Statistical modeling of retinal optical coherence tomography using the Weibull mixture model

Sahar Jorjandi,[1] Zahra Amini,[2] Gerlind Plonka,[3] and Hossein Rabbani[2,*]

[1] *Student Research Committee, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan 81746-734641, Iran*

[2] *Medical Image and Signal Processing Research Center, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*

[3] *Institute for Numerical and Applied Mathematics, Georg-August-University of Göttingen, Germany*

*Rabbani.h@ieee.org

Abstract: In this paper, a novel statistical model is proposed for retinal optical coherence tomography (OCT) images. According to the layered structure of the retina, a mixture of six Weibull distributions is proposed to describe the main statistical features of OCT images. We apply Weibull distribution to establish a more comprehensive model but with fewer parameters that has better goodness of fit (GoF) than previous models. Our new model also takes care of features such as asymmetry and heavy-tailed nature of the intensity distribution of retinal OCT data. In order to test the effectiveness of this new model, we apply it to improve the low quality of the OCT images. For this purpose, the spatially constrained Gaussian mixture model (SCGMM) is implemented. Since SCGMM is designed for data with Gaussian distribution, we convert our Weibull mixture model to a Gaussian mixture model using histogram matching before applying SCGMM. The denoising results illustrate the remarkable performance in terms of the contrast to noise ratio (CNR) and texture preservation (TP) compared to other peer methods. In another test to evaluate the efficiency of our proposed model, the parameters and GoF criteria are considered as a feature vector for support vector machine (SVM) to classify the healthy retinal OCT images from pigment epithelial detachment (PED) disease. The confusion matrix demonstrates the impact of the proposed model in our preliminary study on the OCT classification problem.

1. Introduction

Nowadays enormous datasets in different fields cause difficulty in their processing or interpretation. Hence, many data science experts are looking for a suitable model and pattern to take the advantages to specify the characteristics of each data category as particularly and uniquely [1,2]. So, modeling can be considered as the main core of many signal/image processing tasks because choosing different models lead to different strategies for data analysis. Particularly in medical image analysis which prior information about the properties of the organs under imaging is available, modeling shows its importance more. According to nature and noise effect, stochastic models are generally suggested in these images [3–6].

Optical Coherence Tomography (OCT) has appeared as the most promising ocular imaging technique that provides high-resolution, cross-sectional, and microstructure visualizing of the retina tissue [7,8]. It is applied to detect and manage more retinal diseases in the early stages [9,10]. Like other coherent systems, the OCT images have a multiplicative noise called speckle that dramatically degrades their quality as grainy appearance [8,11]. Here, a novel tailored statistical model for retinal OCT images has been proposed and its performance is surveyed in noise reduction application.

Until now several statistical distributions have been suggested for retinal structure modeling in OCT images and used in various processing applications.

The first investigation can be assigned to Grzywacz et al. that introduced the stretched exponential distribution for the homogenous rectangular regions or parallel regions on retinal layers. The modulation of parameters was served to detect Diabetic Retinopathy (DR). The dependence of parameters was defined as a position function and the spikes in the place of blots were observed in the DR data. However, the parameters were steady within the retinal layers of healthy data [12].

The Generalized Extreme Value (GEV) and Gaussian distributions were suggested to model seven layers of the retina. These models were exploited to build a synthetic OCT dataset [13].

Amini et al. defined a symmetric Normal Laplace Mixture (NLM) model, the convolution of Normal and Laplace distribution, to enhance the contrast of OCT images. It was handled matching the NLM to the Gaussian Mixture Model (GMM) using Gaussianization Transform (GT) [14]. In another study of the researchers, the GT was combined with Spatially Constrained Gaussian Mixture Model (SCGMM) algorithm to reduce the noise of retinal OCT images [15]. In [16], the suggested statistical model in [14] was modified to the asymmetric version due to the histogram of retinal layers' intensities and was called, Asymmetric Normal Laplace Mixture (ANLM) distribution. The denoising results indicated the increase of Contrast to Noise Ratio (CNR) compared to the symmetric model [17].

Jesus et al. described a statistical model for corneal OCT speckle in Region of Interests (ROIs) to evaluate the properties of microstructure in vivo. The generalized Gamma distribution was defined as the distribution of corneal OCT intensities and its parameters have statistically demonstrated significant differences for short and age-related corneal changes [18].

Dubose et al. [19] took into account the segmentation problem of OCT retina layers utilizing the Cramer-Rao Lower Bounds (CRLBs). The authors defined the intensity distribution of each retinal layer as K-family distribution to derive the CRLBs.

Recently, Samieinasab et al. [20] noted a horizontal dependency between neighbor pixels of retinal OCT images at specific distances modeled by a multivariate asymmetric Bessel K Form (BKF). Finally, the result of denoising has been reported by the combination of gaussianization and SCGMM. On the other hand, statistical modeling can also be applied in the transform domain. In the following, we would review them for OCT modeling.

Cauchy distribution has been suggested to model wavelet coefficients of the retinal OCT images. The distribution has been converted to Gaussian by applying histogram matching, and the noiseless coefficients have been recovered by Minimum Mean Square Error(MMSE) estimator [21]. Also, two recent literatures were reported on statistical modeling of retinal OCT images in transform domain [22,23]. Ghaderi et al. revealed the scale mixture model for the group of similar nonlocal patches of OCT image in the sparse domain. The sparse coefficients have been considered as the product of a random vector and a positive scale variable, which are modeled by Laplace and GEV distributions, respectively. The efficiency has been surveyed in denoising and obtained the high resolution OCT images captured by Bioptigen cooperation [22]. In another paper, Amini et al. extended their previous works to a multivariate mixture NL in the dual-tree-complex wavelet transform (DTCWT) domain. The researchers' innovation compared to their previous studies was to propose a multivariate model and gaussianization in the sparse domain. All above mentioned reviewed studies are summarized in Table 1 which is comprised the name of the statistical model, type of OCT device, the domain of the modeling, and application of modeling in the mentioned papers.

Furthermore, some other studies modeled the speckle noise of OCT images by Rayleigh [24], negative exponential [25], and gamma distribution [26]. With this in mind, we undertook this study to propose a more comprehensive model but with fewer parameters that has a better Goodness of Fit (GoF) than previous models as well as considering features such as asymmetry and heavy-tailed of the intensity distribution of retinal OCT data. Finally, to evaluate the performance of the proposed model, it is applied to reduce the noise from the retinal OCT

Table 1. Summary of proposed statistical Distributions for retinal OCT images Modeling

Proposed model	OCT device	Process Domain	Application	Ref.
Stretched Exponential	RTVue-100 FD- OCT	Spatial	Detection of DR	Grzywacz,2010
Normal +GEV	—	Spatial	synthetic dataset	Kulkarni,2011
Mixture of NL	Topcon 3-D OCT-1000/Cirrus HD-OCT	Spatial	Contrast enhancement	Amini,2016
		Spatial	Denoising	Amini,2017
		Sparse	Denoising	Amini,2020
K-Family	Bioptigen	Spatial	CRLB calculation for the OCT layer segmentation problem	Dubose,2017
Mixture of ANL	Topcon 3-D OCT-1000	Spatial	Denoising	Jorjandi,2019
Asymmetric BKF	Topcon 3-D OCT-1000/ Heidelberg/Bioptigen	Spatial	Denoising	Samieinasab,2020
Scale mixture model	Bioptigen	Sparse	Super-resolution/Denoising	Ghaderi,2020

images. In addition, a feasibility study to illustrate the effect of this model in OCT classification is provided.

The paper is organized as follows. Section II dedicates to the proposed model and its application on our dataset denoising. The result of GoF and noise reduction are reported in Section III. Finally, the discussion and conclusion are presented in Section IV and V, respectively.

2. Method

2.1. Dataset

In this study, two datasets are used that introduced in [27,28] and available in [29]. The first dataset was obtained from Topcon 3D OCT-1000 imaging system at the Ophthalmology Dept. of Feiz Hospital, Isfahan, Iran. Thirteen 3-D macular (spectral domain) SD-OCT images collected from eyes without pathology were included in this dataset. The x, y, z scale of the collected volumes is $650 \times 512 \times 128$ voxels, $7\,mm \times 3.125\,mm \times 3.125\,mm$, voxel size $13.67\,\mu m \times 4.81\,\mu m \times 24.41\,\mu m$. We have randomly selected 10 B-scans from each volume so we have totally130 B-Scans which the GoF of candidate distributions are compared on them. Also, the results of denoising are reported on 60 randomly selected B-scans from different volumes. It should be noted that 50 healthy data from this dataset is also selected for classification application. The second dataset contains six 3D OCT data that has recorded from the same device and with the same characteristics. They were diagnosed to have retinal Pigment Epithelial Detachment (PED). This dataset is used to survey the ability of the proposed model to classify healthy data from abnormal ones. Hence, 50 B-scans with pathological symptoms are selected from this dataset.

2.2. Statistical modeling

To provide a statistical model for OCT images, at first a logarithm operator is used to convert the speckle noise to an additive Gaussian noise [30, 31].

2.2.1. Retinal layer modeling

To improve the modeling of a retinal OCT image, we have to find a suitable distribution for each layer that provides an appropriate insight into the special structure. To this end, the eleven retinal

layers in OCT images were manually segmented by an expert and their normalized histograms were plotted as illustrated in Fig. 1.

The most important observed features were the asymmetry and heavy-tailed nature of data distribution. Among the existing distributions, our goal is to present a model with less complexity and parameters that can be well fitted to OCT data.

In this regard, based on the data structure, different distributions were investigated that two-parameter Weibull distribution can better address the mentioned properties than other candidate distributions. The probability distribution function (pdf) of a Weibull random variable is as
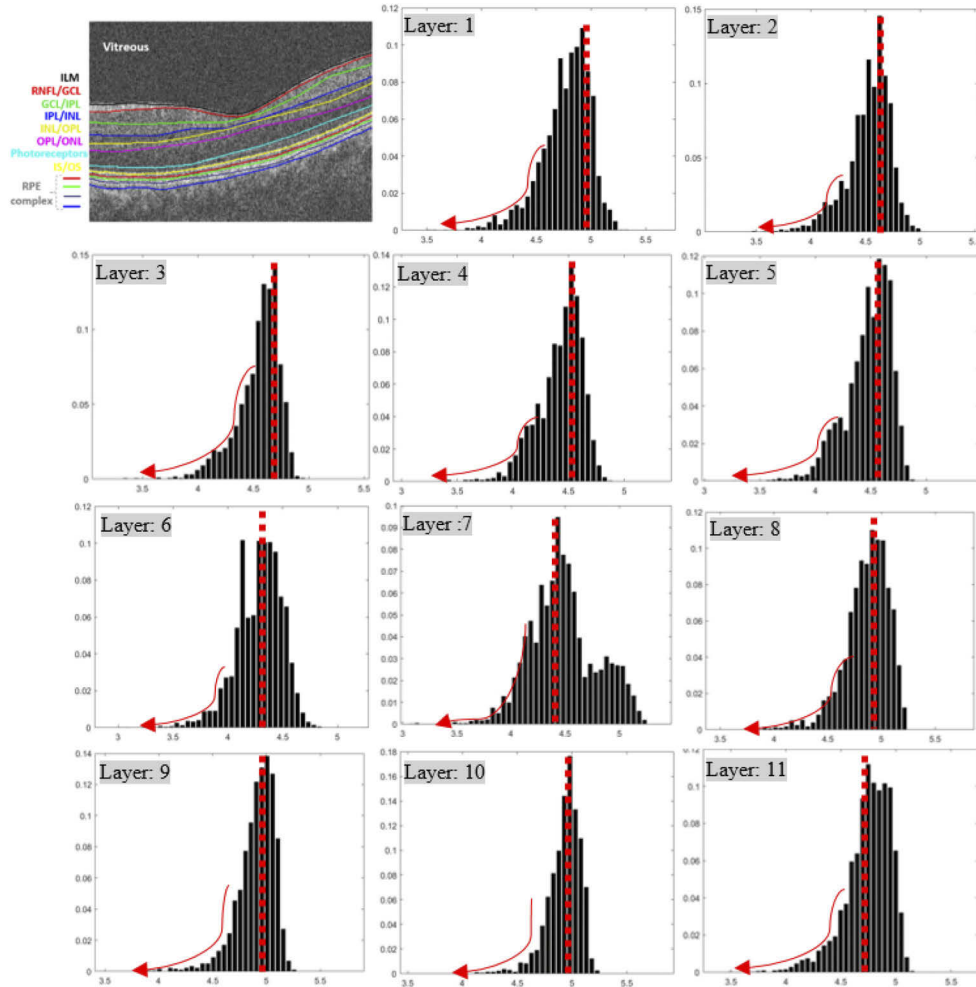
Fig. 1. A sample B-scan of retinal OCT with displaying the manual segmentation of the twelve boundaries as well as their nomination and the normalized histogram of each layer. The 11 layers are numbered from ILM to the last RPE complex, respectively. Heavy-tail and asymmetry nature of data distribution are marked as curved red line and vertical dashed red line in each graph, respectively. The x-axis and y-axis are the range of intensity values of pixels and the normalized histogram of these intensities, respectively.

follow [32,33].

$$f_{weibull}(x; \alpha, \beta) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} & e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, x \geq 0 \\ 0, x < 0 \end{cases} \tag{1}$$

where α, β are parameters of scale and shape, respectively and x denotes the pixel intensity in each B-scan. The advantages are the comprehensiveness, i.e., changing of the values of Weibull parameters leads to a range of well-known distributions as special cases such as exponential ($\beta = 1$) and Rayleigh ($\beta = 2, \alpha = \sqrt{2}\sigma$). The proposed pdf has the ability to model the features of data (asymmetry and heavy-tailed properties). This is compared with Gamma, symmetric and asymmetric Normal Laplace, Generalized Gaussian Distribution (GGD), and GEV pdfs that have been used in the previous studies. The visual and numerical results of GoF demonstrated the ability of Weibull distribution in capturing the main statistical properties of each retinal layer of OCT images in comparison to other mentioned distributions. In the following, we introduce a statistical retinal OCT image model based on Weibull distribution.

2.2.2. Retinal OCT image modeling

Since retina has several various cellular layers, we suggest a mixture distribution [34–37] to model the retinal OCT images. The number of mixture components can in the simplest case, be equivalent to the total eleven anatomic layers of retina. However, these layers are located in different geometric positions, some of them have similar statistical characteristics and can be described with an identical pdf. Due to the one by one (alternate) changing of retinal layer's intensities, we have selected 6 components as nearly optimal value for the number of components. Since the Weibull pdf has a significant GoF for each layer, it was considered as a tailored distribution of each component in the mixture model. Therefore, the mixture of Weibull distribution with different values of parameters in each component is chosen in our proposed model of retinal OCT images.

The proposed mixture model is formulated as

$$g(x) = \sum_{i=1}^{I} w_i f(x; \alpha_i, \beta_i) \tag{2}$$

where $f(x; \alpha_i, \beta_i) = f_{weibull}(x; \alpha_i, \beta_i)$ as in (1), w_i are non-negative weights with $\sum_{i=1}^{I} w_i = 1$. $\alpha_i > 0$, $\beta_i > 0$, $i = 1, \ldots, I$ are the parameters of the Weibull pdf. In our experiments we have used $I = 6$ in model Eq. (2). In total, the parameter sets $\alpha = (\alpha_1, \ldots, \alpha_I)$, $\beta = (\beta_1, \ldots, \beta_I)$ and $w = (w_1, \ldots, w_I)$ have to be estimated from the observed data $x_j, j = 1, \ldots, N$, where in our case N is the number of pixels in an image. The accurate estimation of these parameter sets plays an important role for the GoF results of the mixture model and for the performance of the processing tasks.

Due to incomplete data, we employ the Expectation Maximization (EM) [38,39] algorithm to estimate the model parameters. The EM works iteratively. We start with an initial estimate $\alpha^{(0)}$, $\beta^{(0)}$, $w^{(0)}$ according to our observations from statistical behavior of OCT layers. At the (k+1)-th iteration step of the EM algorithm, we proceed as follows. See Appendix I for more details.

First, we compute the auxiliary variables

$$r_l^{(k)}(x_j) := \frac{w_l^{(k)} f\left(x_j; \alpha_l^{(k)}, \beta_l^{(k)}\right)}{\sum_{i=1}^{I} w_i^{(k)} f\left(x_j; \alpha_i^{(k)}, \beta_i^{(k)}\right)} \tag{3}$$

for $j = 1, \ldots, N$ and $l = 1, \ldots, I$, (expectation step). In the second step (Maximization step), we update the parameter sets $\alpha^{(k)}, \beta^{(k)}, w^{(k)}$ as follows,

$$w_l^{(k+1)} = \frac{\sum_{j=1}^{N} r_l^{(k)}(x_j)}{N} \tag{4}$$

$$\alpha_l^{(k+1)} = \left(\frac{\sum_{j=1}^N x_j^{\beta_l^{(k)}} \cdot r_l^{(k)}(x_j)}{\sum_{j=1}^N r_l^{(k)}(x_j)} \right)^{1/\beta_l^{(k)}} \tag{5}$$

$$\beta_l^{(k+1)} = - \frac{\sum_{j=1}^N r_l^{(k)}(x_j)}{\sum_{j=1}^N r_l^{(k)}(x_j) \cdot \ln\left(\frac{x_j}{\alpha_l^{(k+1)}}\right)\left[1 - \left(\frac{x_j}{\alpha_l^{(k+1)}}\right)^{\beta_l^{(k)}}\right]} \tag{6}$$

for $l = 1, \ldots, I$. The iteration is repeated until we achieve $|\alpha^{(k+1)} - \alpha^{(k)}| \le \varepsilon$, $|\beta^{(k+1)} - \beta^{(k)}| \le \varepsilon$ and $|w^{(k+1)} - w^{(k)}| \le \varepsilon$. Instead of the given formulas in Eqs. (3)–(6), we can also re-compute the auxiliary variables using the latest estimates. For example, after having computed $w_i^{(k+1)}$ we would have

$$r_l(x_j; w^{(k+1)}, \alpha^{(k)}, \beta^{(k)}) = \frac{w_l^{(k+1)} f(x_j; \alpha_l^{(k)}, \beta_l^{(k)})}{\sum_{i=1}^I w_i^{(k+1)} f(x_j; \alpha_i^{(k)}, \beta_i^{(k)})} \tag{7}$$

and $r_l^{(k)}(x_j)$ can be replaced by $r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)})$ in Eq. (5). Similarly, $r_l^{(k)}(x_j)$ can be replaced by the updated $r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)})$ in Eq. (6). Moreover, the estimate for $\beta^{(k+1)}$ can be improved by solving the nonlinear equations

$$\frac{1}{\beta_l} \sum_{j=1}^N r_l^{(k)}(x_j) + \sum_{j=1}^N r_l^{(k)}(x_j) \cdot ln\left(\frac{x_j}{\alpha_l^{(k)}}\right)\left[1 - \left(\frac{x_j}{\alpha_l^{(k)}}\right)^{\beta_l}\right] = 0 \tag{8}$$

directly, where again, $r_l^{(k)}(x_j)$ and, $\alpha_l^{(k)}$ can be replaced by the updates $r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)})$ and $\alpha_l^{(k+1)}$, respectively. In our numerical experiments, we have solved Eq. (8) iteratively by choosing appropriate initial values. Most benefit of our proposed mixture model compared to the one introduced in [14,17] is related to the form of Eqs. (5)-(6). Here, we computed these parameters in the explicit formulas using the numerical methods and fixed the non-convergence problem of parameters that the impacts are illustrated in the result section. The proposed mixture model is compared with NLMM, ANLMM and, GMM using GoF tests and visual assessment which shows the significant superiority of the proposed model. After finding the suitable model for OCT data, its performance is investigated in denoising application.

2.3. OCT image denoising

As mentioned before, speckle noise has greatly reduced the quality of OCT images, hence several articles have surveyed on the noise reduction of these images [40–44]. Because of the effect of denoising in the main steps of OCT image processing such as segmentation or abnormalities detection, we aim to provide an efficient method of noise reduction using the MAP estimator based on the proposed mixture model. For this purpose, we selected the SCGMM algorithm which is a well-known denoising algorithm based on Gaussian assumption. So, we first convert the Weibull mixture model (WMM) to the Gaussian mixture model and then apply the SCGMM algorithm.

2.3.1. Gaussianization

Converting non-Gaussian data to Gaussian type is considered in multiple processing applications [45–52]. Here, Gaussianization is performed using the histogram matching which transforms an image so that its histogram matches to a specified histogram. Here, the Cumulative Distribution Function (CDF) of the Weibull model is equivalent with the CDF of Gaussian distribution

according to the following equation,

$$F_W(x_W) = F_G(x_G)$$

$$1 - e^{-\left(x_W/\alpha_i\right)^{\beta_i}} = \frac{1}{2}[1 + \mathrm{erf}(\frac{x_G - \mu_{iG}}{\sigma_{iG}\sqrt{2}})] \tag{9}$$

The Gaussian random variable x_G is derived in terms of the Weibull random variable x_W, and Weibull and Gaussian parameters $\{\mu_{iG}, \sigma_{iG}\}$ as

$$x_G = \mu_{iG} + \sqrt{2}\sigma_{iG}(\mathit{erfinv}(1 - 2e^{-\left(x_W/\alpha_i\right)^{\beta_i}}) \tag{10}$$

where *erfinv* is the inverse of error function, $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$. The Gaussian parameters, $\{\mu_{iG}, \sigma_{iG}\}$, are calculated in terms of Weibull parameters as

$$\mu_{iG} = \alpha_i\Gamma\left(1 + 1/\beta_i\right) \tag{11}$$

$$\sigma_{iG} = \alpha_i\sqrt{\left[\Gamma\left(1 + 2/\beta_i\right) - \left(\Gamma\left(1 + 1/\beta_i\right)\right)^2\right]} \tag{12}$$

where Γ is gamma function [53].

In this way, each Weibull component in mixture model is converted to a Gaussian component. Finally, all Gaussian components are combined using weighted summation proposed in [54] with weights $\frac{w_i f_i}{\sum_{i=1}^m w_i f_i}$. So, the final formula for producing the enhanced OCT image with Gaussian mixture model is

$$\hat{x} = \sum_{i=1}^6 w_i x_{iG} f_i(x) \Big/ \sum_{i=1}^6 w_i f_i(x) \tag{13}$$

where x_{iG} is the ith gaussianized component and $f_i(.)$ denotes the Weibull distribution using the parameters of ith component. Finally, the exponential operator is used to return the image into the main domain due to use of the logarithm operator at the beginning of the study.

2.3.2. SCGMM

The SCGMM algorithm [55] is a patch-based method [56–59] in which the image is divided into a number of overlapping patches. Then the number of exemplar patches are uniformly chosen in the image, and similar patches are selected using K-Nearest Neighbors algorithm for each exemplar and form a cluster. It is assumed that each cluster has a Gaussian distribution with mean vector and specific covariance matrix. Therefore, the image would follow a Gaussian mixture model. To calculate these areas as well as to estimate the Gaussian parameters (mean vector and correlation matrix), a two-step algorithm is used, the first step of which is the clustering step and the second step is noise reduction (Fig. 2 shows the main steps of SCGMM).

Finally, the steps of denoising based on Weibull mixture model are summarized in Algorithm I.

Algorithm I:

| *Proposing denoising algorithm based on WMM* |
| --- |
| ***Statistical Modeling*** |
| Fitting Weibull mixture model on OCT data and parameter estimation by EM algorithm (Eqs. (2-6)) |
| ***Denoising*** |
| *First step:* Converting Weibull mixture distribution to Gaussian Mixture Model using the histogram matching method (Eqs. (10-13)) |
| *Second Step:* Utilizing SCGMM method [54] for OCT image denoising (Fig. 2) |
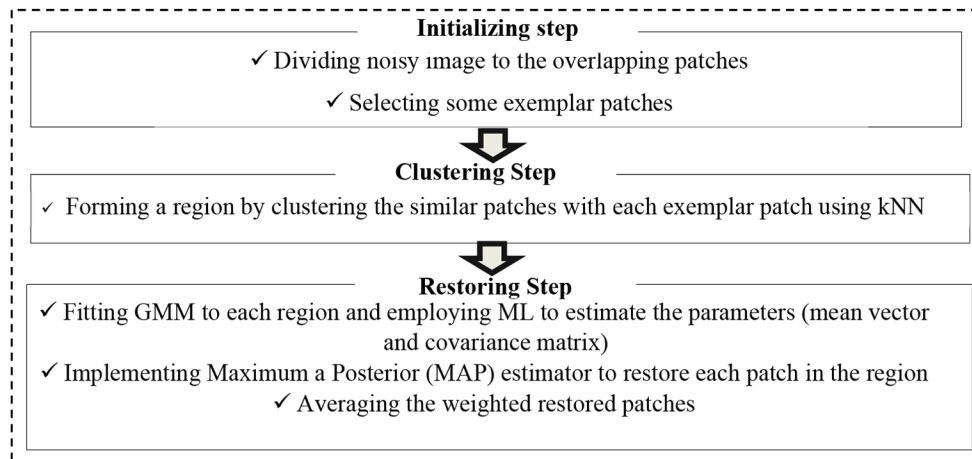
Fig. 2. The steps of SCGMM algorithm for image denoising

3. Results

In the two next subsections, the results of the study are reported. Subsection *A* is related to the GoF of the proposed model as visual and numerical measurements. In Subsection *B* the application of the proposed method in denoising of retinal OCT images is investigated.

3.1. Goodness of fit

Two numerical measurements were calculated to show how well the various above-mentioned statistical models (i.e. GEV, NL, ANL, Gamma, GGD and Weibull) fit into the intensity distribution of retina layers and B-scans. Chi square examines the difference between the observation and expected values according to the following equation:

$$\chi^2 = \sum_{i=1}^{B} \frac{(q_i - p_i)^2}{q_i} \tag{14}$$

where p_i and q_i point out to data distribution and the statistical model, respectively. B is the number of bins for the histogram. Another numerical measurement is Kullback-Leibler Divergence (KLD) which is called relative entropy and measures the discrepancy between two probability distributions. So, it can serve as a GoF test.

$$D_{KL}(P\|Q) = \sum_{i=1}^{B} p_i ln\frac{p_i}{q_i} \tag{15}$$

The smaller value of the criteria indicates the less discrepancy. In this study, they were used to compare six distributions to model the retina layers as well as four mixture models applied to model the whole OCT B-scan. Table 2–4 and Fig. 3–4 are related to the quantitative and qualitative evaluations of the results, respectively. From these results we can conclude that WMM has the lowest value of the quantitative criteria and it is demonstrated more remarkable superiority than other models in the modeling of retinal OCT images.

3.2. Denoising of OCT images

The results of the proposed method on denoising application were evaluated by three numerical criteria called CNR, Texture Preservation (TP), and Edge Preservation (EP). They were calculated in the chosen regions of interest (ROI) on the B-scans.
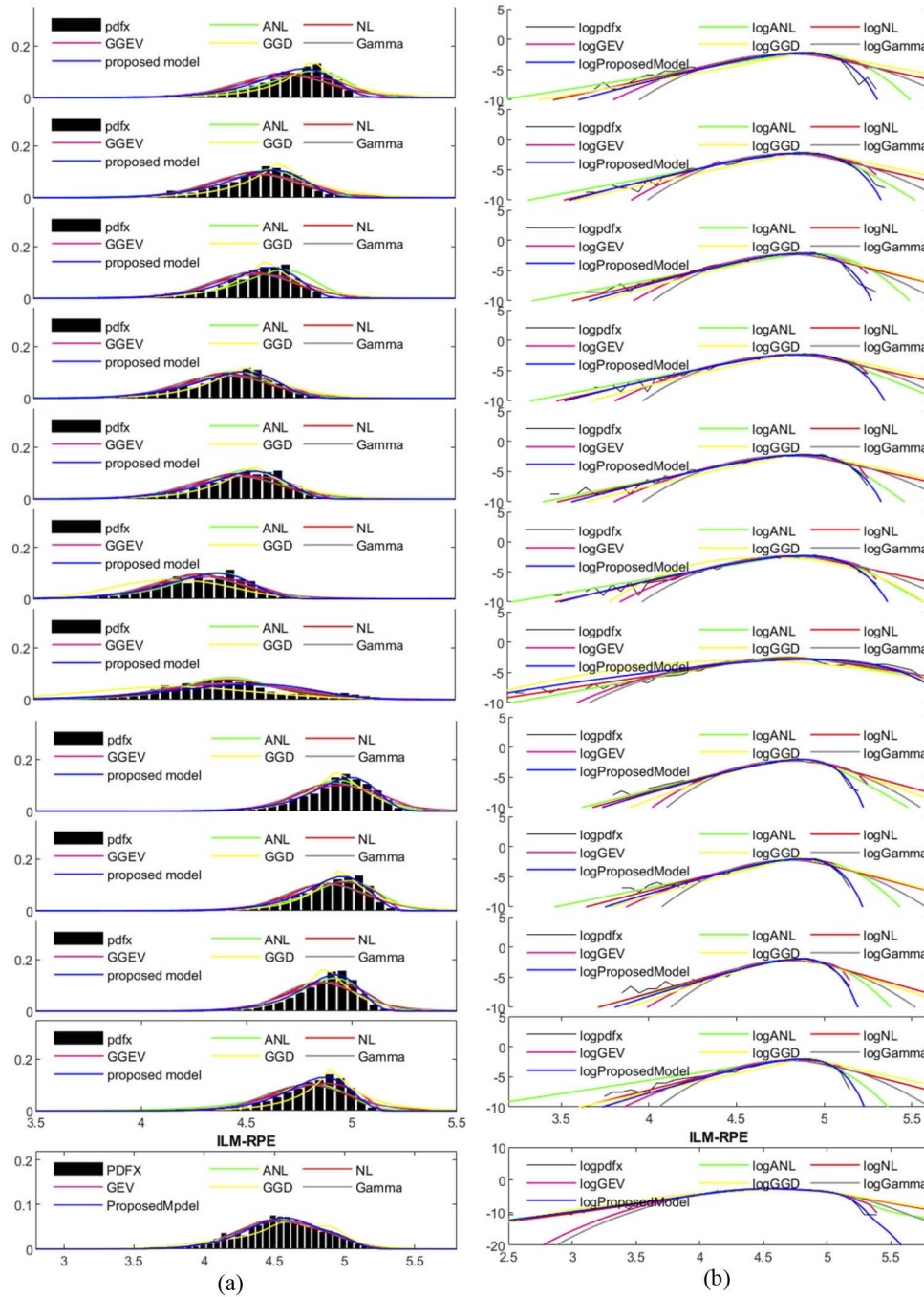
Fig. 3. (a) comparison of the goodness of fit between proposed model, symmetric NL (NL), GEV, Gamma, GGD and asymmetric NL (ANL) in eleven retinal layers of OCT. The x-axis and y-axis are the range of intensity values of pixels and the normalized histogram of these intensities, respectively. (b) logarithm of intensity distribution to survey the goodness of fit in more details and tails between proposed model, symmetric NL (NL), GEV, Gamma, GGD and asymmetric NL (ANL) in eleven retina layers of OCT. The x-axis is same to the (a) and y-axis denotes the logarithm value of intensity distribution of (a).
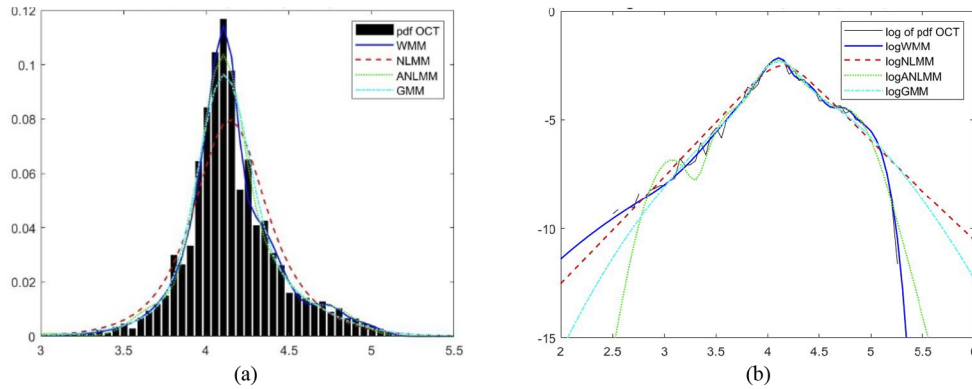
Fig. 4. (a) Comparison the goodness of fit between the proposed mixture model (WMM), symmetric NL mixture model (NLMM), Gaussian Mixture Model (GMM), and Asymmetric NL Mixture Model (ANLMM) for a sample B-scan. The x-axis and y-axis are the range of intensity values of B-scan's pixels and the normalized histogram of B-scan intensities, respectively. (b) Logarithmic intensity distribution of the sample B-scan to survey the goodness of fit in more details (specially in tails). The x-axis is the range of intensity values of B-scan's pixels and y-axis denotes the logarithm value of B-scan's intensity distribution of (a).

Table 2. The mean of Chi-Square values in 130 B-scans to compare goodness of fit for 6 pdfs which are fitted to each retina layer

Layers	Gamma	NL	GEV	GGD	ANL	Weibull
1st	0.0101	0.0090	0.0065	0.0045	0.0034	**0.0019**
2nd	0.0082	0.0071	0.0064	0.0035	0.0031	**0.0016**
3rd	0.0099	0.0091	0.0062	0.0044	0.0037	**0.0016**
4th	0.0070	0.0068	0.0043	0.0044	0.0033	**0.0015**
5th	0.0073	0.0070	0.0035	0.0041	0.0031	**0.0014**
6th	0.0042	0.0048	0.0033	0.0104	0.0048	**0.0023**
7th	**0.0025**	0.0026	0.0032	0.0068	0.0037	0.0063
8th	0.0064	0.0057	0.0029	0.0043	0.0034	**0.0010**
9th	0.0102	0.0096	0.0044	0.0054	0.0052	**0.0014**
10th	0.0118	0.0121	0.0061	0.0059	0.0078	**0.0014**
11th	0.0079	0.0074	0.0027	0.0048	0.0034	**0.0011**

Table 3. The mean of KLD values in 130 B-Scans to compare goodness of fit for 6 pdfs which are fitted to each retina layer

Layers	Gamma	NL	GEV	GGD	ANL	Weibull
1st	0.1882	0.1711	0.0930	0.2973	0.0509	**0.0284**
2nd	0.1074	0.1054	0.0663	0.1534	0.0331	**0.0202**
3rd	0.1680	0.1692	0.0839	0.2678	0.0673	**0.0215**
4th	0.1046	0.1141	0.0497	0.2060	0.0428	**0.0197**
5th	0.1360	0.1417	0.0457	0.2518	0.0512	**0.0198**
6th	0.0821	0.0993	0.0503	0.1965	0.0798	**0.0257**
7th	0.0516	0.0694	**0.0462**	0.2375	0.0783	0.0952
8th	0.1357	0.1314	0.0450	0.1704	0.0556	**0.0169**
9th	0.1514	0.1430	0.0479	0.1866	0.0642	**0.0233**
10th	0.1562	0.1841	0.0596	0.1846	0.0842	**0.0225**
11th	0.1882	0.1711	0.0930	0.2973	0.0509	**0.0284**

CNR measures the contrast between the pixels of m-th ROI and background (noisy region) as given by following definition.

$$CNR = 10 \log \frac{\mu_{mROI} - \mu_{background}}{\sqrt{\sigma_{mROI}^2 + \sigma_{background}^2}} \tag{16}$$

where μ_{mROI}, σ_{mROI}^2 denote the mean and variance of m-th ROI chosen in the intra-layers and the $\mu_{background}$, $\sigma_{background}^2$ are related to the noisy region on the image (vitreous). It is obvious if the m-th region has a high mean and less standard deviation, it will make a high CNR. So, the region with high texture (large σ^2) calculated mistakenly the contrast low. Hence the next criterion (TP) is computed to consider this property in the denoising algorithms and show the preservation of texture after denoising. It is between 0 and 1, and is close to 0 if the texture is become so flatten. The texture regions, layers area, are favor to calculate the TP based on equation in [60].

The last indicator, EP, was calculated to display the maintaining of edges after applying the noise reduction method. EP ranges are similar to TP and the related equation in [60] calculates the correlation of the edge in the denoised and noisy image over the ROIs. The numerical denoising results were reported in Table 5. Also, the quality of the denoising method was shown on a sample image in Fig. 5. From these results we can conclude that our proposed denoising method (WMM-GT-SCGMM) provided the images with high CNR and TP compared to the other methods, i.e. SCGMM, NLMM-GT-SCGMM and, ANLMM-GT-SCGMM.

Table 4. The mean value of Chi-Square values and KLD metrics in 130 B-Scans to compare goodness of fit for mixture models

Criterion	NLMM	GMM	ANLMM	WMM
Chi-Square	0.0050	0.0025	0.0018	**0.0015**
KLD	0.0365	0.0204	0.0224	**0.0115**

Also, the computational time of the aforementioned denoising methods has been calculated for a sample 271×512 OCT image using a 2.40 GHz Intel Core i7 CPU and reported in Table 6. Our method has a less computational time than ANLMM-GT-SCGMM method however it needs more computational time than SCGMM and NLGTSCGMM. This is due to the more accurate parameter estimation method which leads to better results on Goodness-of-Fit. Furthermore,
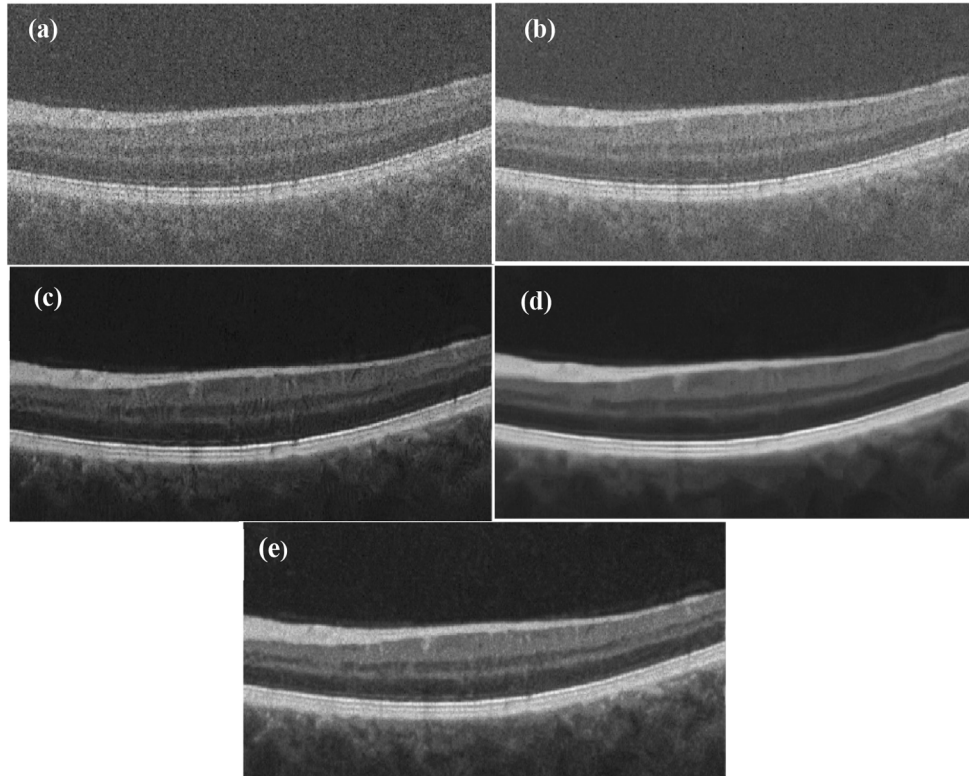
Fig. 5. Comparison of denosing results in a retinal Bscan.(a) original image (b) SCGMM (c) NLMM-GT-SCGMM (d)ANLMM-GT-SCGMM (e) WMM-GT-SCGMM.

Table 5. The mean of the denoising criteria on 60 B-Scans

Method	CNR	TP	EP
Original image	8.52	1.00	1.00
SCGMM	11.45	0.79	0.98
NLMM-GT-SCGMM	11.54	0.61	0.92
ANLMM-GT-SCGMM	14.71	0.75	0.87
WMM-GT-SCGMM	**14.87**	**0.81**	0.91

our denoising results are also better than other competitive methods. Therefore, this additional processing time is justified, especially for offline applications like denoising.

Table 6. The comparison of computational time of the denoising methods

SCGMM	NLMM-GT-SCGMM	ANLMM-GT-SCGMM	WMM-GT-SCGMM
402 sec.	564 sec.	880 sec.	713 sec.

4. Discussion

In this study, a new statistical mixture model has been addressed for retinal OCT images in the spatial domain. For this purpose, we first explored a suitable model for each retinal layer using 2-parameter Weibull distribution. It is more compatible regarding the intensity distribution of each layer (excluded the 7th layer) than other candidate distributions like Gamma, GEV, NL, ANL, and GGD. Regarding the 7th layer, none of the previously examined distributions fitted well because it is a very thin layer and it is difficult to propose a mixture state for it. The proposed Weibull model with few parameters and computational complexity, is able to model the asymmetry and heavy-tailed nature of the intensity distribution of the layers as well as possible, therefore the mixture of them has been selected for modeling of retinal OCT images. This mixture model is composed of 6 components of Weibull distribution. In addition to aforementioned advantages of the proposed model, a robust estimation of WMM's parameters has been provided by EM algorithm due to the closed form of equations and convergence of parameters. Eventually, WMM was compared to three other mixture models, GMM, NLMM, and ANLMM. It had more remarkable superiority than the other candidate mixture models in terms of both qualitative and quantitative criteria. In addition, the efficiency of the proposed model was investigated in noise reduction. For this purpose, SCGMM algorithm was used to decrease the noise. The proposed algorithm, WMM-GT-SCGMM, has demonstrated better CNR and TP compared to other examined denoising methods. Therefore, it seems that the proposed method can be used as a suitable preprocessing step for future operations such as diagnosis from texture analysis. Although most previous modeling studies on OCT images have been purposed for noise reduction, some researchers have used modeling to diagnose disease. Therefore, in this study a preliminary study was conducted to classify normal and abnormal data (PED dataset) using the proposed model.

In order to differentiate between normal and abnormal datasets, the region of retina from the first to the last boundary (ILM-RPE) was considered which segmented automatically using method in [28]. Figure 6 illustrated an OCT image with the pathological symptoms as well as extracted ILM and RPE boundaries.
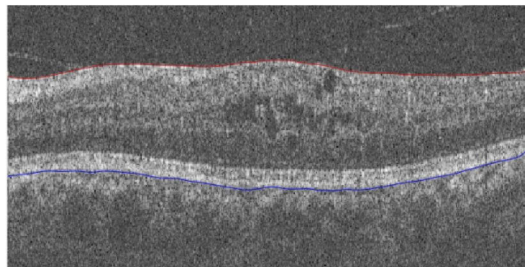
Fig. 6. Separating the ILM and RPE borders on the retinal OCT image with pathological symptoms in red and blue, respectively.

The proposed mixture model was fitted to the intensity distribution associated with the pixels of ILM-RPE region and the parameters and GoF criteria values $\{\alpha_1, \ldots, \alpha_6, \beta_1, \ldots, \beta_6, w_1, \ldots, w_6, \chi^2, D_{kl}\}$ were extracted as feature vectors of 50 healthy B-scans from 5 subjects and 50 B-scans with pathological symptoms from 5 subjects and then given to a linear SVM classifier as input vector. In order to reduce the effect of correlation which is related to selected B-scans from same subject in train and test phases, the classifier was repeatedly trained and tested on ten different selections of subjects' combination so that 80 percent of subjects were selected for train step and remained 20 percent for test step. Eventually, the average of results of accuracy, precision, recall and F1-score are reported in Table 7.

Table 7. the performance of linear SVM classifier for classification of healthy and PED disease in OCT images.

Accuracy	Precision	Recall	F1-score
88.5%	91.1%	88.5%	86.8%

The obtained results of classification procedure show that a simple classifier can result in an acceptable performance of distinguishing between healthy data from patients which illustrate the ability of the proposed Weibull mixture model in classification of OCT data.

5. Conclusion

Modeling is considered as an important and notable topic in medical imaging, so that the more specific and precise model achieves the ultimate purpose of image processing tasks, i.e., discrimination of healthy from abnormal data. Hence, in this study the Weibull mixture statistical model was addressed for retinal OCT images that despite of less complexity and parameters, was considerably more consistent with dataset than previous ones, and is due to the well GoF. Also, the problem of EM algorithm that were directly solved using a numerical method and GoF results illustrated the remarkable effect.

In addition to address the statistical model for retinal OCT images, its efficiency has also been investigated for noise reduction and classification. The denoising results illustrated the proficiency of proposed method (WMM-GT-SCGMM) to improve the CNR while preserving the texture. In the initial study to utilize the proposed model in classification, the parameters and criteria of GoF as a feature vector achieved the promising results in the machine learning problem. Therefore, the proposed model can be taken into account as a way to discriminate the healthy from unhealthy data. In future work, adding more specific features, such as the geometry of disease symptoms as well as more data volumes, would allow for using sophisticated classification methods such as deep learning networks and discrimination of more diseases.

Appendix I

Here we shortly explain the EM algorithm for iterative calculation of the parameters of the Weibull Mixture Model $g(x; w, \alpha, \beta) = \sum_{i=1}^{I} w_i f(x, \alpha_i, \beta_i)$ considered in Eq. (2) for I=6 with $f(x; \alpha_i, \beta_i) = f_{Weibull}(x; \alpha_i, \beta_i)$, and with the unknown parameters $\alpha = (\alpha_1, \ldots, \alpha_I)$, $\beta = (\beta_1, \ldots, \beta_I)$, and $w = (w_1, \ldots, w_I)$ and the side condition $\sum_{i=1}^{I} w_i = 1$. For given observed data x_j, $j = 1, \ldots, N$, (in our case, N is the number of pixels in an image) we aim at estimating all $w_i, \alpha_i, \beta_i, i = 1, \ldots, I$.

Assuming that the realizations x_j are independent and identically distributed, the maximum likelihood estimation gives us

$$(\hat{w}, \hat{\alpha}, \hat{\beta}) = arg \max_{w, \alpha, \beta} L_N(w, \alpha, \beta) \qquad \text{s.t.} \quad \sum_{i=1}^{I} w_i = 1 \qquad (17)$$

with the log-likelihood function $L_N(w, \alpha, \beta) = \log(\prod_{j=1}^{N} g(x_j; w, \alpha, \beta)) = \sum_{j=1}^{N} \log\left(\sum_{i=1}^{I} w_i f(x_j, \alpha_i, \beta_i)\right)$.

Since this maximization leads to a nonlinear optimization problem, we use the EM algorithm which is an iterative alternating maximization approach. We start with an initial set of parameters as $w^{(0)} = (0.13, 0.17, 0.1, 0.26, 0.04, 0.3)$, $\alpha^{(0)} = (3, 3.5, 4, 4.5, 5, 5.5)$ and $\beta^{(0)} = (21, 22, 23, 24, 40, 26)$, and iteratively compute the parameter sets $w^{(k)}$, $\alpha^{(k)}$ and $\beta^{(k)}$, where in each step, two parameter sets stay to be fixed while the remaining one is optimized.

In the (k+1)-th iteration step, we calculate $w^{(k+1)}$ by approximating

$$w^{(k+1)} = arg \max_w \sum_{j=1}^{N} \log \left(\sum_{i=1}^{I} w_i f(x_j; \alpha_i^{(k)}, \beta_i^{(k)}) \right), \qquad \text{s.t.} \sum_{i=1}^{I} w_i = 1. \qquad (18)$$

The constraint gives for the last weight $w_I = 1 - \sum_{i=1}^{I-1} w_i$. Taking the derivatives with regard to w_l, $l = 1, \ldots, I - 1$, we find the necessary conditions

$$\begin{aligned}
\frac{\partial}{\partial w_l} L_N(w, \alpha^{(k)}, \beta^{(k)}) &= \frac{\partial}{\partial w_l} \sum_{j=1}^{N} \log \left(\sum_{i=1}^{I} w_i f(x_j; \alpha_i^{(k)}, \beta_i^{(k)}) \right) \\
&= \sum_{j=1}^{N} \frac{1}{\sum_{i=1}^{I} w_i f(x_j; \alpha_i^{(k)}, \beta_i^{(k)})} \left(\frac{\partial}{\partial w_l} (\sum_{i=1}^{I} w_i f(x_j; \alpha_i^{(k)}, \beta_i^{(k)})) \right) \\
&= \sum_{j=1}^{N} \frac{1}{\sum_{i=1}^{I} w_i f(x_j; \alpha_i^{(k)}, \beta_i^{(k)})} (f(x_j; \alpha_l^{(k)}, \beta_l^{(k)}) - f(x_j; \alpha_I^{(k)}, \beta_I^{(k)})) = 0
\end{aligned} \qquad (19)$$

where we have used in the last line that $\frac{\partial w_I}{\partial w_l} = -1$. We set

$$r_l(x_j, w, \alpha, \beta) := \frac{w_l f(x_j; \alpha_l, \beta_l)}{\sum_{i=1}^{I} w_i f(x_j; \alpha_i, \beta_i)} \qquad (20)$$

for $l = 1, \ldots, I$, and $j = 1, \ldots, N$, and observe that $\sum_{l=1}^{I} r_l(x_j, w, \alpha, \beta) = 1$ for all $j = 1, \ldots, N$. Then we can conclude from Eq. (19) that for $l = 1, \ldots, I - 1$

$$\sum_{j=1}^{N} \frac{r_l(x_j, w, \alpha^{(k)}, \beta^{(k)})}{w_l} - \sum_{j=1}^{N} \frac{r_I(x_j, w, \alpha^{(k)}, \beta^{(k)})}{w_I} = 0, \qquad (21)$$

i.e., with $\lambda := \sum_{j=1}^{N} \frac{r_I(x_j, w, \alpha^{(k)}, \beta^{(k)})}{w_I}$,

$$\lambda w_l = \sum_{j=1}^{N} r_l(x_j, w, \alpha^{(k)}, \beta^{(k)}), l = 1, \ldots, I. \qquad (22)$$

Taking the sum over all l in the last equation, we obtain from the side condition for the weights

$$\lambda = \lambda \left(\sum_{l=1}^{I} w_l \right) = \sum_{l=1}^{I} \sum_{j=1}^{N} r_l(x_j, w, \alpha^{(k)}, \beta^{(k)}) = \sum_{j=1}^{N} \left(\sum_{l=1}^{I} r_l(x_j, w, \alpha^{(k)}, \beta^{(k)}) \right) = N. \qquad (23)$$

Using the previous iteration $w^{(k)}$ instead of w in this sum in Eq. (22), we get the new estimate

$$w_l^{(k+1)} = \frac{1}{N} \sum_{j=1}^{N} r_l(x_j, w^{(k)}, \alpha^{(k)}, \beta^{(k)}). \qquad (24)$$

Next, we calculate $\alpha_i^{(k+1)}$, $i = 1, \ldots, I$, by considering the derivative of the log-likelihood function, where $w^{(k+1)}$ and $\beta^{(k)}$ stay to be fixed. We find for the Weibull density function f

$$\begin{aligned}
\frac{\partial}{\partial \alpha} f(x; \alpha, \beta) &= \frac{\partial}{\partial \alpha} \left(\frac{\beta x^{\beta-1}}{\alpha^{\beta}} e^{-(\frac{x}{\alpha})^{\beta}} \right) \\
&= f(x; \alpha, \beta) \frac{\beta}{\alpha} \left(\left(\frac{x}{\alpha} \right)^{\beta} - 1 \right).
\end{aligned}$$

Thus, we obtain for $l = 1, \ldots, I$ the necessary conditions

$$
\begin{aligned}
\frac{\partial}{\partial \alpha_l} L_N(w^{(k+1)}, \alpha, \beta^{(k)}) &= \sum_{j=1}^N \frac{1}{\sum_{i=1}^I w_i^{(k+1)} f(x_j; \alpha_i, \beta_i^{(k)})} w_l^{(k+1)} \frac{\partial}{\partial \alpha_l} f(x_j; \alpha_l, \beta_l^{(k)}) = \\
\sum_{j=1}^N &\frac{w_l^{(k+1)} f(x_j; \alpha_l, \beta_l^{(k)})}{\sum_{i=1}^I w_i^{(k+1)} f(x_j; \alpha_i, \beta_i^{(k)})} \left(\frac{\beta_l^{(k)}}{\alpha} \left(\left(\frac{x_j}{\alpha} \right)^{\beta_l^{(k)}} - 1 \right) \right) = \\
\frac{\beta_l^{(k)}}{\alpha} &\sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha, \beta^{(k)}) \left(\left(\frac{x_j}{\alpha} \right)^{\beta_l^{(k)}} - 1 \right) = 0.
\end{aligned}
\tag{25}
$$

Again, we replace $r_l(x_j, w^{(k+1)}, \alpha, \beta^{(k)})$ by $r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)})$ in the last sum and find

$$
\sum_{j=1}^N x_j^{\beta_l^{(k)}} r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)}) = \alpha^{\beta_l^{(k)}} \sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)}).
\tag{26}
$$

This leads to the new estimates

$$
\alpha_l^{(k+1)} = \left(\frac{\sum_{j=1}^N x_j^{\beta_l^{(k)}} r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)})}{\sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)})} \right)^{1/\beta_l^{(k)}}
\tag{27}
$$

for $l = 1, \ldots, I$.

Finally, we calculate $\beta_i^{(k+1)}$ *for* $i = 1, \ldots, I$ by considering the derivative of the log-likelihood function, where the parameter sets $w^{(k+1)}$ and $\alpha^{(k+1)}$ stay to be fixed. From

$$
\begin{aligned}
\frac{\partial}{\partial \beta} f(x; \alpha, \beta) &= \frac{\partial}{\partial \beta} \left(\frac{\beta}{\alpha} \frac{x^{\beta-1}}{\alpha^{\beta-1}} e^{-\left(\frac{x}{\alpha} \right)^\beta} \right) \\
&= f(x; \alpha, \beta) \left(\frac{1}{\beta} + \ln \left(\frac{x}{\alpha} \right) \left[1 - \left(\frac{x}{\alpha} \right)^\beta \right] \right)
\end{aligned}
\tag{28}
$$

for the Weibull density function, we find

$$
\begin{aligned}
\frac{\partial}{\partial \beta_l} &= L_N(w^{(k+1)}, \alpha^{(k+1)}, \beta) = \\
\sum_{j=1}^N &\frac{1}{\sum_{i=1}^I w_i^{(k+1)} f(x_j; \alpha_i^{(k+1)}, \beta_i)} w_l^{(k+1)} \frac{\partial}{\partial \beta_l} f(x_j; \alpha_l^{(k+1)}, \beta_l) = \\
\sum_{j=1}^N &\frac{w_l^{(k+1)} f(x_j; \alpha_l^{(k+1)}, \beta_l)}{\sum_{i=1}^I w_i^{(k+1)} f(x_j; \alpha_i^{(k+1)}, \beta_i)} \left(\frac{1}{\beta_l} + \ln \left(\frac{x_j}{\alpha_l^{(k+1)}} \right) \left[1 - \left(\frac{x_j}{\alpha_l^{(k+1)}} \right)^{\beta_l} \right] \right) = \\
\sum_{j=1}^N &r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta) \left(\frac{1}{\beta_l} + \ln \left(\frac{x_j}{\alpha_l^{(k+1)}} \right) \left[1 - \left(\frac{x_j}{\alpha_l^{(k+1)}} \right)^{\beta_l} \right] \right) = 0.
\end{aligned}
\tag{29}
$$

Again, we simplify the calculation be replacing $r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta)$ by $r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)})$ and obtain

$$
\frac{1}{\beta_l} \sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)}) = - \sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)}) \ln \left(\frac{x_j}{\alpha_l^{(k+1)}} \right) \left[1 - \left(\frac{x_j}{\alpha_l^{(k+1)}} \right)^{\beta_l} \right]
\tag{30}
$$

Finally, replacing β_l by $\beta_l^{(k)}$ on the right-hand side of the last equation, we arrive at the new estimate

$$
\beta_l^{(k+1)} = - \frac{\sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)})}{\sum_{j=1}^N r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)}) \ln \left(\frac{x_j}{\alpha_l^{(k+1)}} \right) \left[1 - \left(\frac{x_j}{\alpha_l^{(k+1)}} \right)^{\beta_l^{(k)}} \right]}.
\tag{31}
$$

Observe that in the EM algorithm, one usually computes

$$r_l^{(k)}(x_j) := r_l(x_j, w^{(k)}, \alpha^{(k)}, \beta^{(k)}) := \frac{w_l^{(k)} f(x_j, \alpha_l^{(k)}, \beta_l^{(k)})}{\sum_{i=1}^{I} w_i^{(k)} f(x_j, \alpha_i^{(k)}, \beta_i^{(k)})} \tag{32}$$

in the so-called expectation step and just uses this approximation instead of the improved function values $r_l(x_j, w^{(k+1)}, \alpha^{(k)}, \beta^{(k)})$ in Eq. (27) or $r_l(x_j, w^{(k+1)}, \alpha^{(k+1)}, \beta^{(k)})$ in Eq. (31).

References

1. J. Gunawardena, "Models in systems biology: the parameter problem and the meanings of robustness," Elem. Comput. Syst. Biol. **1**, 21–47 (2010).
2. C. Grimm and G. E. Marai, "Biomedical applications: from data capture to modeling," IEEE Computer Graphics and Applications **32**, 20–21 (2012).
3. M. E. Leveton, "Statistical Models in Medical Image Analysis," *MIT* (2000).
4. T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," Med. Image Anal. **13**(4), 543–563 (2009).
5. T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Medical Imaging 2001: Image Processing* (International Society for Optics and Photonics, 2001), 4322, pp. 236–248.
6. J. Koikkalainen, T. Tölli, K. Lauerma, K. Antila, E. Mattila, M. Lilja, and J. Lötjönen, "Methods of artificial enlargement of the training set for statistical shape models," IEEE Trans. Med. Imaging **27**(11), 1643–1654 (2008).
7. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," Science **254**(5035), 1178–1181 (1991).
8. M. E. J. van Velthoven, D. J. Faber, F. D. Verbraak, T. G. van Leeuwen, and M. D. de Smet, "Recent developments in optical coherence tomography for imaging the retina," Prog. Retin. Eye Res. **26**(1), 57–77 (2007).
9. M. Bhende, S. Shetty, M. K. Parthasarathy, and S. Ramya, "Optical coherence tomography: A guide to interpretation of common macular diseases," Indian J. Ophthalmol. **66**(1), 20 (2018).
10. M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," IEEE Rev. Biomed. Eng. **3**, 169–208 (2010).
11. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," J. Biomed. Opt. **4**(1), 95–106 (1999).
12. N. M. Grzywacz, J. De Juan, C. Ferrone, D. Giannini, D. Huang, G. Koch, V. Russo, O. Tan, and C. Bruni, "Statistics of optical coherence tomography data from human retina," IEEE Trans. Med. Imaging **29**(6), 1224–1237 (2010).
13. P. Kulkarni, D. Lozano, G. Zouridakis, and M. Twa, "A statistical model of retinal optical coherence tomography image data," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (IEEE, 2011), pp. 6127–6130.
14. Z. Amini and H. Rabbani, "Statistical modeling of retinal optical coherence tomography," IEEE Trans. Med. Imaging **35**(6), 1544–1554 (2016).
15. Z. Amini and H. Rabbani, "Optical coherence tomography image denoising using Gaussianization transform," J. Biomed. Opt. **22**(08), 1 (2017).
16. S. Jorjandi, H. Rabbani, R. Kafieh, and Z. Amini, "Statistical modeling of Optical coherence tomography images by asymmetric normal Laplace mixture model," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (IEEE, 2017), pp. 4399–4402.
17. S. Jorjandi, H. Rabbani, Z. Amini, and R. Kafieh, "OCT image denoising based on asymmetric normal Laplace mixture model," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2019), pp. 2679–2682.
18. D. A. Jesus and D. R. Iskander, "Assessment of corneal properties based on statistical modeling of OCT speckle," Biomed. Opt. Express **8**(1), 162 (2017).
19. T. B. DuBose, D. Cunefare, E. Cole, P. Milanfar, J. A. Izatt, and S. Farsiu, "Statistical Models of Signal and Noise and Fundamental Limits of Segmentation Accuracy in Retinal Optical Coherence Tomography," IEEE Trans. Med. Imaging **37**(9), 1978–1988 (2018).
20. M. Samieinasab, Z. Amini, and H. Rabbani, "Multivariate statistical modeling of retinal optical coherence tomography," IEEE Trans. Med. Imaging **39**(11), 3475–3487 (2020).
21. S. Sahu, H. V. Singh, B. Kumar, and A. K. Singh, "Statistical modeling and Gaussianization procedure based de-speckling algorithm for retinal OCT images," J Ambient Intell Human Comput **1**, 1–14 (2018).

22. P. G. Daneshmand, H. Rabbani, and A. Mehridehnavi, "Super-resolution of optical coherence tomography images by scale mixture models," IEEE Trans. Image Process. **29**, 5662–5676 (2020).

23. Z. Amini, H. Rabbani, and I. Selesnick, "Sparse domain Gaussianization for multi-variate statistical modeling of retinal OCT images," *IEEE Trans. Image Process.* (2020).

24. B. Karamata, K. Hassler, M. Laubscher, and T. Lasser, "Speckle statistics in optical coherence tomography," J. Opt. Soc. Am. A **22**(4), 593 (2005).

25. M. Pircher, E. Götzinger, R. Leitgeb, A. F. Fercher, and C. K. Hitzenberger, "Speckle reduction in optical coherence tomography by frequency compounding," J. Biomed. Opt. **8**(3), 565 (2003).

26. M. Y. Kirillin, G. Farhat, E. A. Sergeeva, M. C. Kolios, and A. Vitkin, "Speckle statistics in OCT images: Monte Carlo simulations and experimental studies," Opt. Lett. **39**(12), 3472 (2014).

27. R. Kafieh, H. Rabbani, and I. Selesnick, "Three dimensional data-driven multi scale atomic representation of optical coherence tomography," IEEE Trans. Med. Imaging **34**(5), 1042–1062 (2015).

28. R. Kafieh, H. Rabbani, M. D. Abramoff, and M. Sonka, "Intra-retinal layer segmentation of 3D optical coherence tomography using coarse grained diffusion map," Med. Image Anal. **17**(8), 907–928 (2013).

29. S. Jorjandi, Z. Amini, G. Plonka, and H. Rabbani, "Isfahan MISP Dataset," Isfahan University of Medical Sciences 2021, https://misp.mui.ac.ir/bank.

30. H. Rabbani, R. Nezafat, and S. Gazor, "Wavelet-domain medical image denoising using bivariate laplacian mixture model," IEEE Trans. Biomed. Eng. **56**(12), 2826–2837 (2009).

31. A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," IEEE Trans. Med. Imaging **20**(8), 772–783 (2001).

32. E. E. Elmahdy and A. W. Aboutahoun, "A new approach for parameter estimation of finite Weibull mixture distributions for reliability modeling," Applied Mathematical Modelling **37**(4), 1800–1810 (2013).

33. H. Rinne, *The Weibull Distribution: A Handbook* (CRC Press, 2008).

34. T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of fMRI time series," IEEE Trans. Med. Imaging **29**(4), 1059–1074 (2010).

35. F. Destrempes, J. Meunier, M.-F. Giroux, G. Soulez, and G. Cloutier, "Segmentation in ultrasonic B-mode images of healthy carotid arteries using mixtures of Nakagami distributions and stochastic optimization," IEEE Trans. Med. Imaging **28**(2), 215–229 (2008).

36. A. Furui, R. Onishi, A. Takeuchi, T. Akiyama, and T. Tsuji, "Non-Gaussianity detection of EEG signals based on a multivariate scale mixture model for diagnosis of epileptic seizures," IEEE Trans. Biomed. Eng. **68**(2), 515–525 (2021).

37. M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Segmentation of skin lesions in 2-D and 3-D ultrasound images using a spatially coherent generalized Rayleigh mixture model," IEEE Trans. Med. Imaging **31**(8), 1509–1520 (2012).

38. J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," IEEE Trans. signal Process. **42**(10), 2664–2677 (1994).

39. T. K. Moon, "The expectation-maximization algorithm," IEEE Signal Process. Mag. **13**(6), 47–60 (1996).

40. M. Li, R. Idoughi, B. Choudhury, and W. Heidrich, "Statistical model for OCT image denoising," Biomed. Opt. Express **8**(9), 3903 (2017).

41. L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu, "Sparsity based denoising of spectral domain optical coherence tomography images," Biomed. Opt. Express **3**(5), 927–942 (2012).

42. M. A. Mayer, A. Borsdorf, M. Wagner, J. Hornegger, C. Y. Mardin, and R. P. Tornow, "Wavelet denoising of multiframe optical coherence tomography data," Biomed. Opt. Express **3**(3), 572–589 (2012).

43. Y. Ma, X. Chen, W. Zhu, X. Cheng, D. Xiang, and F. Shi, "Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN," Biomed. Opt. Express **9**(11), 5129–5146 (2018).

44. M. Esmaeili, A. M. Dehnavi, F. Hajizadeh, and H. Rabbani, "Three-dimensional curvelet-based dictionary learning for speckle noise removal of optical coherence tomography," Biomed. Opt. Express **11**(2), 586–608 (2020).

45. S. Chen and R. Gopinath, "Gaussianization," Adv. Neural Inf. Process. Syst. **13**, 423–429 (2000).

46. A. P. Condurache and A. Mertins, "Elastic-transform based multiclass Gaussianization," IEEE Signal Process. Lett. **18**(8), 482–485 (2011).

47. D. F. D. Martins, "Gaussianization methods for structural brain imaging," Universidade de Coimbra (2014).

48. M. C. Neyrinck, "Transformationally decoupling clustering and tracer bias," Proc. Int. Astron. Union **10**(S306), 251–254 (2014).

49. I. Mezghani-Marrakchi, G. Mahé, S. Djaziri-Larbi, M. Jaidane, and M. T.-H. Alouane, "Nonlinear audio systems identification through audio input Gaussianization," IEEE/ACM Trans. audio, speech, Lang. Process. **22**(1), 41–53 (2013).

50. S. Lyu and E. P. Simoncelli, "Nonlinear extraction of independent components of natural images using radial Gaussianization," Neural Comput. **21**(6), 1485–1519 (2009).

51. X. Zhou, N. Cui, Z. Li, F. Liang, and T. S. Huang, "Hierarchical gaussianization for image classification," in *2009 IEEE 12th International Conference on Computer Vision* (IEEE, 2009), pp. 1971–1977.

52. Y. Kawakami, T. Hattori, D. Kutsuna, H. Matsushita, Y. Imai, H. Kawano, and R. P. C. J. Rajapakse, "An automated color image arrangement method based on histogram matching," in *2013 International Conference on Biometrics and Kansei Engineering* (IEEE, 2013), pp. 31–34.

53. B. Deng and D. Jiang, "Determination of the Weibull parameters from the mean value and the coefficient of variation of the measured strength for brittle ceramics," J. Adv. Ceram. **6**(2), 149–156 (2017).

54. H. Rabbani, M. Vafadust, and S. Gazor, "Image denoising based on a mixture of Laplace distributions with local parameters in complex wavelet domain," in *Proceedings - International Conference on Image Processing,* ICIP (2006), pp. 2597–2600.

55. M. Niknejad, H. Rabbani, and M. Babaie-Zadeh, "Image restoration using Gaussian mixture models with spatially constrained patch clustering," IEEE Trans. Image Process. **24**(11), 3624–3636 (2015).

56. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process. **15**(12), 3736–3745 (2006).

57. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. image Process. **16**(8), 2080–2095 (2007).

58. A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005* (IEEE, 2005), II, pp. 60–65.

59. J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *Proceedings of the IEEE International Conference on Computer Vision* (2015), pp. 244–252.

60. A. Pizurica, L. Jovanov, B. Huysmans, V. Zlokolica, P. De Keyser, F. Dhaenens, and W. Philips, "Multiresolution denoising for optical coherence tomography: a review and evaluation," Curr. Med. Imaging Rev. **4**(4), 270–284 (2008).