves the reliability of our automatic segmentation method and its possible use for clinical diagnosis, since all validation measures discussed above are in the range of inter-physicians variability.

The proposed method was a basis for a preliminary study of variation of the morphology of foveal and perifoveal layers within a population of healthy subjects. We found that the profile of the fovea changed with axial length, with a relative increase of central ONL thickness which underlies a corresponding increase of foveal thickness. Independently of axial length, a large variability of the OPL/ONL complex was noted. On the opposite, the maximal retinal thickness was stable over a large range of axial length and had a low coefficient of variation. No gender-related differences were noted.

4. Conclusion

In this paper, a novel algorithm for segmenting eight layers in OCT retinal images has been proposed. Our algorithm is based mainly on global methods, which go beyond the limits of local methods by combining local information with more global features, ensuring a consistent behavior for all layers in all images. Although the method involves a number of parameters, they are either set to a constant value for all images, according to some anatomical knowledge (such as the number of classes in the k-means algorithm) or determined automatically for each case based on information extracted from the image to be processed (such as for the ONL/IS junction or for noise parameters). Hence no user interaction is required to tune the parameters. A quantitative evaluation has been performed on 14 images from the Topcon database and 16 images from the Spectralis database covering the usually observed variability, by comparing the results with those obtained manually. The quantitative results show a very good agreement and high correlation between measures. Also the distance between the manually segmented interfaces and the automated ones proves the accuracy of our segmentation method, including for interfaces that were not addressed before in the literature, such as the interface between the inner segments of photoreceptors (IS) and the outer nuclear layer (ONL). The inter and intra-physician variability was also studied to show that the results obtained by the proposed method are in the range of values of manual segmentation made by physicians. All these measures are necessary to assist ophthalmologists to analyze the structure of the retinal tissue and its variability. Such segmentation will help ophthalmologists to improve the accuracy of diagnosis of blinding retinal diseases, especially for the early detection of alteration of specific cell populations and/or substructures (for instance, photoreceptor outer segments). In the near future, this will be of interest for the follow-up of neuroprotective and regenerative treatments. A more distant yet reachable perspective is automated diagnosis.

Acknowledgments

The authors would like to thank the medical team of the Clinical Investigation Center of the Quinze-Vingts hospital for contributing to the manual segmentations. This work was partially funded by a CIFRE Grant from FOVEA and the French Ministry of Research (Contract CIFRE 897-2008).

References

- [1] D. Kiernan, W. Mieler, S. Hariprasad, Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems, *American Journal of Ophthalmology* 149 (1) (2010) 18–31.
- [2] M. Born, E. Wolf, *Interference and Diffraction with Partially Coherent Light*, fourth ed., Principles of Optics, Pergamon Press, United Kingdom, 1970 pp. 491–505.
- [3] Heidelbergengineering, Spectralis HRA+OCT.
- [4] Topcon, Topcon Europe Medical BV.
- [5] H. Maître (Ed.), *Image Processing*, ISTE Wiley, London, UK, 2008.
- [6] H. Ishikawa, D.M. Stein, G. Wollstein, S. Beaton, J.G. Fujimoto, J.S. Schuman, Macular segmentation with optical coherence tomography, *Investigative Ophthalmology & Visual Science* 46 (2005) 2012–2017.
- [7] D. Koozekanani, K. Boyer, C. Roberts, Retinal thickness measurements from optical coherence tomography using a Markov boundary model, *IEEE Transactions on Medical Imaging* 20 (9) (2001) 900–916.
- [8] M. Baroni, P.F.A.L. Torre, Towards quantitative analysis of retinal features in optical coherence tomography, *Medical Engineering & Physics* 29 (4) (2007) 432–441.
- [9] A.M. Bagci, M. Shahidi, R. Ansari, M. Blair, N.P. Blair, R. Zelkha, Thickness profile of retinal layers by optical coherence tomography image segmentation, *American Journal of Ophthalmology* 146 (5) (2008) 679–687.
- [10] D. Fernandez, H.M. Salinas, C.A. Puliafito, Automated detection of retinal layer structures on optical coherence tomography images, *Optics Express* 13 (25) (2005) 10200–10216.
- [11] M. Garvin, M. Abramoff, R. Kardon, S. Russell, W. Xiaodong, M. Sonka, Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search, *IEEE Transactions on Medical Imaging* 27 (10) (2008) 1495–1505.
- [12] M. Haeker, X. Wu, M. Abramoff, R. Kardon, M. Sonka, Incorporation of regional information in optimal 3-d graph search with application for intraretinal layer segmentation of optical coherence tomography images, *International Conference on Information Processing in Medical Imaging (IPMI)* (2007) 607–618.
- [13] M. Shahidi, Z. Wang, R. Zelkha, Quantitative thickness measurement of retinal layers imaged by optical coherence tomography, *American Journal of Ophthalmology* 139 (6) (2005) 1056–1061.
- [14] A. Mishra, A. Wong, K. Bizheva, D.A. Clausi, Intra-retinal layer segmentation in optical coherence tomography images, *Optics Express* 17 (26) (2009) 23719–23728.
- [15] M. Baroni, S. Diciotti, A. Evangelisti, P. Fortunato, A.L. Torre, Texture classification of retinal layers in optical coherence tomography, *International Federation for Medical and Biological Engineering (IFMBE)*, vol. 16, Boston, USA, 2007, pp. 847–850.
- [16] F. Rossant, I. Ghorbel, I. Bloch, M. Paques, S. Tick, Automated segmentation of retinal layers in OCT imaging and derived ophthalmic measures, in: *IEEE International Symposium on Biomedical Imaging (ISBI)*, Boston, USA, 2009, pp. 1370–1373.
- [17] J. Weickert, B.H. Romeny, M. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Transactions on Image Processing* 7 (3) (1998) 398–410.
- [18] C. Xu, J.L. Prince, Snakes, shapes, and gradient vector flow, *IEEE Transactions on Image Processing* 7 (3) (1998) 359–369.
- [19] R. Brown, P. Hwang, *The Discrete Kalman Filter, State-Space Modeling and Simulation, Introduction to Random Signals and Applied Kalman Filtering*, third ed., John Wiley & Sons, New York, 1997, pp. 214–220 (Chapter 5).
- [20] S. Geman, D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6) (1984) 721–741.
- [21] J. Besag, On the statistical analysis of dirty pictures, *Journal of the Royal Statistical Society B* 48 (1986) 259–302.
- [22] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (4598) (1983) 671–680.
- [23] R. Klein, B. Klein, S. Moss, D. DeMets, Inter-observer variation in refraction and visual acuity measurement using a standardized protocol, *Ophthalmology* 90 (11) (1983) 1357–1366.
- [24] D. Hartong, E. Berson, T. Dryja, Retinitis pigmentosa, *Lancet* 368 (9549) (2006) 1795–1809.
- [25] A. Zijdenbos, B. Dawant, R. Margolin, A. Palmer, Morphometric analysis of white matter lesions in MR images: method and validation, *IEEE Transactions on Medical Imaging* 13 (4) (1994) 716–724.