IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 25, 2017, accepted December 28, 2017, date of publication January 5, 2018,
date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2789526

Automated Detection and Measurement of Corneal Haze and Demarcation Line in Spectral-Domain Optical Coherence Tomography Images

AHMAD R. DHAINI[1], (Member, IEEE), MANAL CHOKR[1], SARA MARIA EL-OUD[2], MAAMOUN ABDUL FATTAH[3], AND SHADY AWWAD[3]

[1]Department of Computer Science, American University of Beirut, Beirut 1107 2020, Lebanon
[2]Department of Physics, The George Washington University, Washington, DC 20052, USA
[3]Department of Ophthalmology, American University of Beirut Medical Center, Beirut 1107 2020, Lebanon

Corresponding author: Ahmad R. Dhaini (ahmad.dhaini@aub.edu.lb)

This work was supported by the Lebanese National Council for Scientific Research Grant Research Program.
(Ahmad R. Dhaini and Shady Awwad contributed equally to this work.)

ABSTRACT Keratoconus is a progressive eye disease that may lead to significant loss of visual acuity. Corneal cross-linking (CXL) is a surgical procedure that halts the progression of keratoconus. One commonly used clinical indicator of CXL success, albeit being an indirect one, is the presence and depth of stromal demarcation line. In addition, corneal haze beyond the demarcation line can be an ominous sign of loss of corneal transparency, which is a much dreaded side effect of CXL. To date, ophthalmologists evaluate the presence and depth of the demarcation line and grade corneal haze using slit lamp biomicroscopy and/or optical coherence tomography (OCT). Interpreting the output of the former is very biased at best, while analyzing the information presented by the latter is time consuming, potentially error prone, and observer dependent. In this paper, we propose the first method that employs image analysis and machine learning to automatically detect and measure corneal haze and demarcation line presence and depth in OCT images. The automated method provides the user with haze statistics as well as visual annotation, reflecting the shape and location of the haze and demarcation line in the cornea. Our experimental results demonstrate the efficacy and effectiveness of the proposed techniques vis-a-vis manual measurements in a much faster, repeatable, and reproducible manner.

INDEX TERMS Medical image analysis, optical coherence tomography, keratoconus, crosslinking, corneal haze.

I. INTRODUCTION

The cornea is the transparent front part of the eye and the outer wall of the anterior chamber, with the iris and pupil forming the posterior wall. It accounts for approximately two-thirds of the eye's total optical power. As illustrated in Fig. 1, the cornea can be observed using spectral-domain optical coherence tomography (SD-OCT), which is a remarkable tool for capturing more retinal and corneal data in less time and at a higher axial image resolution than the older time-domain OCT (TD-OCT) [1].

As illustrated in Fig. 2, Keratoconus is a common ectatic disorder that is characterized by progressive thinning and steepening of the cornea, which leads to apical scarring and eventual loss of visual acuity [2]. Consequently, corneal collagen cross-linking (CXL) has been applied as the most successful treatment modality for halting or slowing the progression of Keratoconus. This surgical procedure increases

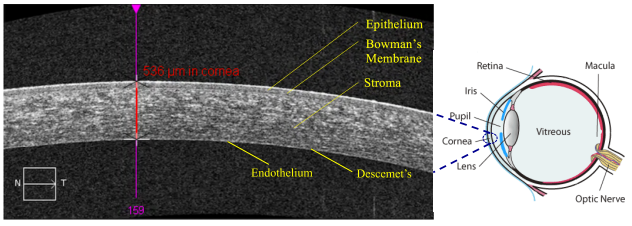


FIGURE 1. Anatomy of the eye and SD-OCT scan of the cornea.

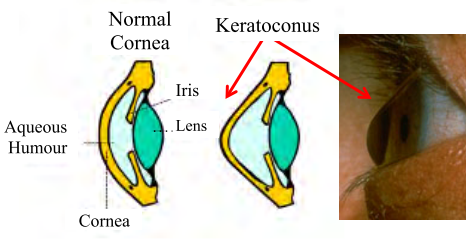


FIGURE 2. Normal cornea vs. Keratoconus with cornea protrustion.

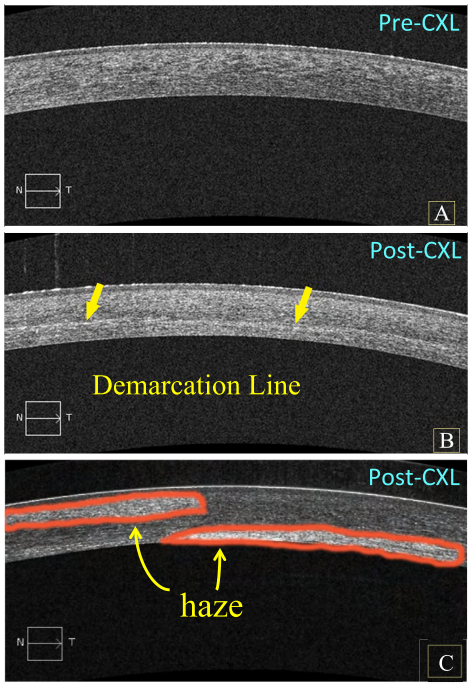


FIGURE 3. OCT scans depicting [A] a pre-CXL clear cornea; [B] post-CXL demarcation line (a bright line that follows the curvature of the cornea); and [C] post-CXL corneal haze (a stromal region brighter than the rest of the stroma).

the biomechanical stiffening and biochemical resistance of the cornea [3].

Corneal stromal haze has been generally observed on clinical examination as the most common side-effect of CXL, potentially affecting vision [4]. As illustrated in Fig. 3, it can be defined as a cloudy or opaque appearance of the cornea and is most often caused by inflammatory cells and other molecules activated during trauma, infection, or eye procedures and surgeries such as corneal cross-linking for Keratoconus eyes and kerato-refractive surgeries like photo-refractive keratectomy. Stromal haze leads to forward scatter that can potentially degrade vision. Therefore, objective observation and tracking of corneal haze is very valuable to ophthalmologists [4]. On the other hand, measuring corneal cross-linking success relies on the presence and depth of a thin, linear, hazy, hyper-reflective line, which typically forms between the cross-linked and non-cross stromal tissue, dubbed the "demarcation line". Evaluating the presence and measuring the depth of the demarcation line on SD-OCT is a useful tool for clinicians to assess the depth of the cross-linking effect, and to compare the efficacy of the many technical variations of the procedure [5]–[9]. However, this process is tedious and time consuming, as the clinician or experienced operator has to go through many OCT sections of the cornea to determine and measure the demarcation line depth using a digital caliper, among other reactive and non-specific hazy lines that can mimic the actual demarcation line [10], [11]. More importantly, the whole evaluation is in part subjective and operator-dependent, with intra-observer repeatability and inter-observer reproducibility yet to be investigated. Therefore, quantification of corneal haze and demarcation line by means of an objective automated method that can adequately assess the overall amount of opacification would lead to better understanding of the phenomenon.

Previous studies have employed automated detection and segmentation techniques, yet on retinal OCT images [12]–[16]. However in recent years, automated segmentation of corneal layers using OCT has been gaining attention in the research community, due to the importance of information related to the curvature and thickness of different corneal layers after procedures such as refractive surgery [17]–[22]. Nevertheless, none of the previous work addressed the problem of detecting and measuring corneal haze and demarcation line in corneal OCT images for Keratoconus patients after cross-linking. In this paper, we propose the first solution that offers the measurements needed for corneal haze and demarcation line examination in a fast, automated and objective manner with good agreement compared to manual measurements, while being more repeatable and reproducible. Using image analysis and machine learning, our scheme is able to automatically 1) validate OCT scans; 2) detect corneal boundaries; 3) detect and classify corneal haze; and 4) detect and measure the depth of the demarcation line, such that all these are performed in less than 1 second [23]. For the reader's convenience, we summarize the most relevant notations used in this paper in Table 1; the rest are defined in the text.

The remainder of the paper is organized as follows. Section II discusses the OCT image validation process. Section III describes the module responsible for detecting and classifying corneal haze. In Section IV, we present the module responsible for detecting and measuring the depth of the demarcation line. In Section V, we present our experimental results and analysis. In Section VI, we conclude our work and discuss future extensions.

II. OCT IMAGE VALIDATION USING MACHINE LEARNING

Our OCT imaging system is a Cirrus High-Definition SD-OCT (HD-OCT) scanner (Carl Zeiss Meditec, Dublin, CA, USA) with a wavelength of 840 nm (which is the default setting), and a scanning speed up to 27,000 axial scans per second. Higher wavelengths, such as 1310 nm, can alternately be more fit for images of the anterior chamber angle as they better penetrate the sclera. However, for corneal imaging, both wavelengths perform fair equally well.

TABLE 1. Most relevant notations used in the paper.

Notation	Description
$\mathbf{M}$	Matrix representing the OCT image with dimensions $w \times h$
θ	Brightness parameter θ used in the *Thresholding* technique
$p_{i,j}$	Pixel intensity value at position $[i, j]$ in $\mathbf{M}$
$c(x)$	Function depicting the coordinates of curve c (e.g., $top(x)$, $bottom(x)$, $mid_1(x)$, $mid_2(x)$, $dem(x)$)
A_{cor}	Total area of the cornea
$A_{cor-\mathcal{H}}$	Total area of cornea without haze
I_{cor}	Total corneal pixel intensity
$I_{cor-\mathcal{H}}$	Total pixel intensity of cornea without haze
$\overline{I_{cor}}$	Average corneal pixel intensity
T_{cor}	Thickness of the cornea
r	Corneal region $r = \begin{cases} 1, & \text{Anterior stroma} \\ 2, & \text{Mid stroma} \\ 3, & \text{Posterior stroma} \end{cases}$
e	Thickness of the Epithelium's bright surface
C_r	The set of contours in corneal region r such that $C_r = \{c_{r,1}, c_{r,2}, ..., c_{r,i}, ..., c_{r,\lvert C_r \rvert}\}$
$\mathcal{H}_r$	The set of haze in corneal region r: $\mathcal{H}_r = \{\hbar_{r,1}, \hbar_{r,2}, ..., \hbar_{r,i}, ..., \hbar_{r,\lvert \mathcal{H}_r \rvert}\}$
$A(\hbar)$	Area of haze $\hbar$
$I(\hbar)$	Pixel intensity of haze $\hbar$
$\overline{I(\hbar)}$	Average pixel intensity of haze $\hbar$
$S(c)$	Score of curve $c(x)$
$\mathfrak{T}$	Outliers' demarcation line score threshold
$\mathfrak{s}$	Minimum demarcation line score

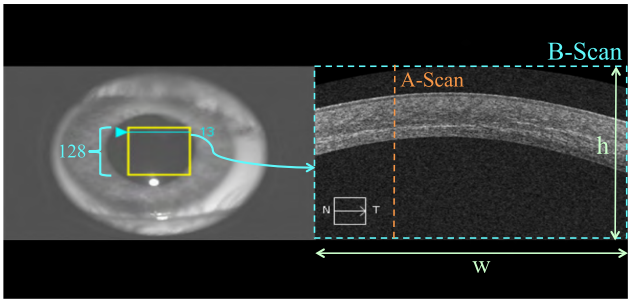


FIGURE 4. An OCT B-scan (i.e., 13 out of 128) taken using the HD-OCT scanner in our lab.

The corneal cube mode uses a 128 × 512 horizontal raster protocol. As depicted in Fig. 4, it generates a $4mm \times 4mm$ grid by performing 128 horizontal 2-dimensional (2D) B-scans, each comprising 512 A-scans, of 5 microns axial resolution. The transverse resolution is 8 microns. The OCT B-scans used by the automated solution have a width $w = 540$ pixels, a height $h = 280$ pixels, and a length $l = 4000$ μm (i.e., 4-mm cube). Thus, the *pixel-to-*μm conversion ratio can be obtained by l/w. Each OCT B-scan (also referred to in the text as "cut") is modeled as a matrix $\mathbf{M}$ of dimensions $w \times h$. The matrix consists of pixel intensity values denoted as $p_{i,j}$ (also referred to as $[i,j]$; e.g., $\mathbf{M}[1, 4] = p_{1,4}$), where $i = 0, \ldots, h-1$ and $j = 0, \ldots, w-1$. For an 8-bit greyscale image, $0 \leq p_{i,j} \leq 255$.

$$\mathbf{M} = \begin{bmatrix} p_{0,0} & \cdots & p_{0,w-1} \\ \vdots & \ddots & \vdots \\ p_{h-1,0} & \cdots & p_{h-1,w-1} \end{bmatrix}.$$

For analyzing corneal haze, only one clear central cut is evaluated, as clinically, only central corneal haze has visual implications [4]. However, detecting and measuring the depth of the demarcation line requires evaluating the cut in which the demarcation line appears the "clearest" (the demarcation line varies significantly across scans). Thus, all 128 cuts must be processed to select the "best" cut, which would contain the "suggested" demarcation line (details are provided in Section IV).

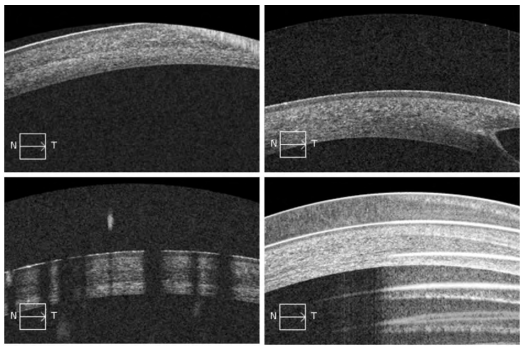


FIGURE 5. Examples of invalid/damaged OCT cuts.

A. ELIMINATION OF INVALID CUTS

As shown in Fig. 5, some OCT images may have some damaged cuts, or ones with artifacts. However, this should not disqualify the whole image, since other cuts of the same image may be valid for haze analysis. Thus, only the damaged scans must be eliminated so as not to falsify the reported output. To achieve so, we apply machine learning to classify each cut as either *valid* or *invalid*.

1) DATA PREPARATION AND FEATURE SELECTION

Our data is composed of 8,064 scans extracted from 63 OCT images in our data bank; each scan is represented by 36,450 features extracted using the image-processing library OpenCV [24]. The width and height of each scan are reduced to half so as to decrease the machine training time. Each scan is annotated by class "0" as invalid, or class "1" as valid. This annotation is performed manually and diligently twice by experienced operators and on two separate occasions to ensure accurate learning.

2) DATA COMPRESSION

Due to the large feature space, feature reduction is performed to guarantee good generalization of the learnt models and reduce the training time. To achieve the latter, Principal Component Analysis (PCA) is employed [25], such that our dataset is split into two subsets: a *training* set consisting of 80% of the images, and a *testing* set consisting of the remaining 20%. PCA is then applied on the training set, and the first m Eigen vectors that preserve 95% of the data variance are retrieved. In recent studies, validation sets have no longer been used; thus, we do not consider them. Instead, we employ the 10-folds cross-validation technique, which has been the commonly used alternative [26]. Since PCA is cubic in the number of original dimensions (thus, does

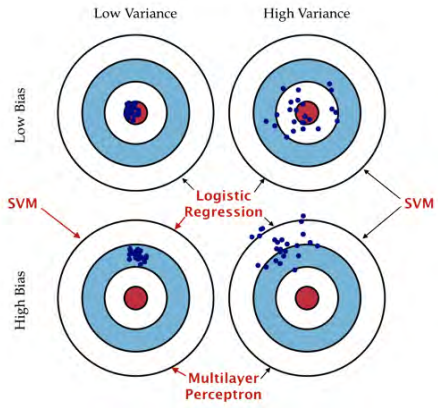


FIGURE 6. Illustration of bias and variance and the corresponding training models typically used.

not scale well), we employ Information Gain (InfoGain) to obtain a list of features ranked according to their influence on the prediction accuracy [27]. Due to the large dataset size, the top 5^{th} is used as our cut-off (i.e., 8,192 features per instance) as ranked by InfoGain, which is observed to perform well when later combined with PCA. Subsequently, PCA is applied for further dimensionality reduction, which suffices to have 172 features for representing each instance.

3) BIAS VARIANCE ANALYSIS

To improve data fitting and train highly accurate models, the different sources of error leading to bias and variance are identified. We denote the class to-be-predicted as y and the vector of features as X. We assume there is a function f such that $y = f(X)$. Hence, our goal is to estimate a function $\hat{f}(X)$ that emulates $f(X)$. Thus, for a feature vector X, the expected out-of-sample squared prediction error $Err(X)$ is given by:

$$Err(X) = \underbrace{(E[\hat{f}(X)] - f(X))^2}_{bias^2} + \underbrace{E[(\hat{f}(X) - E[\hat{f}(X)])^2]}_{variance}. \quad (1)$$

To identify whether the data suffers from high bias or high variance (illustrated in Fig. 6), we plot in Fig. 7 the curve of E_{in}, the measured error in the training data, and E_{val}, the measured error in the 10-folds cross-validation data. As observed, experiments show that our data suffers from high bias.[1] Thus, we apply Logistic Regression [28], Support Vector Machines (SVM) [29] and Multilayer Perceptron [30] to our data set; these powerful models have been commonly used for accurate classification of datasets suffering from high bias.

4) EMPLOYED MACHINE LEARNING MODEL

The experiments were performed on the training data using the machine learning library Weka [31]. By applying the 10-folds cross-validation technique, we tune-and-experiment with the hyperparameters for each of the considered models.

[1] In general, the learning curve of high bias data is identified by a small gap between E_{in} and E_{val}, and high values for both E_{in} and E_{val}; whereas the learning curve of high variance data is identified by a constant decrease of E_{val}, and a value of E_{in} that is much lower than E_{val}.

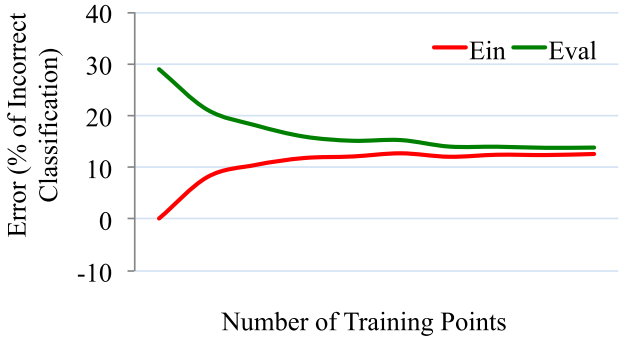


FIGURE 7. Learning curve of the training data using Logistic Regression, which corresponds to high bias.

TABLE 2. Best performance measurements of applied models.

	Accuracy	F-Measure	Precision
Logistic Regression	0.8620	0.9236	0.9835
Support Vector Machines	0.8628	0.9250	0.9974
Multilayer Perceptron	0.8504	0.9151	0.9508

For model evaluation, the accuracy, precision, recall and F-measure metrics are obtained as follows:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + FP + FN + TN}, \\ \text{Precision} &= \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \\ \text{F-Measure} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}; \end{aligned} \quad (2)$$

where, classifying a valid cut is taken to be positive P, and a damaged/invalid cut to be negative N. Hence, TP, FP, FN, and TN are the number of true positive, false positive, false negative, and true negative instances, respectively. Our aim is to minimize the number of invalid cuts that are classified as valid cuts (i.e., FP instances), rather than to minimize the number of incorrectly classified instances. This is due to the fact that valid cuts would not affect the overall results of an OCT image, whereas one damaged cut would detect imprecise demarcation line(s), and thus provide wrong output. Therefore, the machine learning model is evaluated based on the largest precision value.

Table 2 summarizes the best performance of each applied model. As observed, all of the models are able to attain good performance after several rounds of tuning the hyperparameters. However, since our objective is to minimize the number of FP instances, we adopt the SVM model, as it is able to provide only 14 FP instances among all cuts of all images.

The hyper-parameters of the chosen SVM model were obtained as follows: *polynomial kernel* with degree 1; C value equal to 1; *epsilon* equal to $1.0E^{-10}$ and *tolerance* equal to 0.01. The selected SVM model resulted in only 3 FP instances for all cuts, with a precision equal to 0.9978, F-measure equal to 0.9366, and Accuracy equal to 0.8828; which are considered to be excellent, since the impact of the 3 FP instances on determining the consistency of the demarcation line would be negligible. Subsequently, the model is

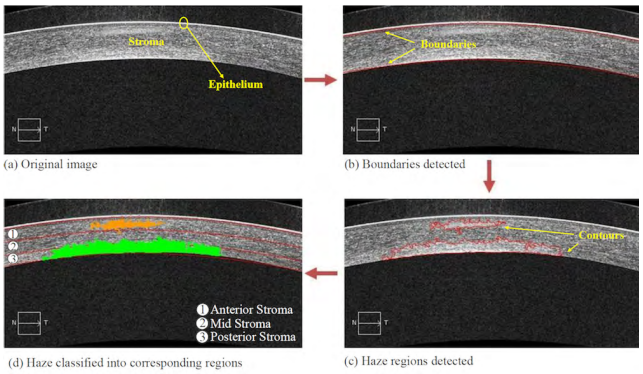


FIGURE 8. Steps taken to detect and classify haze.

saved and embedded in the automated scheme as part of the OCT image filtering process, which ensures that only valid cuts are used for analysis.

III. CORNEAL HAZE DETECTION AND CLASSIFICATION

To offer automated objective detection and classification of corneal haze, we employ custom image analysis algorithms and make use of methods from OpenCV. This process is achieved via the following steps:

1) Defining the corneal region: Fig. 8b;
2) Computing corneal image properties such as area(s), intensity, and average brightness;
3) Splitting the cornea into three regions: *anterior stroma*, *mid stroma*, and *posterior stroma*;
4) Detecting haze regions (contours): Fig. 8c;
5) Aggregating "close" haze regions;
6) Classifying haze into corresponding region(s): Fig. 8d.

Details about the foregoing steps are provided in the following sub-sections.

A. DEFINING THE CORNEAL REGION

Defining the corneal region in the image (i.e., where haze must be detected) requires finding the coordinate vectors of the top and bottom boundaries. Once identified, the top boundary must be shifted down so that the Epithelium's surface is not considered as part of the analyzed region (the epithelium typically has a very bright surface that may greatly alter the average brightness calculated). Each of the boundary vectors is found using a different method. This is due to the fact that the upper section of the image has a uniform shape due to the high continuous intensity of the Epithelium's bright surface, whereas the lowest section does not always have this uniformity; thus, it may not have a continuous curvature.

1) FINDING THE TOP BOUNDARY

To find the top boundary, the image is first processed using the *Thresholding* technique [32], which is based on a brightness parameter θ, such that pixels with values smaller than θ are referred to as *background* pixels, and those with values larger than θ are referred to as *foreground* pixels. Applying *Thresholding* generates a binary version of the matrix, denoted $\mathbf{M}_B$, by setting each pixel $p'_{i,j}$ as follows:

$$p'_{i,j} = \begin{cases} 1, & p_{i,j} \geq \theta \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

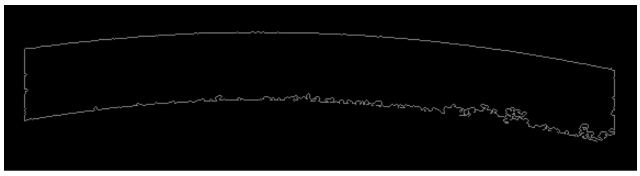


FIGURE 9. Image after applying *Thresholding* and *Contouring*.

The new image enables to extract the *contrours* (i.e., borders) of the cornea. To achieve so, we employ Suzuki's algorithm [33] for obtaining external contours (i.e., a contour that is not encompassed within any other contour). However, the contouring function requires a "connected" pixel region (in our case, it's the corneal region) enclosed in another "connected" region (in our case, it's the "black+legend" region). Since this is not the case for our image, we augment the image with black borders of arbitrary thickness to the left and right of $\mathbf{M}_B$. The border thickness can be of any value as long as the cornea can be seen as a *complete* object. Using contouring and after trimming out the smaller corneal contours (i.e., of size smaller than $(w \times h)/\beta$; we set $\beta = 10$), we obtain the image shown in Fig. 9. Finally, after eliminating the extra thickness borders that we added, we extract the top boundary vector $top(x)$ by searching for the first pixel at the y-axis that has a non-zero value. The vector $top(x)$ is in fact a function that, for a given x-coordinate, it returns the y-coordinate falling on the top boundary. To ignore the Epithelium's surface, every y-coordinate in the vector $top(x)$ is increased by the thickness of the Epithelium's surface e, which we measure dynamically by applying a technique similar to the one used for detecting a demarcation line (typically, the average value of e is about 3 pixels $\approx 22\mu m$).

2) FINDING THE BOTTOM BOUNDARY

As observed in Fig. 9, since the bottom boundary does not have a well-defined curvature, applying the technique used for the top boundary to find it will not achieve the desired results. Thus, we apply a different technique using the original image matrix $\mathbf{M}$ to capture it. Here, we assume the following:

- The bottom boundary of the cornea follows the same curvature as the top boundary (which is normally the case for almost all images);
- The change in pixel intensity around the top and bottom boundaries is the highest among all other regions.

Based on these assumptions, we employ a new technique, which performs the following steps:

1) For every pixel curve $c(x)$ starting from $top(x)$ downwards, compute the change in image intensity at a distance y_d above and below $c(x)$. The change in image

intensity is computed as the ratio $\overline{I_r}$ of intensities above and below $c(x)$, denoted $\overline{I_a}$ and $\overline{I_b}$ respectively, as follows:

$$\overline{I_a} = \sum_{x=0}^{w} \sum_{y=c(x)-y_d}^{c(x)-1} M[x,y],$$
$$\overline{I_b} = \sum_{x=0}^{w} \sum_{y=c(x)+1}^{c(x)+d} M[x,y],$$
$$\overline{I_r} = \frac{\overline{I_a}}{\overline{I_b}}; \tag{4}$$

2) Find the minimum ratio $\overline{I_r}$ (thus, the curve with the highest pixel intensity value) and mark its y-position as y_{shift};
3) The bottom boundary curve is thus represented by $bottom(x) = top(x) + y_{shift}$.

B. CORNEAL IMAGE PROPERTIES

With the corneal region identified, we calculate the corneal area, thickness and brightness, which are three important properties for the detection and classification of corneal haze, as well as for relevant statistical data needed in longitudinal studies.

1) CORNEAL THICKNESS

Corneal thickness is used to analyze corneal thinning over time. It is computed as follows:

$$T_{cor} = y_{shift} - e. \tag{5}$$

2) CORNEAL AREA

The corneal area is computed by counting the number of pixels between the top and bottom boundaries as follows:

$$A_{cor} = \sum_{x=0}^{w} \sum_{y=top(x)}^{bottom(x)} 1. \tag{6}$$

3) CORNEAL INTENSITY

Similar to the corneal area, the total corneal intensity and average corneal intensity are computed as follows:

$$I_{cor} = \sum_{x=0}^{w} \sum_{y=top(x)}^{bottom(x)} M[x,y]; \tag{7}$$
$$\overline{I_{cor}} = \frac{I_{cor}}{A_{cor}}. \tag{8}$$

C. SPLITTING THE CORNEA INTO THREE REGIONS

The cornea comprises three "depth" regions in which haze may appear (i.e., anterior stroma, mid stroma, and posterior stroma). Identifying the corneal region, in which detected haze may exist is key for classifying the identified haze, which may affect clinical decision-making. The foregoing stroma regions are normally of equal size and follow the curvature of the top and bottom boundaries. By dividing the cornea into three equal regions ($T_{cor}/3$), the two curves defining the forgoing regions are thus given using the following functions:

$$mid_1(x) = top(x) + T_{cor}/3;$$
$$mid_2(x) = top(x) + 2 \times T_{cor}/3. \tag{9}$$

D. DETECTING HAZE REGIONS

Haze appears in the cornea as dense regions of bright pixel intensity. To find these regions, we use the contouring technique used for finding the top boundary. However, here, contouring is applied within the corneal region only. Thus, we define a new image matrix $\mathbf{M}_C$ consisting of pixels $p''_{i,j}$ inside the cornea only; these are computed as follows:

$$p''_{i,j} = \begin{cases} p_{i,j}, & top(i) \le j \le bottom(i) \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

The new image created from $\mathbf{M}_C$, is then blurred to obtain smoother boundaries for the haze, and to fade away any possible noisy fluctuations in high pixel intensity regions [24]. To extract regions of "brighter" pixel intensity, we apply the *Thresholding* technique. However, since haze regions may vary in terms of pixel intensity, we compute a dynamic threshold based on the *Isodata* algorithm [32]. The latter takes the histogram $H(i)$ of the image pixels' intensities as input and provides the threshold θ as output. Our region(s) of interest have high intensity values; thus, we extract the histogram of the image such that it covers the region between the average intensity and the maximum intensity as follows:

$$H(i) = n_i, \quad i = \overline{I_{cor}}, \overline{I_{cor}} + 1, \overline{I_{cor}} + 2, \ldots, p_{max}; \tag{11}$$

where, n_i is the number of corneal pixels with intensity value i, and p_{max} is the maximum pixel intensity value (e.g., 255 for 8-bit greyscale images). The *Isodata* algorithm requires setting an initial threshold value θ_0 and then computing the mean pixel intensity values below and above θ_0. In our customized version of the *Isodata* algorithm, θ_0 is computed as follows:

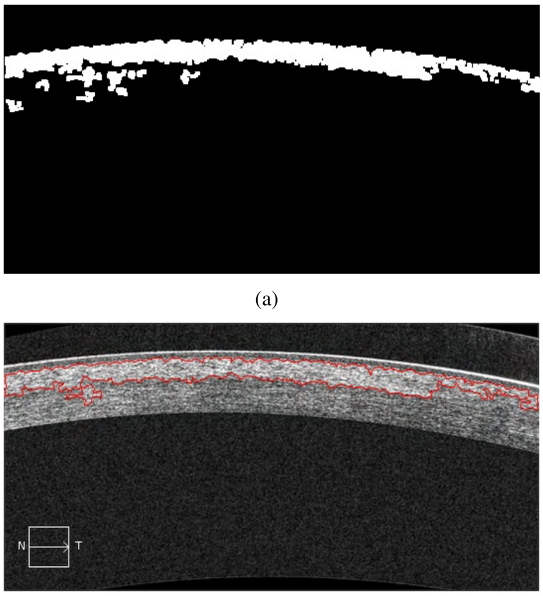
$$\theta_0 = \overline{I_{cor}} + \frac{p_{max} - \overline{I_{cor}}}{2}. \tag{12}$$

Subsequently, these parameters are used to obtain new values of θ_k ($k = 1, 2, 3, \ldots$) until θ_k is converged into a "steady" value (i.e., until $\theta_k = \theta_{k-1}$). The algorithm then stops and the output threshold value θ is set to the last θ_k value. We also compute θ_k, and the mean *background* and *foreground* pixels values ($m_{b,k-1}$ and $m_{f,k-1}$, respectively) as follows:

$$\theta_k = (m_{b,k-1} + m_{f,k-1})/2;$$
$$m_{b,k} = \frac{1}{\Lambda_{b,k}} \sum_{i=\overline{I_{cor}}}^{\theta_{k-1}} i \times H(i); \quad m_{f,k} = \frac{1}{\Lambda_{f,k}} \sum_{i=\theta_{k+1}}^{\overline{I_{cor}}} i \times H(i), \tag{13}$$

where,

$$\Lambda_{b,k} = \sum_{i=\overline{I_{cor}}}^{\theta_{k-1}} H(i); \quad \Lambda_{f,k} = \sum_{i=\theta_{k-1}}^{\overline{I_{cor}}} H(i).$$



(a)

(b)

FIGURE 10. a) Image after applying thresholding and morphology operators; b) Image with the haze detected and outlined.

To eliminate noise and join distant pixels that belong to the same haze, we apply *morphology operators* [24]. Finally, we apply *Contouring* to extract a set C representing all external contours that have an area greater than half of the average area of all contours. This is done to ensure that only the contours that dominate the image are extracted. The set C then represents the list of detected corneal haze regions as outlined in Fig. 10. Each contour $c \in C$ may span across multiple corneal regions. Moreover, two or more contours may belong to the same haze region if they are "close enough" to each other. Thus, contours must be first aggregated to form a single haze (where applicable), which then gets classified based on the area it belongs to the most (i.e., where the majority of its pixels is concentrated).

E. HAZE AGGREGATION

Since the aggregation process must occur per corneal region, our aim is to construct a set C_r per region $r = 1, 2, 3$ such that, $c_r \in C_r$ if c_r has the majority of its pixels in region r. Consequently, using C_r, we construct the set of haze per region r, denoted $\mathcal{H}_r = \{\hbar_{r,1}, \hbar_{r,2}, \ldots, \hbar_{r,|\mathcal{H}_r|}\}$, such that a haze set $\hbar_{r,k} \in \mathcal{H}_r$ comprises one or more contours. The construction of $\mathcal{H}_r$ is performed via the iterative Algorithm 1. Here, based on medical observation and verification, the default value of d_{max} is set to 50 pixels. However, please note that our program provides the user with the option to adjust the value of d_{max} dynamically to obtain more coarse-grained or fine-grained aggregation results.

By applying the new algorithm, we obtain the set of all haze $\mathcal{H} = \bigcup_{r=1}^{3} \mathcal{H}_r$, of size $|\mathcal{H}|$, as well as the set of all haze in region r, denoted as $\mathcal{H}_r$, of size $|\mathcal{H}_r|$.

Algorithm 1 Haze Aggregation Algorithm

Initialize: $r = 1$;
repeat
  Step 1:
  Construct a binary matrix X_r of dimensions $|C_r| \times |C_r|$, indicating if any two contours $c_{r,i}$ and $c_{r,j}$ in region r are to be combined as follows:

$$X_r[i,j] = X_r[j,i] = \begin{cases} 1, & d(c_{r,i}, c_{r,j}) \leq d_{max} \\ 0, & \text{otherwise} \end{cases}$$

  where, $d(c_{r,i}, c_{r,j})$ is the (smallest) measured distance between contours $c_{r,i}$ and $c_{r,j}$; and d_{max} is the maximum distance allowed between any two contours. $X_r[i,i] = 1$;
  Step 2:
  Construct $C_r = \{c_{r,1}, c_{r,2}, \ldots, c_{r,i}, \ldots, c_{r,|C_r|}\}$;
  Step 3:
  Initialize: $k = 1$;
  repeat
    Initialize: $\hbar_{r,k} = \emptyset$;
    for $c_{r,i} \in C_r$ **do**
      for $c_{r,j} \in C_r \mid X_r[i,j] = 1, c_{r,j} \notin \hbar_{r,k}$ **do**
        $\hbar_{r,k} = \hbar_{r,k} \cup \{c_{r,j}\}$;
        $C_r = C_r - \{c_{r,i}\}$;
        repeat 12–14 for the added $c_{r,j}$
    $\mathcal{H}_r = \mathcal{H}_r \cup \{\hbar_{r,k}\}$;
    $k = k + 1$;
  until $C_r = \emptyset$;
  $r = r + 1$;
until $r = 3$;

F. HAZE CLASSIFICATION AND PROPERTIES

Haze classification requires identifying the corneal region in which the haze exists (which can be easily obtained based on the value of r for every $\hbar_{r,k} \in \mathcal{H}_r$), as well as collecting some salient haze properties such as the area of the haze in each region, the overall corneal haze intensity, and the percentage of haze brightness in comparison to the rest of the cornea.

To collect these properties, we make use of a function $b(x, y, c)$ provided by OpenCV, which for a given pixel of coordinates (x, y) and a contour c, it returns whether the pixel belongs to this contour or not, as follows:

$$b(x,y,c) = \begin{cases} 1, & (x,y) \text{ inside } c \\ 0, & (x,y) \text{ on } c \\ -1, & \text{otherwise} \end{cases} \tag{14}$$

Using $b(x, y, c)$, we compute the area of haze $\hbar_{r,k}$ in region r by counting the number of pixels as follows:

$$A(\hbar_{r,k}) = \sum_{x=0}^{w} \sum_{y=a}^{b} \begin{cases} 0, & b(x,y,c_{r,i}) = -1, \quad \forall c_{r,i} \in \hbar_{r,k} \\ 1, & \text{otherwise} \end{cases} \tag{15}$$

where,

$$a = \begin{cases} top(x), & r = 1 \\ mid_1(x), & r = 2 \\ mid_2(x), & r = 3 \end{cases} \quad b = \begin{cases} mid_1(x), & r = 1 \\ mid_2(x), & r = 2 \\ bottom(x), & r = 3 \end{cases}$$

Thus, the total area of haze $\hbar_k$ across all regions can be obtained as follows:

$$A(\hbar_k) = \sum_{r=1}^{3} A_r(\hbar_{r,k}). \tag{16}$$

Subsequently, the total and average pixel intensity of haze $\hbar_k$ can be computed as follows:

$$I(\hbar_k) = \sum_{x=0}^{w}\sum_{y=0}^{h} \begin{cases} 0, & b(x, y, c_i) = -1, \ \forall c_i \in \hbar_k \\ M[x, y], & \text{otherwise} \end{cases}$$
$$\overline{I(\hbar_k)} = I(\hbar_k)/A(\hbar_k). \tag{17}$$

To compare the pixel intensity (i.e., percentage of brightness) of the overall haze with the rest of the cornea, we compute the total area of corneal haze $A(\mathcal{H})$, and the area of the cornea without haze $A_{cor-\mathcal{H}}$ as follows:

$$A(\mathcal{H}) = \sum_{x=0}^{w}\sum_{y=0}^{h} \begin{cases} 0, & b(x, y, c_i) = -1, \quad \forall c_i \in C \\ 1, & \text{otherwise} \end{cases}$$
$$A_{cor-\mathcal{H}} = A_{cor} - A(\mathcal{H}). \tag{18}$$

Similarly, we compute the total haze pixel intensity and total corneal intensity with and without haze as follows:

$$I(\mathcal{H}) = \sum_{x=0}^{w}\sum_{y=0}^{h} \begin{cases} 0, & b(x, y, c_i) = -1, \quad \forall c_i \in C \\ M[x, y], & \text{otherwise} \end{cases}$$
$$I_{cor-\mathcal{H}} = I_{cor} - I(\mathcal{H}). \tag{19}$$

Finally, the percentage of brightness of a haze $\hbar_k$ to the rest of the cornea with and without haze are obtained by $\frac{I(\hbar_k)}{I_{cor}} \times 100$, and $\frac{I(\hbar_k)}{I_{cor-\mathcal{H}}/A_{cor-\mathcal{H}}} \times 100$, respectively.

IV. DEMARCATION LINE DETECTION AND DEPTH MEASUREMENT

The demarcation line does not normally appear in corneal hazy images. However when it does, it indicates that the eye is healing as expected. The process of detecting it requires identifying "the brightest" curves that could potentially be the demarcation line, by computing a curve score $S(c)$ that is greater than a minimum score, denoted $\mathfrak{s}$; and also observing the OCT cut in which the demarcation line appears the "clearest", such that it consistently appears across valid cuts (e.g., in at least 15% of the cuts).

To account for the fact that the demarcation line may not exist at the same depth across different cuts, we consider a demarcation line as being the same across cuts if it appears at a fluctuating depth of a certain small range obtained statistically from the available images. Subsequently, the detected lines that get a score $S(c) > \mathfrak{s}$, are of the highest group of scores (i.e., *outliers*), and also are consistent, are taken as potential demarcation lines. Finally, the potential line that gets the highest score, is picked as *suggested* demarcation line and its corresponding cut is taken as the *best* cut; and the line with the second high score is reported as *candidate* line.

To detect potential demarcation lines in an OCT image and measure their depth, we make the following two assumptions:

1) The demarcation line follows the same curvature as the top boundary (which is normally the case in almost all images, especially for 4mm-wide cross-sections of the cornea);
2) The total intensity value and continuity of bright pixels of the demarcation line (except the Epithelium's bright surface) is the highest among all other regions.

As before, the first assumption implies the pixel vector of the demarcation line can be found by shifting $top(x)$ down, such that the appropriate position is found based on the second assumption. This approach is valid since the demarcation line comprises a dense curve of high-intensity pixels with few pixels in thickness (unlike haze, which is usually thicker and does not necessarily span the width of the cornea).

Details about the foregoing scheme are provided in the following sub-sections.

A. IMAGE PRE-PROCESSING FOR IDENTIFYING BRIGHT CURVES

To identify curves of high-intensity pixels, we first blur the image so as to eliminate any sharp intensity fluctuation. Next, using an OpenCV function, we apply the Sobel Derivative [24] in the y-direction (i.e., vertical derivative), which multiplies each pixel value by the kernel matrix **KM**, so that it gets assigned a value representing the vertical change across the pixel; the greater the change, the higher the value.

$$\mathbf{KM} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

B. COMPUTATION OF S(c)

The computation of score $S(c)$ for each curve at depth d is obtained by scanning every pixel $p_{i=1..n} \in c(x)$, where n is the width of the curve, from left to right and computing the total pixel intensity value of $c(x)$. The computation of $S(c)$ must account for continuity and brightness; however, continuity should have a higher weight such that if, for example, a line at depth d_1 is continuous throughout the whole curve yet at the same time has a low overall pixel intensity value, and a line at depth d_2 has high intensity value but only appears in a small section of the curve, then the line at depth d_1 would have a higher score since it is more probable to be a potential demarcation line.

Given a curve $c(x)$ at depth d of pixels values $p_1, p_2, \ldots, p_n$, such that $0 \leq p_i \leq p_{max} \ \forall i \in \{1, \ldots, n\}$, the computation of its score $S(c)$ is performed as follows:

1) Compute the mean position, μ, of the non-zero-pixel values. For example, for $n = 540$, if at depth d, $c(x)$

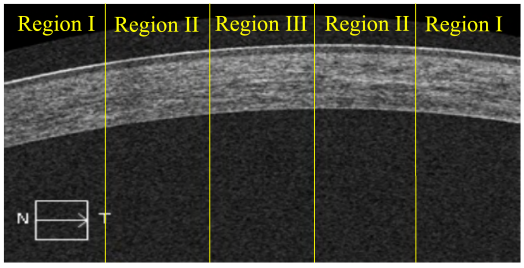


FIGURE 11. The different regions used in the score computation.

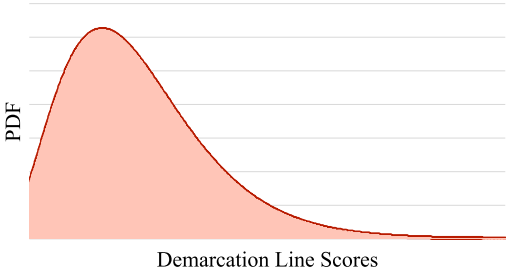


FIGURE 12. PDF of demarcation line scores.

has pixel values $\{p_3 = 18, p_{29} = 103, p_{413} = 107\}$ and $p_i = 0\ \forall i \in \{1, \ldots, n\} \setminus \{3, 29, 413\}$, then the mean position $\mu = 29$.

2) As illustrated in Fig. 11, we divide a cut into three regions and classify the mean value via a coefficient, $\mathfrak{c}_\mu$, computed as follows:

$$\mathfrak{c}_\mu = \begin{cases} 3, & \text{if } \mu \in \text{Region III} : \{\frac{2n}{5}, \frac{3n}{5}\} \\ 2, & \text{if } \mu \in \text{Region II} : \{[\frac{n}{5}, \frac{2n}{5}], [\frac{3n}{5}, \frac{4n}{5}]\} \\ 1, & \text{if } \mu \in \text{Region I} : \{[0, \frac{n}{5}], [\frac{4n}{5}, n]\} \end{cases} \tag{20}$$

By setting the parameter $\mathfrak{c}_\mu$, more weight is given to the curves with μ value in the center of the cornea, which most likely defines a wide-spread distribution of bright pixels across the curve, thereby exhibiting a continuous potential demarcation line.

3) Compute the standard deviation, σ, of the position of the non-zero pixel intensity values as follows:

$$\sigma^2 = \frac{\mathcal{S}}{n}, \quad \mathcal{S} = \sum_{i=1}^{n} (i - n)^2 \text{ if } p_i > 0. \tag{21}$$

That is, $\mathcal{S}$ is the summation of the squared difference between the mean, μ, and the different bright pixels; σ^2 is the average difference between the bright pixels and the mean μ.

4) Compute $S(c)$ as follows:

$$S(c) = \mathfrak{c}_\mu \times \sigma \times \sum_{i=1}^{n} p_i. \tag{22}$$

C. DETERMINING THE MINIMUM SCORE

In some cuts, a demarcation line must not be detected, as it is "barely visible". However, in the same cut, a small bright region in the stroma may have a "high-enough" score and can exist in several cuts (i.e., is consistent). In this case, the automated solution must not report it as a demarcation line; therefore, a minimum score must be set, below which a curve is not considered a potential demarcation line. To statistically set this score, denoted $\mathfrak{s}$, we first calculate the scores for all potential demarcation lines of all available OCT images. The total number of valid entries in our dataset is 5,228. To reduce the set space, we aggregate the scores such that, starting with the minimum score $\mathfrak{m}$ (of the available values), all scores belonging to $\mathfrak{m} + \mathfrak{o}$, are grouped (e.g., $\mathfrak{o} = 2000$). Subsequently, $\mathfrak{m}$ is constantly updated by $\mathfrak{o}$ (i.e., $\mathfrak{m} = \mathfrak{m} + \mathfrak{o}$) until the maximum score (which is obtained from the available values) is reached. The histogram of the grouped scores then helps determine the distribution of scores. As depicted in Fig. 12, the distribution belongs to the Gamma family, which requires computing its shape and scale parameters α and β. For any positive real number x, it is defined as follows:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt. \tag{23}$$

The cumulative distribution (CDF) function F is given by:

$$F(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^x t^{\alpha-1} e^{\frac{-t}{\beta}} dt. \tag{24}$$

The probability density function (PDF) is then obtained with $\alpha = 4.128$ and $\beta = 175718.39$. Finally, the 16^{th} percentile minimum risk is computed [34], by finding x such that $F(x; 4.128, 175718.39) = 0.16$. Hence, $\mathfrak{s} = x = 384489.6$. Subsequently, a line is considered a potential demarcation line if and only if its score $S(c) > \mathfrak{s}$; otherwise it is ignored.

D. IDENTIFYING POTENTIAL DEMARCATION LINE(s)

To identify the scores that are considered *outliers* (and thus taken as potential demarcation lines), we apply a statistical approach, where the set of scores is divided into four quartiles with three delimiters $Q1$, $Q2$ and $Q3$ [35]. The delimiter $Q1$ represents the lower 25^{th} percentile score $S(c)$; $Q2$ denotes the mid-50^{th} percentile score, and $Q3$ denotes the 75^{th} percentile score. Upper and lower *fences* (i.e., the cut-off scores that separate two quartiles) are usually set at a fixed distance from the interquartile range (i.e., $Q3 - Q1$). These are set by default as 1.5 times the interquartile range. Subsequently, any score that falls outside these fences is considered a potential outlier. Even when the scores are not normally distributed, this approach can be used safely as it depends on the median and not the average of the scores. In our scheme, only the upper fence is used as the threshold for considering a curve as a potential demarcation line. This threshold $\mathfrak{T}$ is computed as $\mathfrak{T} = Q3 + (Q3 - Q1) \times 1.5$. Hence, a curve $c(x)$ is considered a potential demarcation line if its score $S(c) > \mathfrak{T}$.

In a hazy image, the scores of potential demarcation line curves would be high, which makes them belong to $Q3$ subject to a higher upper fence value. This ensures that

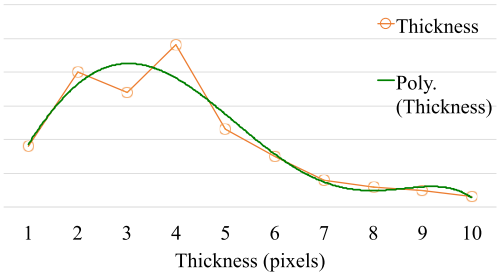


FIGURE 13. Histogram of demarcation line thickness.

these curves are not considered as part of a haze region. Thus, our mechanism is able to separate potential demarcation line from bright regions that should be taken as haze.

Finally, to ensure that our scheme does not consider multiple curves as multiple potential demarcation lines, when in fact they are the same line, we have evaluated the average thickness of a demarcation line in our data set. As shown in Fig. 13, the histogram of average thickness of all extracted potential lines for all images, determines the average thickness of a demarcation line to be approximately 4 pixels. Thus, while scanning the cut, whenever a curve is determined as *potential*, the scores of the next three curves are "tagged"; consequently, the one with the highest score among the four tagged curves, is taken as potential; the rest are discarded.

E. SETTING THE DETECTED DEMARCATION LINE(s)

Using the foregoing techniques on the pre-processed image, the following steps are applied to set and extract the suggested and candidate demarcation lines:

1) For each curve $c(x)$ from $top(x)$ to $bottom(x)$, calculate $S(c)$;
2) Identify and mark the potential demarcation lines, i.e., with $S(c) > \max(\mathfrak{s}, \mathfrak{T})$; these will be the *outliers* in the data set collected from each image;
3) Scan through the potential demarcation lines, and mark the position of the suggested demarcation line as y_d as the highest value of $S(c)$ (or second highest for reporting the candidate);
4) The suggested/candidate demarcation is thus set by: $dem(x) = top(x) + y_d$.

Note that in case a potential line is detected near or at the bottom boundary (i.e., in the last 35 microns of the cornea), it is considered the suggested line if and only if it is the only potential line detected; otherwise, it is considered as candidate even if it has the highest score (the second highest score is considered as the suggested). This is due to the fact that, as confirmed by manual operators, a high intensity (i.e., high $S(c)$ score) potential falling in the last 35 microns will most likely be a reflection in the cornea and not a demarcation line, except if it was the only potential line.

Fig. 14 illustrates an example of applying the foregoing method. Here, the suggested demarcation line $dem(x)$ highlighted in color yellow, and the candidate demarcation line is highlighted in color orange. The figure also shows how the depth is measured in microns.

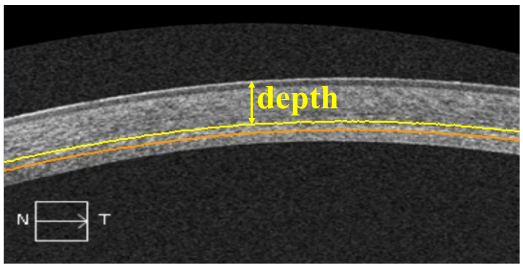


FIGURE 14. Suggested and the demarcation lines highlighted in yellow and orange, respectively.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we present and analyze the results collected using the automated method for OCT images of 40 patients' eyes whom underwent crosslinking for Keratoconus between May 2013 and Nov 2014, at the American University of Beirut Medical Center, Lebanon (28 males and 12 females; mean age: 22.0 $\pm$ 6.1 years). 10 eyes (i.e., 25%) had grade 1 Keratoconus according to the Amsler-Krumeich classification, whereas 24 eyes (i.e., 60%) had grade 2, and 6 eyes (i.e., 15%) had grade 3. Corneal OCT was performed as part of their routine postoperative protocol measurements. The images used for haze analysis were collected using our Cirrus HD-OCT at pre-operation and then at 2-weeks, 3-months, 6-months and 12-months post-operation; whereas the scans used for the demarcation line experiments were collected at 1-month and 3-months post-operation. The corneal OCT image capture is performed by asking the patient to fixate a light, which is not necessarily the geometric center of the cornea. This is applicable to any ophthalmic measuring device available on the market, as these devices center on the coaxially sighted subject fixated light reflex, and not on the geometric center of the cornea. In Keratoconus, more often than not, the visual axis is mildly deviated from the geometric center of the cornea, often towards the cone, and hence the measurement (square) looks decentered [36]. Clinically, the area of interest is around the visual axis, which has been common standard in the ophthalmic community and industry; thus, it would not affect the measurement repeatability or image quality. However, the capturing process is not automated as the image is not automatically dead-centered on target; so there may be room for human error in centration, which has the potential to affect repeatability. Nevertheless, the measurement encompasses a $4\,mm \times 4\,mm$ area of scanned cornea with the software making its pick among 128 cuts. This somehow corrects, at least partially, the non-automated centration in terms of repeatability. We note that repeatability and reproducibility may be enhanced when the analysis is applied to OCTs that perform automated corneal position detection such as AS-OCT.

The machine learning model, and the detection and measurement software (Java program) were implemented on a personal computer (Intel Xeon 32 cores, 32GB RAM,

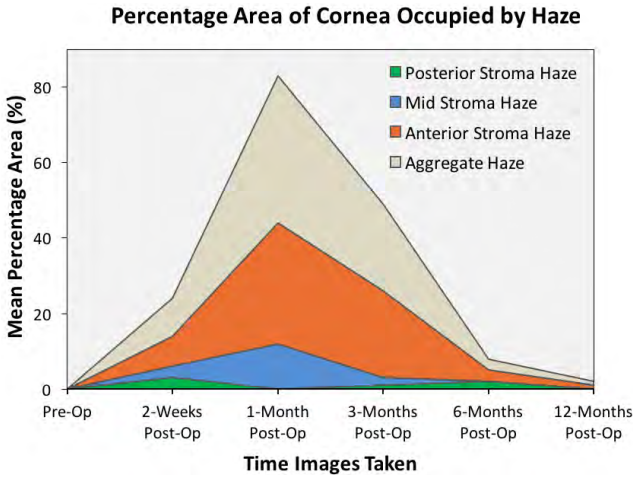


FIGURE 15. Change of corneal area occupied by haze from baseline for all patients over time.

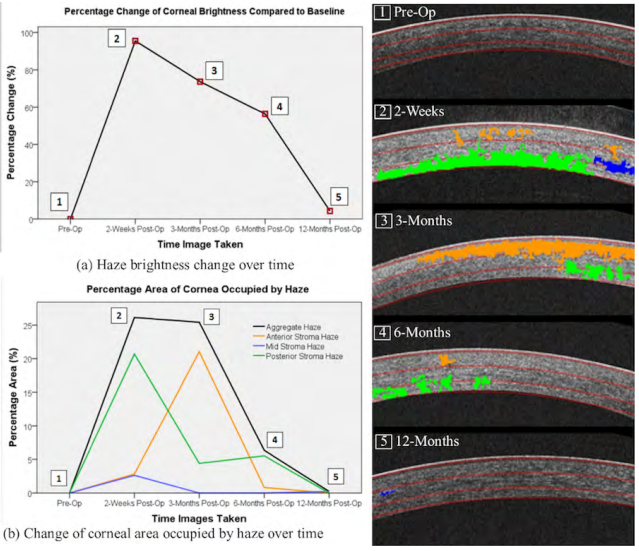


FIGURE 16. Analysis of sample patient haze over time.

32-bit OS). Corneal haze is evaluated by the area occupied by haze, and haze location (i.e., anterior, mid stroma, and posterior stroma); whereas the demarcation line is evaluated by its depth in microns as obtained by the automated solution vis-a-vis manual measurements. This study was approved by the Institutional Review Board at the American University of Beirut and adhered to the Declaration of Helsinki principles.

A. CORNEAL HAZE RESULTS

Fig. 15 shows the trends in haze location and density in the cornea as obtained by the automated technique. Overall, the haze is mostly present at 1-month post surgery for all regions; especially in the anterior and mid stroma regions. In contrast, haze in the posterior stroma region peaks at 2-weeks and 6-months post surgery.

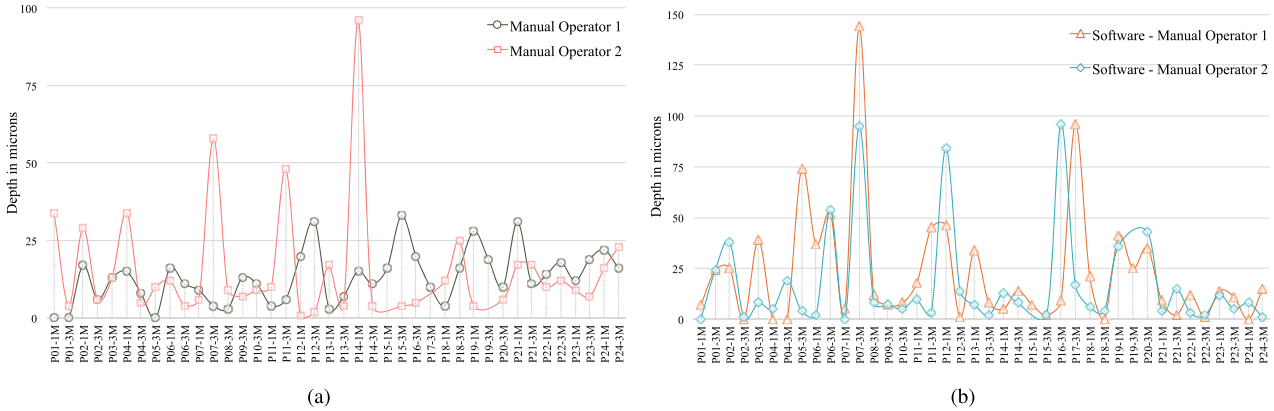
To highlight the ability of the automated method to aid in tracking each patient's haze progress over time, we show in Fig. 16 an example patient case. Here, the haze of the analyzed patient peaks at 2-weeks and 3-months post surgery compared to pre-op, with significant decrease to pre-op values after 6-months and 12-months of a typical steroid treatment course. In particular, posterior haze develops directly after treatment (from 2-weeks to 1-month). Subsequently, later in the patient's course (from 1-month to 3-months), the haze mainly occupies the anterior part of the cornea.

The results showcased in Fig. 15 and Fig. 16 are consistent with the ones obtained by Greenstein *et al.* [4] using Scheimpflug imaging and slit lamp biomicroscopy, where only corneal densitometry of the whole cornea was assessed. This bolsters further the credibility of the automated technique in assessing, categorizing, and staging corneal haze. More importantly, the accuracy of the OCT technology in assessing corneal haze is superior to Scheimpflug, as the latter, in addition to having a lower resolution, it suffers much more than OCT from distortion and erroneous data due to potential light scattering and absorption associated with anterior and mid stroma haze [37].

B. DEMARCATION LINE RESULTS

Two independent human operators and the automated detection software examined the corneal OCT scans. Operators evaluated the presence of the demarcation line and measured its depth by looking at OCT images on two separate occasions one week apart. The operators were blinded to patients' names, the results of other examiners (human or software), and their own previous results. Intra-observer measurements repeatability and intra-observer reproducibility are calculated and compared.

Table. 3 shows the depths of the demarcation lines in microns for all patients. The absence of a reported value indicates that the operator (software/human) did not detect any (candidate or suggested) demarcation line (e.g., P19 at 3-month post-operation). As noticed, some measurements show approximately equal values between the automated solution and the two manual operators (e.g., P24 at 1-month post-operation), or the margin of difference is very small. Here, we note that every pixel is $\approx$7.4 microns, and it was observed that a human operator commonly "drifts" by $\pm$5 pixels when doing manual measurements, which is equivalent to a drift of $\sim$40 microns. We also observe that in some cases the automated solution provides a measurement close to one of the human operators' measurements only (e.g., P13 at 1-month post-operation), or it provides a measurement that is far from both manual ones (e.g., P7 at 3-months post-operation). Interestingly, since the software is designed to detect and then select best potential line (i.e., the suggested one), as well as the second best potential (if applicable), those two can be fairly used to assess the performance of the software vs. the different measurements of human operators. More importantly, although a potential demarcation line might also be chosen by ophthalmologists as the most plausible one, for a line to be chosen as suggested and not candidate by the automated solution, this means that it has the highest score among all consistent potential demarcation lines.



(a) (b)

FIGURE 17. Differences of demarcation lines depths as measured by manual operators and software for all OCT images. (a) Among manual operators on two occasions. (b) Between software and average manual.

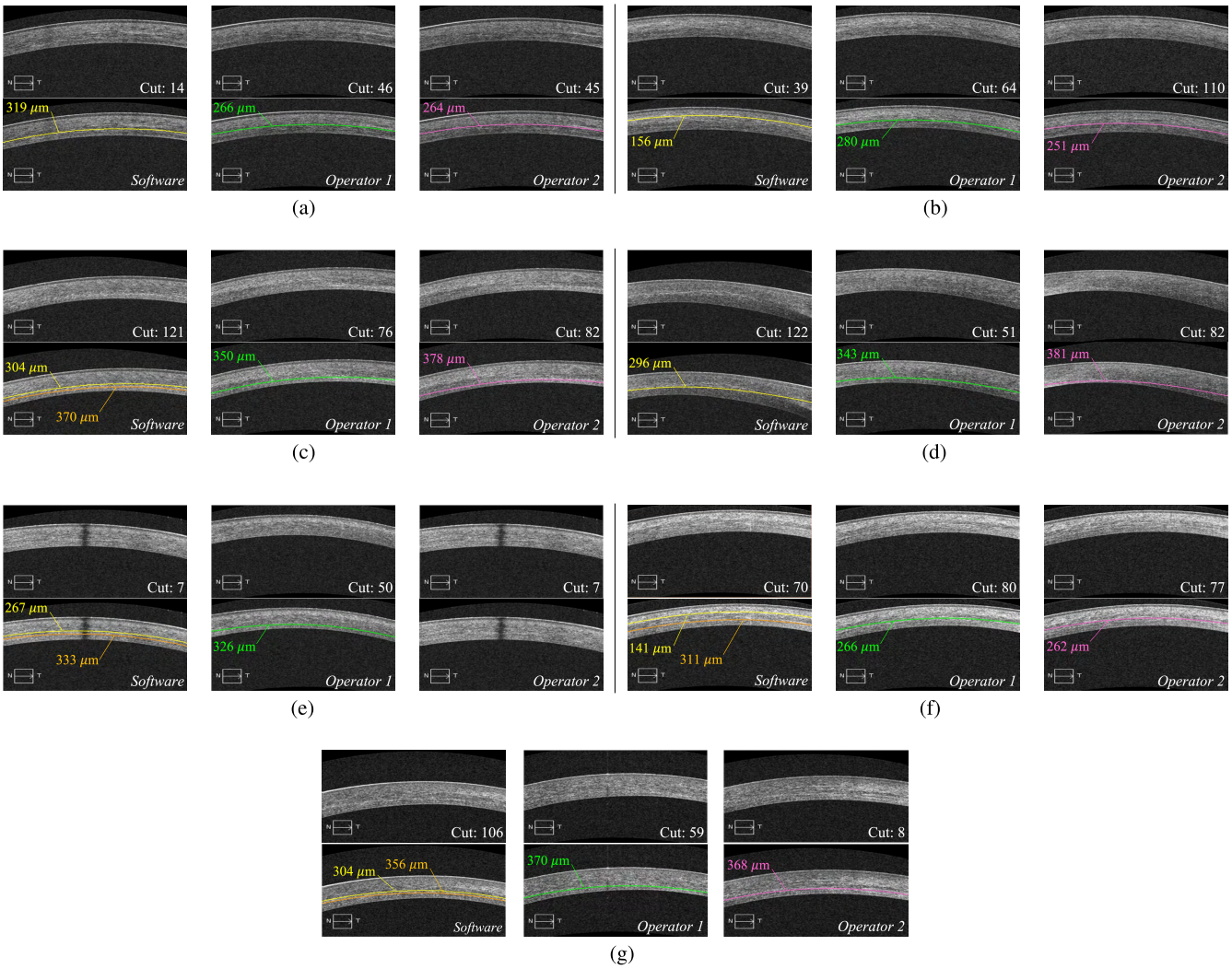


FIGURE 18. Raw and annotated cuts with suggested/candidate demarcation line depth(s) based on human and software operators for all outstanding cases. The annotation is made on the operator's chosen OCT cut. (a) P6: 3-months post-op. (b) P7: 3-months post-op. (c) P9: 3-months post-op. (d) P12: 1-month post-op. (d) P12: 1-month post-op. (e) P15: 1-month post-op. (f) P20: 3-months post-op. (g) P23: 1-month post-op.

The difference of measurements in microns between the different operators is further quantified in Fig. 17. As observed in Fig. 17(a), the measurements provided by the human operators differ between the two measurements of the same operator, and across operators. These results highlight the human error and inconsistency in obtaining manual

TABLE 3. Depth of *suggested* and *candidate* demarcation lines (in microns) as obtained by the software vs. manual measurements of two operators on two separate occasions.

	Software		Manual Op. 1		Manual Op. 2	
	Suggested	Candidate	(1)	(2)	(1)	(2)
P01-1M	259	244	252	252	260	294
P01-3M	237	–	262	262	262	266
P02-1M	244	–	287	270	283	312
P02-3M	274	–	274	280	273	279
P03-3M	341	415	393	380	407	394
P04-1M	267	304	304	289	299	333
P04-3M	267	–	274	266	286	291
P05-3M	356	–	281	281	341	351
P06-1M	222	207	276	260	232	220
P06-3M	319	356	266	255	264	260
P07-1M	274	–	289	280	281	275
P07-3M	156	141	304	300	309	251
P08-3M	311	–	296	299	303	294
P09-3M	304	370	363	350	378	385
P10-3M	319	356	311	300	333	324
P11-1M	378	363	341	345	388	398
P11-3M	356	–	304	310	352	400
P12-1M	296	–	363	343	381	382
P12-3M	289	–	259	290	303	305
P13-1M	378	–	415	412	387	370
P13-3M	319	348	333	340	350	346
P14-1M	304	370	365	350	287	383
P14-3M	326	–	311	300	338	334
P15-1M	267	333	326	342	–	–
P15-3M	296	378	266	299	290	294
P16-3M	319	–	348	328	420	415
P17-3M	385	437	289	299	420	420
P18-1M	326	341	304	300	332	320
P18-3M	319	274	274	290	290	315
P19-1M	148	–	218	190	190	186
P19-3M	326	400	281	300	–	–
P20-3M	141	311	266	276	262	268
P21-1M	222	281	259	290	277	260
P21-3M	259	274	287	276	307	290
P22-1M	363	385	336	350	398	388
P22-3M	341	237	339	321	326	338
P23-1M	304	356	370	382	377	368
P23-3M	289	–	319	300	301	294
P24-1M	356	370	378	356	378	362
P24-3M	385	341	326	310	341	364

subjective measurements, which motivates the need for automated objective measurements. Similarly, Fig. 17(b) shows the differences in measurement between the software and the average measurements provided by the two human operators. As observed, although blinded to the human operator's measurements, the software is able to provide very close (and sometimes almost identical) results taking into account the aforementioned human drift in measurement. This highlights the accuracy and robustness of the automated method. However, in Table 3, we also highlight the cases where discrepancy between the automated and manual operators is observed. This is due to the fact that unlike human operators, automated measurement takes into account multiple factors (i.e., valid cuts, consistency among cuts, continuity of brightness, etc.) for considering a demarcation line as a ''legal'' one, thereby having the human operators selecting ''imperfect'' cut and curve. These outstanding cases (i.e., P6-3M, P7-3M, P9-3M, P12-1M, P15-1M, P20-3M and P23-1M) are further analyzed and reported in Fig. 18. Our analysis concludes the following:

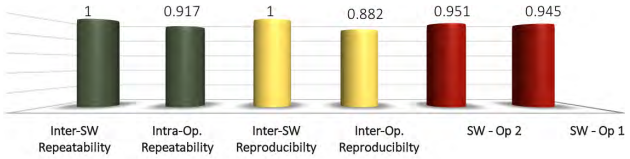


FIGURE 19. The ICC for all patients.

- Manual measurements do not systematically capture the clearest line, especially that ophthalmologists tend to mainly observe central cuts, as they assume that a clear line will most likely appear in middle cuts (cases of P7-3M, P12-1M); nevertheless, some manual measurements matched with the software's candidate line's depth as opposed to the suggested one (case of P9-3M);
- The thickness of the cornea sometimes varies among different cuts; however, the relative depths of the line between the cuts selected by the software and humans were approximately the same (i.e., most likely the same line was chosen by all operators; case of P6-3M);
- The selection of the cut has high impact on the measurement of the demarcation line depth, as in some cases the line may fluctuate in position among different cuts (e.g., P23-1M); or seen as ''non-existent'' by some operators (case of P15-1M).

These observations highlight the advantage of the automated tool in consistently selecting the best cut based on a well-defined criterion (which incorporates the score of the line and the consistency of the line among multiple cuts) as opposed to the error-prone human approximation.

Although a minimum score value was set, we notice that the software is still able to detect a suggested (and a candidate) demarcation line in all OCT images, even though one of the human operators did not detect a demarcation line in exactly two images: P15-1M and P19-3M. We examined the scores of the demarcation lines in the two images, and noticed that the demarcation line chosen by the software and one (or two) human operator(s) is ''not clear'' and with relatively the same pixel intensity. Hence, the software is able to give a well-computed score that accurately assesses the brightness and continuity of a stromal line; whereas, a human operator might give two different examinations for two lines with the same characteristics.

Finally, to assess the overall performance of the software vs. manual measurements, we compute the Pearson Correlation Coefficients (PCC) and the Intraclass Correlation Coefficients (ICC) using the two-way random-effects model and consistency of average measures for inter-operator relationship (i.e., between manual and automated operators), and the two-way mixed-effects model with absolute agreement of single measures for intra-operator relationship (i.e., for the different measures of the same manual operator). For an ICC value less than 0.4, groups are considered to have a poor agreement; whereas for an ICC value greater than 0.75, the data is said to have excellent agreement. In Fig. 19, we report the ICC values computed for the 40 Kerato-

conus eyes considered in our study. As noticed, the human-software (for both human operators) ICC value is very high (i.e., > 0.94, with $p < 0.001$). Furthermore, our experiments show that the mean corneal demarcation line depth is 295.9 $\pm$59.8 microns as computed by the automated solution; whereas it is 314.5 $\pm$48.4 microns as computed by the human operators. The PCC values between the software and the first and second manual operators are 0.918 and 0.910 ($p < 0.001$), respectively. The average time per OCT examination by the human operator is 31 seconds; whereas it is less than a second by the software. All these results validate the effectiveness and efficiency of the automated tool.

VI. CONCLUSIONS AND FUTURE WORK

We proposed a new automated solution that employs customized image processing and machine learning methods to detect and measure corneal haze and demarcation line in OCT images for Keratoconus patients after crosslinking. Our experimental results highlighted the effectiveness and efficiency of the automated solution, which enables ophthalmologists to obtain fast, accurate and objective information on corneal haze. The proposed solution has the promise to be employed as standardized method of care for stromal haze measurement of individual patients or in aggregate data for the purpose of longitudinal studies, and may improve clinical decision making after corneal surgeries such as cross-linking.

As future work, we plan to extend the automated tool to support most recent OCT machines, which also generate 8mm/9mm scans of the cornea. More notably, automated analysis will have an even more important role in newer, upcoming OCT technologies, like the polarization-sensitive OCT (ps-OCT). Furthermore, since the demarcation line depth exhibits a small variation across time, we plan to amend the algorithm's selection criteria such that it takes into consideration n sets of suggested/candidate demarcation lines at n different periods, so as to pick the most probable combination of lines. Finally, we plan to model the OCT images in 3D, which will provide a holistic view of corneal haze and demarcation line, and thus may enable optimal evaluation of their presence and depth.

ACKNOWLEDGMENTS

The authors would like to thank Ahmad Sibai for his help in the scoring technique of the demarcation line depth.

REFERENCES

[1] D. Huang *et al.*, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.
[2] Y. S. Rabinowitz, "Keratoconus," *Survey Ophthalmology*, vol. 42, no. 4, pp. 297–319, 1998.
[3] N. Sorkina and D. Varssano, "Corneal collagen crosslinking: A systematic review," *Ophthalmologica*, vol. 232, no. 1, pp. 10–27, Jun. 2014.
[4] S. A. Greenstein, K. L. Fry, J. Bhatt, and P. S. Hersh, "Natural history of corneal haze after collagen crosslinking for keratoconus and corneal ectasia: Scheimpflug and biomicroscopic analysis," *J. Cataract Refractive Surgery*, vol. 36, no. 12, pp. 2105–2114, 2010.
[5] M. Mita, G. O. Waring, and M. Tomita, "High-irradiance accelerated collagen crosslinking for the treatment of keratoconus: Six-month results," *J. Cataract Refractive Surgery*, vol. 40, no. 6, pp. 1032–1040, 2014.
[6] M. Doors, N. G. Tahzib, F. A. Eggink, T. J. M. Berendschot, C. A. B. Webers, and R. M. M. A. Nuijts, "Use of anterior segment optical coherence tomography to study corneal changes after collagen cross-linking," *Amer. J. Ophthalmol.*, vol. 148, no. 6, pp. 844–851, 2009.
[7] J. C. Yam *et al.*, "Corneal collagen cross-linking demarcation line depth assessed by visante OCT after CXL for keratoconus and corneal ectasia," *J. Refractive Surgery*, vol. 28, no. 7, pp. 475–481, 2012.
[8] G. D. Kymionis *et al.*, "Evaluation of the corneal collagen cross-linking demarcation line profile using anterior segment optical coherence tomography," *Cornea*, vol. 32, no. 7, pp. 907–910, 2013.
[9] G. D. Kymionis *et al.*, "Correlation of the corneal collagen cross-linking demarcation line using confocal microscopy and anterior segment optical coherence tomography in keratoconic patients," *Amer. J. Ophthalmol.*, vol. 157, no. 1, pp. 110–115, 2014.
[10] T. Seiler and F. Hafezi, "Corneal cross-linking-induced stromal demarcation line," *Cornea*, vol. 25, no. 9, pp. 1057–1059, 2006.
[11] E. B. Ozgurhan, B. I. S. Akcay, Y. Yildirim, G. Karatas, T. Kurt, and A. Demirok, "Evaluation of corneal stromal demarcation line after two different protocols of accelerated corneal collagen cross-linking procedures using anterior segment optical coherence tomography and confocal microscopy," *J. Ophthalmol.*, vol. 2014, Nov. 2014, Art. no. 981893.
[12] A.-H. Karimi, A. Wong, and K. Bizheva, "Automated detection and cell density assessment of keratocytes in the human corneal stroma from ultra-high resolution optical coherence tomograms," *OSA Biomed. Opt. Exp.*, vol. 2, no. 10, pp. 2905–2916, Oct. 2011.
[13] Z. Amini and H. Rabbani, "Statistical modeling of retinal optical coherence tomography," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1544–1554, Jun. 2016.
[14] J. Cheng *et al.*, "Speckle reduction in 3D optical coherence tomography of retina by a—Scan reconstruction," *IEEE Trans. Med. Imag.*, vol. 35, no. 10, pp. 2270–2279, Oct. 2016.
[15] C. Dongye *et al.*, "Automated detection of dilated capillaries on optical coherence tomography angiography," *OSA Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 1101–1109, Feb. 2017.
[16] L. de Sisternes, G. Jonna, M. A. Greven, Q. Chen, T. Leng, and D. L. Rubin, "Individual drusen segmentation and repeatability and reproducibility of their automated quantification in optical coherence tomography images," *ARVO Transl. Vis. Sci. Technol.*, vol. 6, no. 1, pp. 1–20, Feb. 2017.
[17] J. Eichel, A. Mishra, D. Clausi, P. Fieguth, and K. Bizheva, "A novel algorithm for extraction of the layers of the cornea," in *Proc. Can. Conf. Comput. Robot Vis.*, Kelowna, BC, Canada, May 2009, pp. 313–320.
[18] F. LaRocca, S. J. Chiu, R. P. McNabb, A. N. Kuo, J. A. Izatt, and S. Farsiu, "Robust automatic segmentation of corneal layer boundaries in SD-OCT images using graph theory and dynamic programming," *OSA Biomed. Opt. Exp.*, vol. 2, no. 6, pp. 1524–1538, Jun. 2011.
[19] D. Williams, Y. Zheng, F. Bao, and A. Elsheikh, "Automatic segmentation of anterior segment optical coherence tomography images," *SPIE J. Biomed. Opt.*, vol. 18, no. 5, pp. 056003-1–056003-7, May 2013.
[20] L. Fang, S. Li, D. Cunefare, and S. Farsiu, "Segmentation based sparse reconstruction of optical coherence tomography images," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 407–421, Feb. 2017.
[21] H. Fu *et al.*, "Segmentation and quantification for angle-closure glaucoma assessment in anterior segment OCT," *IEEE Trans. Med. Imag.*, vol. 36, no. 9, pp. 1930–1938, Sep. 2017.
[22] T. Zhang *et al.*, "A novel technique for robust and fast segmentation of corneal layer interfaces based on spectral-domain optical coherence tomography imaging," *IEEE Access*, vol. 5, pp. 10352–10363, Jun. 2017.
[23] A. R. Dhaini, M. Chokr, S. M. E. Oud, M. A. Fattah, and S. Awwad, "System and method for automated detection of corneal haze and demarcation line in optical coherence tomography images," U.S. Patent 6 161 021, Jul. 2017.
[24] D. L. Baggio, *OpenCV 3.0 Computer Vision With Java*. Birmingham, U.K.: Packt, 2015.
[25] G. H. Dunteman, *Principal Components Analysis*. Newcastle upon Tyne, U.K.: Sage, 1989, p. 69.
[26] R. R. Picard and R. D. Cook, "Cross-validation of regression models," *J. Amer. Statist. Assoc.*, vol. 79, no. 387, pp. 575–583, 1984.
[27] J. T. Kent, "Information gain and a general measure of correlation," *Biometrika*, vol. 70, no. 1, pp. 163–173, 1983.
[28] D. W. Hosmer, Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398. Hoboken, NJ, USA: Wiley, 2013.

[29] B. Scholkopf and A. J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001.
[30] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural Network Design*. Stillwater, OK, USA: Oklahoma State Univ., 2014.
[31] E. Frank *et al.*, *Weka*. Boston, MA, USA: Springer, 2005.
[32] I. T. Young, J. J. Gerbrands, and L. J. van Vliet, *Fundamentals of Image Processing*, 2nd ed. Delft, The Netherlands: Delft Univ. Technol., 2007.
[33] S. Suzuki *et al.*, "Topological structural analysis of digitized binary images by border following," *Comput. Vis., Graph., Image Process.*, vol. 30, no. 1, pp. 32–46, 1985.
[34] T. E. Rhodes, O. Chumburidze, and S. Qiao, *Percentile Methodology for Probability Distributions as Applied to the Representative Scenario Method*. Braintree, MA, USA: MIB Solutions, 2015.
[35] S. Walfish, "A review of statistical outlier methods," *Pharmaceutical Technol.*, vol. 30, no. 11, p. 82, 2006.
[36] K. Mihaltz, K. Kranitz, I. Kovacs, A. Takacs, J. Nemeth, and Z. Nagy, "Shifting of the line of sight in keratoconus meadured by a hartmann-shack sensor," *Ophthalmology*, vol. 117, no. 1, pp. 41–48, 2010.
[37] H. Huebscher, W. Fink, D. Steinbruck, and T. Seiler, "Scheimpflug records without distortion–a mythos?" *Ophthalmic Res.*, vol. 31, no. 2, pp. 134–139, 1999.

AHMAD R. DHAINI (M'04) received the B.Sc. degree in computer science from the American University of Beirut (AUB) in 2004, the M.Sc. degree in electrical and computer engineering from Concordia University, Canada, with a best thesis award nomination in 2006, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2011. From 2006 to 2007, he was a Software Analyst and a Consultant with TEKsystems, Canada, and from 2007 to 2008, he was a Software Designer with Ericsson, Canada. From 2011 to 2012, he was a Research Associate with the University of Waterloo and a Consultant with KAUST, Saudi Arabia. He is currently an Assistant Professor of computer science with AUB. He has received several awards such as the Ontario Graduate Scholarship in Science and Technology and other various teaching and research awards from the University of Waterloo.

From 2012 to 2014, he was a Post-Doctoral Scholar with the Photonics and Networking Research Laboratory, Stanford University. He received the prestigious NSERC Postdoctoral Fellowship. He also completed the Stanford Ignite Program for entrepreneurship and innovation, which teaches scientists how to convert an idea into a business.

Dr. Dhaini has authored or co-authored over 40 highly cited research articles in top IEEE journals and conferences. He holds two U.S. patents. He serves as an Editor for *Photonics Networks Communications* (Springer) and a reviewer and a Technical Program Committee Member for several major IEEE journals and conferences. He is a reviewer for NSF, NSERC, and several U.S. universities' internal grants. His research interests cover several themes of optical networks such as fiber-wireless broadband access networks, mission-critical networks, green communications, and software-defined networking. He is the PI and Co-PI of several projects, more specifically in the areas of mobile health and medical image analysis.

MANAL CHOKR received the bachelor's degree in computer science from Lebanese University in 2015, and the master's degree in computer science from the American University of Beirut (AUB) in 2017. Her current research interests include medical image analysis, computer vision, optimization, and machine learning. She was a recipient of the Best Graduate Student Award. She received the Outstanding Poster Award in the field of clinical research during the 7th Annual AUB Biomedical Research Day, Lebanon, for her research entitled "Automated Detection and Depth Measurement of Demarcation Line Using Optical Coherence Tomography in Keratoconus Patients After Crosslinking."

SARA MARIA EL-OUD received the bachelor's degree in physics from the American University of Beirut in 2015. She then pursued a year of research with Dr. A. Dhaini, during which they tackled the problem of objective analysis of medical images by developing a program for the automated detection and analysis of corneal haze. She joined The George Washington University in 2016, where she is pursuing the master's degree in physics. Her current research interests include computational biophysics, modeling of biological systems, and medical technology.

MAAMOUN ABDUL FATTAH received the medical degree from the University of Damascus. He is currently a Second Year Resident with the Department of Ophthalmology, American University of Beirut Medical Center. He has authored over eight full length articles in well-recognized scientific journals. These articles were specialized in ophthalmology. He received multiple awards for his research and participation in scientific meetings (Denver, Colorado ARVO 2015; ASCRS, San Deigo, CA, USA, 2015; and Beirut, AUBMC, 2015 and 2017). In addition, within his residency training, he has several ongoing research projects.

SHADY AWWAD received the medical degree from the American University of Beirut Medical Center (AUBMC). He is currently an Associate Professor, Director of the Laser Eye Center, and the Head of the Cornea and Refractive Surgery Division with AUBMC. He has authored or co-authored 57 peer-reviewed publications. His research focuses on refractive surgery, keratoconus, intraocular lens calculations, and corneal measurements. He was a recipient of the American Academy of Ophthalmology Achievement Award. He received the Alpha Omega Alpha Award, an internship in general surgery, and a residency in ophthalmology from 1999 to 2002 at AUBMC. He received a double fellowship in corneal diseases and surgery and refractive surgery from 2002 to 2006 and a fellowship in ophthalmic pathology from 2005 to 2006 at The University of Texas Southwestern Medical Center, Dallas, TX, USA.