

# UC San Diego

## UC San Diego Previously Published Works

### Title

Glaucoma progression detection using nonlocal Markov random field prior

### Permalink

<https://escholarship.org/uc/item/5m93d4j1>

### Journal

Journal of Medical Imaging, 1(3)

### ISSN

2329-4302

### Authors

Belghith, Akram  
Bowd, Christopher  
Medeiros, Felipe A  
[et al.](#)

### Publication Date

2014-12-29

### DOI

10.1117/1.jmi.1.3.034504

Peer reviewed

# Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

## **Glaucoma progression detection using nonlocal Markov random field prior**

Akram Belghith  
Christopher Bowd  
Felipe A. Medeiros  
Madhusudhanan Balasubramanian  
Robert N. Weinreb  
Linda M. Zangwill

# Glaucoma progression detection using nonlocal Markov random field prior

Akram Belghith,<sup>a,\*</sup> Christopher Bowd,<sup>a</sup> Felipe A. Medeiros,<sup>a</sup> Madhusudhanan Balasubramanian,<sup>b,c,d</sup> Robert N. Weinreb,<sup>a</sup> and Linda M. Zangwill<sup>a</sup>

<sup>a</sup>University of California San Diego, Hamilton Glaucoma Center, 9500 Gilman Drive, La Jolla, California 92093-0946, United States

<sup>b</sup>University of Memphis, Department of Electrical and Computer Engineering, 3815 Central Avenue, Memphis, Tennessee 38152 United States

<sup>c</sup>University of Memphis, Department of Biomedical Engineering, 920 Madison Avenue, Memphis, Tennessee 38103 United States

<sup>d</sup>University of Tennessee Health Science Center, Department of Biomedical Engineering, 920 Madison Avenue, Memphis, Tennessee 38103 United States

**Abstract.** Glaucoma is neurodegenerative disease characterized by distinctive changes in the optic nerve head and visual field. Without treatment, glaucoma can lead to permanent blindness. Therefore, monitoring glaucoma progression is important to detect uncontrolled disease and the possible need for therapy advancement. In this context, three-dimensional (3-D) spectral domain optical coherence tomography (SD-OCT) has been commonly used in the diagnosis and management of glaucoma patients. We present a new framework for detection of glaucoma progression using 3-D SD-OCT images. In contrast to previous works that use the retinal nerve fiber layer thickness measurement provided by commercially available instruments, we consider the whole 3-D volume for change detection. To account for the spatial voxel dependency, we propose the use of the Markov random field (MRF) model as a prior for the change detection map. In order to improve the robustness of the proposed approach, a nonlocal strategy was adopted to define the MRF energy function. To accommodate the presence of false-positive detection, we used a fuzzy logic approach to classify a 3-D SD-OCT image into a “non-progressing” or “progressing” glaucoma class. We compared the diagnostic performance of the proposed framework to the existing methods of progression detection. © 2014 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.1.3.034504]

Keywords: glaucoma; nonlocal Markov field; change detection; fuzzy logic classifier.

Paper 14103PR received Aug. 4, 2014; accepted for publication Nov. 20, 2014; published online Dec. 29, 2014.

## 1 Introduction

Glaucoma is a leading cause of visual impairment and blindness around the world.<sup>1</sup> It is characterized by structural changes in the optic nerve head (ONH) (neuroretinal rim and retinal nerve fiber layer (RNFL) thinning and optic nerve cupping are usually the first visible signs), and visual field defects. Glaucoma is painless and gradually damages the optic nerve, which makes patients not aware of any visual loss until the optic nerve is severely damaged. Untreated it can cause an irreversible vision loss, and hence it is critical to early detect glaucoma and to monitor its progression in order to avoid permanent damage to the ONH.

It has now been more than 150 years since physicians were able to visualize damages in the ONH associated with glaucoma using an ophthalmoscope. However, clinical examination of the ONH remains subjective, qualitative, and often with higher variabilities among observers.<sup>2</sup>

Over the past three decades, advances in technology for ocular imaging using the optical properties of the optic nerve and RNFL layer have gained widespread use in the diagnosis and management of glaucoma patients. In particular, the Heidelberg retina tomograph (HRT) (Heidelberg Engineering, Heidelberg, Germany), a confocal scanning laser technology, has been commonly used for glaucoma diagnosis since its commercialization 20 years ago.<sup>3</sup> Several strategies have been

proposed for automated change detection using HRT images. One strategy, the proper orthogonal decomposition,<sup>4</sup> indirectly utilizes the spatial relationship among voxels by controlling the family-wise type I error rate. The Markov model was used in Ref. 5 to model the inter/intra observations dependency allowing a better glaucoma regression detection rate. However, because HRT imaging is limited to ONH surface topography, it cannot directly measure the RNFL thickness.<sup>6</sup>

In order to overcome this drawback, the three-dimensional (3-D) spectral domain optical coherence tomography (SD-OCT), an optical imaging technique, is now the standard of care for obtaining images of both the ONH and RNFL thicknesses. Several studies have been proposed for glaucoma diagnosis using SD-OCT images. Most of the studies use the peripapillary RNFL thickness measurements (i.e., circular B-scans (3.4-mm diameter, 768 A-scans) centered at the optic disk were automatically averaged to reduce speckle noise)<sup>7–11</sup> or the ONH measurements (e.g., rim area, minimum rim width, etc.)<sup>12–14</sup> which are either provided by the commercially available spectral-domain optical coherence tomographers or manually calculated by experts.<sup>13</sup> In Ref. 15, authors used the 3-D volume to extract the whole RNFL thickness map instead of the peripapillary RNFL thickness and proposed a variable-size super pixel segmentation to improve the discrimination between early glaucomatous and healthy eyes. In Ref. 16,

\*Address all correspondence to: Akram Belghith, E-mail: [abelghith@ucsd.edu](mailto:abelghith@ucsd.edu)

authors used the RNFL thickness map to identify and to quantify glaucoma defects. Because of the high reproducibility of the SD-OCT, many studies used the peripapillary RNFL measurements for glaucoma change detection.<sup>17–19</sup> In Ref. 20, authors showed that the instrument built-in segmentation software is relatively robust to the image quality and the noise may lower the accuracy of the RNFL layer thickness estimation. However, the detection of glaucoma progression with optical coherence tomography (OCT) remains a challenge because when assessing structural changes over time, it is difficult to discriminate between glaucomatous structural damage and measurement variability.

In this paper, we propose a new strategy for glaucoma progression detection using 3-D SD-OCT images. This strategy is divided into two steps:

1. *Change detection step*: it consists of detecting changes between a baseline image and a follow-up image. In contrast to previous works that use the RNFL thickness measurement provided by commercially available instruments, we consider the whole 3-D volume for change detection. Note that to the best of our knowledge, this is the first time that the whole 3-D SD-OCT volume is considered for glaucoma progression detection. In this work, we propose a fully Bayesian framework for change detection. Bayesian methods are relatively simple and offer efficient tools to include *priori* through a *posteriori* probability density functions (PDF). In this work, we propose the use of the Markov random field (MRF) to exploit the statistical correlation of intensities among the neighborhood voxels.<sup>21</sup> In particular, consider the use of a nonlocal framework for the MRF energy function definition which has been proposed for the image denoising task.<sup>22</sup> Indeed, the nonlocal approach has been successfully applied in several image processing applications such as image restoration<sup>23</sup> and image segmentation.<sup>24</sup> The main idea of the nonlocal approach is to exploit repetitive structures in the image which leads to a multi-model approach using only a single observation. Moreover, in order to develop a noise robust algorithm, we propose the consideration of the change detection problem as a missing data problem where we jointly estimate the noise hyperparameters and the change detection map. Because we used the MRF model with the change detection map as the prior for the change detection map, the optimization step is intractable. Hence, we propose the use of a Monte Carlo Markov chain (MCMC) technique.<sup>25</sup>
2. *Classification step*: it consists of classifying an SD-OCT image into the “non-progressing” and the “progressing” glaucoma classes using the estimated change detection map. For this, a threshold-based classification method is generally used to accommodate the presence of false-positive detection.<sup>4</sup> However, the choice of the threshold may affect the robustness of the classification method. We will show that the classification results are better when we use a fuzzy set-based classifier compared with other classifiers.

This paper is divided into four sections. In Sec. 2, the proposed glaucoma change detection scheme is presented. In Sec. 3, we describe the classification scheme. Then, in Sec. 4 results obtained by applying the proposed scheme to real data are presented. Specifically, we compare the diagnostic accuracy, robustness, and efficiency of this novel proposed approach to the two existing progression detection RNFL-based approaches: the artificial neural network classifier (ANN) and the support vectors machine classifier.

## 2 Change Detection

Let us consider the detection of changes in a pair of 3-D SD-OCT images. The change detection problem is formulated as a multihypotheses testing problem by denoting the “no-change,” “increase change” (i.e., ONH hypertrophy), and “decrease change” (i.e., ONH atrophy) hypotheses as  $H_0$ ,  $H_1$ , and  $H_2$ , respectively. We denote by  $I_0$  and  $I_1$  two images acquired over the same eye at times  $t_0$  and  $t_1$ , respectively ( $t_1 > t_0$ ), and co-registered. In this paper, we assume that the noise is additive, white, and normally distributed, whereas the noise-free SD-OCT images follow a gamma distribution. Indeed, the gamma distribution  $\mathcal{G}$  is widely used for fitting non-negative data such as the synthetic aperture radar images<sup>26</sup> and the high resolution magic angle spinning nuclear magnetic resonance images<sup>27</sup> due to its shape hyperparameters ( $\alpha > 0, \beta > 0$ ), which allow us to fit spectral data that may present a background.<sup>28</sup> The direct model for both images  $I_0 = [I_0(i)|i = 1, \dots, M]$  and  $I_1 = [I_1(i)|i = 1, \dots, M]$ , where  $M$  is the number of voxels, is then given by

$$I_0 = X_0 + N_0; \quad I_1 = X_1 + N_1, \quad (1)$$

where  $X_0 = [X_0(i)|i = 1, \dots, M]$ ,  $X_0(i) \sim \mathcal{G}(\alpha_0, \beta_0)$ , and  $X_1 = [X_1(i)|i = 1, \dots, M]$ ,  $X_1(i) \sim \mathcal{G}(\alpha_1, \beta_1)$  are the noise-free 3-D SD-OCT ONH images and  $N_0 = [N_0(i)|i = 1, \dots, M]$ ,  $N_0(i) \sim \mathcal{N}(0, \sigma_0)$ ,  $N_1 = [N_1(i)|i = 1, \dots, M]$ , and  $N_1 \sim \mathcal{N}(0, \sigma_1)$  are the Gaussian additive noises. As no simple expression of the distribution of the difference of two gamma distribution variables exists, we adopted for image an ratioing approach instead of the differencing image approach, which generates a ratio image  $R$  resulting in the following direct model:

$$R = \frac{I_0}{I_1} = \frac{X_0 + N_0}{X_1 + N_1}, \quad (2)$$

where  $R = (R(i), i = 1 \dots M)$ . This model is unfortunately intractable. To overcome this problem, we propose a hierarchical change detection framework which consists of estimating the noise-free SD-OCT images  $\hat{X}_0$  and  $\hat{X}_1$  and then using the image ratio approach for change detection. The new direct model is then given by  $R = \hat{X}_0/\hat{X}_1$ . The gamma ratio distribution is expressed as

$$p(R(i); \alpha_0, \beta_0, \alpha_1, \beta_1) = \frac{\Gamma(\alpha_0 + \alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} R(i)^{\alpha_0-1} \left(\frac{\beta_0}{\beta_1}\right)^{\alpha_0} \times \left(R(i) + \frac{\beta_0}{\beta_1}\right)^{-(\alpha_0+\alpha_1)}. \quad (3)$$

**Proof.** The PDF of the ratio  $R = (X_0/X_1)$  is given by

$$\begin{aligned} & p(R(i); \alpha_0, \beta_0, \alpha_1, \beta_1) \\ &= \int_0^\infty \mathcal{G}(X_1(i)R(i), \alpha_0, \beta_0) \mathcal{G}(X_1(i), \alpha_1, \beta_1) X_1(i) dX_1(i) \\ &= \frac{R(i)^{\alpha_0-1}}{\Gamma(\alpha_1)\beta_1^{\alpha_1}} \frac{1}{\Gamma(\alpha_0)\beta_0^{\alpha_0}} \left(\frac{R(i)}{\beta_0} + \frac{1}{\beta_1}\right)^{\alpha_0+\alpha_1} \Gamma(\alpha_1 + \alpha_0) \\ &\quad \times \int_0^\infty \frac{1}{\left(\frac{R(i)}{\beta_0} + \frac{1}{\beta_1}\right)^{\alpha_0+\alpha_1} \Gamma(\alpha_1 + \alpha_0)} X_1^{\alpha_0+\alpha_1-1} \\ &\quad \times \exp\left(-X_1(i) \left(\frac{r}{\beta_0} + \frac{1}{\beta_1}\right)\right) dX_1 \\ &= \frac{\Gamma(\alpha_0 + \alpha_1)}{\Gamma(\alpha_0)\Gamma(\alpha_1)} R(i)^{\alpha_0-1} \left(\frac{\beta_0}{\beta_1}\right)^{\alpha_0} \left(R(i) + \frac{\beta_0}{\beta_1}\right)^{-(\alpha_0+\alpha_1)}, \quad (4) \end{aligned}$$

since the PDF integrates to 1.

For the sake of simplicity, we assume that  $\alpha_0 = \alpha_1 = \alpha$  and  $\beta = \beta_0/\beta_1$ ,  $p(R(i); \alpha, \beta)$  is then given by

$$p(R(i); \alpha, \beta) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} r(i)^{\alpha-1} \beta^\alpha (r(i) + \beta)^{-2\alpha}. \quad (5)$$

The change detection is handled through the introduction of change class assignments  $Q = [Q(i)|i = 1, \dots, M]$ . Hence, the posterior probability distribution of change detection map  $Q$  at each voxel location (i.e.,  $Q(i)$ ) is given by

$$p(Q_i = H_j | R, \Theta) = \frac{1}{Z} \exp(-U(q_i = H_j | R, \Theta)), \quad (6)$$

where  $Z$  is the normalization constant,  $\Theta$  consists of the model hyperparameters and  $U(H_j | R, \Theta)$ , and  $j \in \{0, 1, 2\}$  is the energy function of the MRF model. In the case of an isotropic Potts model-type prior,<sup>29</sup> the corresponding posterior energy to be minimized is

$$\begin{aligned} U(Q_i = H_j | R, \alpha_j, \beta_j, \zeta) &= \sum_{l \in S_i} \{\log[p_l(R(l) | Q(l) = H_j; \alpha_j, \beta_j)]\} \\ &\quad + \zeta \sum_{l \in S_i^*} \delta(q_i, q_l), \quad (7) \end{aligned}$$

where  $p_l(R(l) | Q(l) = H_j)$  is the gamma ratio probability density function of the  $l$ 'th amplitude ratio  $R(l)$  belonging to the source neighboring system  $S_i$  given the class  $H_j$ ,  $(\alpha_j, \beta_j)$  is the set of the gamma ratio distribution hyperparameters of the class  $H_j$ ,  $j \in \{0, 1, 2\}$ ,  $\delta$  is the delta Kroneker function, and  $\zeta$  is a positive parameter. Note that we opted for the 3-D eight-connectivity neighboring system.

In this work, we adopt a nonlocal approach to define the energy function. The nonlocal approach aims to take advantage of the redundancy presented in the images by extending the neighboring system. This is based on the assumption that a

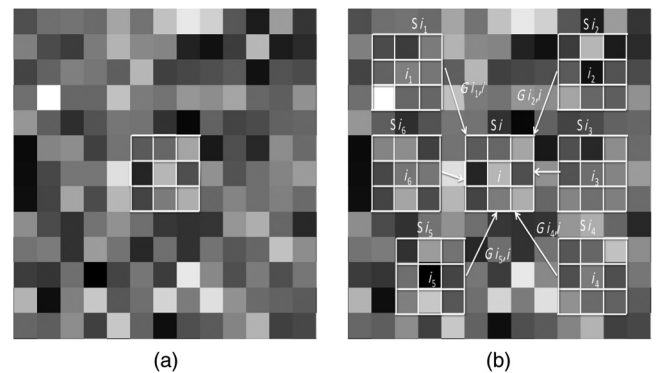
given neighborhood around a voxel may be similar to the neighborhoods around other voxels in the same image. Figure 1 presents an example of the classical and the nonlocal neighboring systems. In contrast to the classical approach, the nonlocal approach allows us to exploit the information available in other neighborhoods called patches. The contribution of each new neighborhood or patch depends on the similarity between the main neighborhood  $S_i$  and the other patches  $S_{i_d}$  where  $i_d$  is the center of the patch [Fig. 1(b)]. The more similar they are, the greater their contribution is. To this end, a weighted graph  $G$  that links together voxels over the image is calculated. In Ref. 30, authors showed that the Euclidean distance is sufficient to reliably measure the similarity. Hence, the similarity between two voxels  $i$  and  $i_d$ , based on the pairwise Euclidean distance, is expressed as follows:

$$G_{i_d, i} = \sum_{l_1 \in S_{i_d}, l_2 \in S_i} (R(l_1) - R(l_2))^2. \quad (8)$$

The new energy function, which integrates the non-local weights  $G$  associated with different patches around the voxel of interest  $i$  and located in an extended neighborhood  $NS_i$ , is expressed as follows:

$$\begin{aligned} U(Q_i = H_j | R, \alpha_j, \beta_j, \zeta, \eta) &= \eta \sum_{S_n \in NS_i} G_{n, i} \sum_{l \in S_n} \{\log[p_l(R(l) | q(l) = H_j, \alpha_j, \beta_j)]\} \\ &\quad + \zeta \sum_{l \in S_i^*} \delta(Q_i, Q_l). \quad (9) \end{aligned}$$

The first part of the energy function models the *a priori* we have on the image ratio  $R(l)$  conditioned to change and no-change hypotheses as well as the information available in other patches (information fusion). The second part models the spatial dependency by the use of the second-order isotropic Potts model with parameter  $\zeta$ . Hence, the definition of the energy function favors the generation of homogeneous areas reducing the impact of the speckle noise, which could affect the classification results of the SD-OCT images.<sup>31</sup> The hyperparameters  $\zeta, \eta$  handle the importance of the energy terms. On one hand,  $\zeta$  allows us to tune the importance of the spatial dependency constraint. In particular, high values of  $\zeta$  correspond to a high spatial constraint. On the other hand,  $\eta$  models



**Fig. 1** Comparison of (a) the classical approach (b) the nonlocal approach where the different neighborhoods/patches  $S_{i_d, d \in \{1, 2, 3, 4, 5, 6\}}$  will contribute to the estimation of the voxel  $i$  based and the their similarities  $G_{i_d, i}$ .

the reliability of the patches and is usually assumed to take on values in  $[0, 1]$ . This constraint can easily be satisfied by choosing the appropriate *a priori* distribution for  $\eta$ .

The whole set of the model hyperparameters is then given by  $\Theta = \{\sigma_{0,1}, \alpha_{0,1}, \beta_{0,1}, \alpha_j, \beta_j, \eta, \zeta\}$ . Without information prior knowledge, the following priors are retained:  $p(\sigma_0^2) = \frac{1}{\sigma_0}$ ,  $p(\sigma_1^2) = \frac{1}{\sigma_1}$ . The non-negativity of the the hyperparameters  $\{\alpha_{0,1}, \beta_{0,1}, \alpha_j, \beta_j, \eta, \zeta\}$  is guaranteed through the use of the exponential densities:  $p(\alpha_0, \beta_0) = \frac{1}{\varphi} \exp(-\alpha_0/\varphi) \frac{1}{\vartheta} \exp(-\beta_0/\vartheta)$ ,  $p(\alpha_1, \beta_1) = \frac{1}{\varphi} \exp(-\alpha_1/\varphi) \frac{1}{\vartheta} \exp(-\beta_1/\vartheta)$ ,  $p(\alpha_j, \beta_j) = \frac{1}{\varphi} \exp(-\alpha_j/\varphi) \frac{1}{\vartheta} \exp(-\beta_j/\vartheta)$ ,  $p(\eta) = \frac{1}{\kappa} \exp(-\eta/\kappa)$ , and  $p(\zeta) = \frac{1}{\kappa} \exp(-\zeta/\kappa)$ . Note that the values of  $\{\varphi, \vartheta, \kappa\}$  are empirically fixed but do not really influence the results. The hierarchical approach we proposed and the prior knowledge we opted for our model lead to an intractable posterior distribution  $p(Q, \Theta|R)$ . Therefore, we propose the use of an MCMC procedure to estimate the model parameters and hyperparameters. Indeed, the principle of the MCMC method is to generate samples drawn from the posterior densities and then to be able to achieve parameter estimation. We use a Gibbs sampler based on a stationary ergodic Markov chain allowing us to draw samples whose distribution asymptotically follows the *a posteriori* densities. Indeed, the Gibbs sampler decomposes the problem of generating a sample from a high dimension PDF by simulating each variable separately according to its conditional PDF. Since no classical expressions for the posterior PDFs are available, there is no direct simulation algorithm for them. To overcome this problem, we perform Hastings–Metropolis steps within the Gibbs sampler<sup>25</sup> to ensure that all convergence properties are preserved.<sup>32</sup> Moreover, since no knowledge on the posterior definition is available, we have opted for the random walk version of the Hastings–Metropolis algorithm. The proposal distribution  $p$  is then defined as  $p(\cdot|Y) = Y + p'(\cdot)$  where  $p'$  is a distribution which does not depend on  $Y$  and is usually centered on 0. Hence, at each iteration, a random move is proposed from the actual position. The choice of a good proposal distribution is crucial to obtain an efficient algorithm, and the literature suggests that, for low dimension variables, a good proposal should lead to an acceptance rate of 0.5.<sup>25</sup> The determination of a good proposal distribution can be solved by using a standard Gaussian distribution with an adaptive scale technique.<sup>33</sup> The main Gibbs sampler steps are described in Algorithm 1. We

#### Algorithm 1 Sampling Algorithm.

1. Initialization of  $X_0^{[0]}, X_1^{[0]}, Q^{[0]}$  and  $\Theta^{[0]}$
2. For each iteration  $h$  repeat:
  - i) Sampling  $X_0^{[h]}$  from  $p(X_0|I_0, \Theta^{[h-1]})$
  - ii) Sampling  $X_1^{[h]}$  from  $p(X_1|I_1, \Theta^{[h-1]})$
  - iii) Calculating  $R^{[h]} = X_0^{[h]}/X_1^{[h]}$
  - iv) Creating a configuration of  $Q$  basing on  $R^{[h]}$
  - v) Calculating  $p^{[h]}(Q = H_1|R^{[h]}, \Theta^{[h-1]})$
  - vi) Sampling  $\Theta^{[h]}$  from  $p(\Theta|X_0^{[h]}, X_1^{[h]}, I_0, I_1)$
  - vii) Until convergence criterion is satisfied

used a burn-in period of  $h_{\min} = 500$  iterations followed by another 1000 iterations for convergence ( $h_{\max} = 1500$ ). The change detection map  $Q$  is estimated using the maximum *a posteriori* MAP estimator:  $\hat{Q} = \underset{H_1}{\operatorname{argmax}} \bar{p}_{H_1}$ , where

$$\bar{p}_{H_1} = \frac{1}{h_{\max} - h_{\min}} \sum_{h=h_{\min}+1}^{h_{\max}} p^{[h]}(Q=H_1|R^{[h]}, \Theta^{[h-1]}), \quad \bar{p}_{H_2} = \frac{1}{h_{\max} - h_{\min}} \sum_{h=h_{\min}+1}^{h_{\max}} p^{[h]}(Q=H_2|R^{[h]}, \Theta^{[h-1]}) \text{ and } \bar{p}_{H_3} = 1 - \bar{p}_{H_3} - \bar{p}_{H_1}.$$

### 3 Classification

This step aims at classifying an image into the nonprogressor or the progressor classes based on the estimated change detection map. In contrast to Ref. 4, we considered two-layer fuzzy classifier rather than a threshold-based classifier. Indeed, the fuzzy set theory is used to quantify the membership degrees of a given image to each class (i.e. progressor and nonprogressor).

As in Ref. 4, we considered two features as input for the classifier: (1) feature 1: the number of changed sites and (2) feature 2: the ratio image intensity  $R$  of changed sites. In this work, only the loss of retinal height in neighboring areas is considered change due to glaucomatous progression because an increase in retinal height is considered improvement (possibly due to treatment or tissue rearrangement).

We now calculate the elementary membership degree to the nonprogressor class given each feature  $\gamma_{\text{nor}, o \in \{\text{feature1}, \text{feature2}\}}$  using an  $S$ -membership function  $f$  whose expression is given in Eq. (10). Note that the range  $[a, c]$  defines the fuzzy region.

$$f(x; a, b, c) = \begin{cases} 0 & x < a \\ \frac{(x-a)^2}{(b-a)(c-a)} & a \leq x < b \\ 1 - \frac{(x-c)^2}{(c-b)(c-a)} & b \leq x < c \\ 1 & \text{otherwise} \end{cases}, \quad (10)$$

where  $a < b < c$ .

$$\gamma_{\text{nor}} = \gamma_{\text{nor}, \text{feature1}} \times \gamma_{\text{nor}, \text{feature2}}, \quad (11)$$

where  $\gamma_{\text{nor}, \text{feature1}} = f(N_c; a_1, b_1, c_1)$ ,  $\gamma_{\text{nor}, \text{feature2}} = f(\sum_{i \in C} R_i; a_2, b_2, c_2)$ ,  $N_c$  and  $C$  stand for the number of changed sites and the changed site class, respectively, and  $(a_1, b_1, c_1, a_2, b_2, c_2)$  are the hyperparameters of the  $S$ -membership functions. The hyperparameters  $(a_1, b_1, c_1, a_2, b_2, c_2)$  are estimated with the genetic algorithms<sup>34</sup> using longitudinal SD-OCT data from a training dataset which contains 10 normal eyes, 5 nonprogressing eyes, and 10 progressing eyes. Note that the training dataset is independent of the test dataset described in Sec. 4.1.

The membership degree to the glaucoma class  $\gamma_{\text{glau}}$  is given by

$$\gamma_{\text{glau}} = 1 - \gamma_{\text{nor}}. \quad (12)$$

To decide if a given SD-OCT image belongs to the glaucoma progressor class, another membership function is used. We opted for the trapezoidal function denoted by  $g$  as a membership function. The expression of  $g$  is given by

$$g(x; a_3, b_3, c_3, d_3) = \begin{cases} \left(\frac{x-a_3}{b_3-a_3}\right) & a_3 \leq x < b_3 \\ 1 & b_3 \leq x < c_3 \\ \left(\frac{d_3-x}{d_3-c_3}\right) & c_3 \leq x < d_3 \\ 0 & \text{otherwise} \end{cases}, \quad (13)$$

where  $a_3 < b_3 < c_3 < d_3$ . Alternately, the genetic algorithm was used to estimate this quadruple. The decision to classify an image into the glaucoma class depends on the output of the function  $g(\gamma_{\text{glau}})$ . As can be observed, the  $g$  function depends on the quadruple  $(a_3, b_3, c_3, d_3)$  as well as on  $(a_2, b_2, c_2)$  and  $(a_1, b_1, c_1)$ . If  $g(\gamma_{\text{glau}}) = 1$  the image is assigned to the glaucoma progressor class. Note that an eye is considered as a progressor if one follow-up image is a progressor.

## 4 Experiments

The different datasets used for validation are described in the next sub-section 4.1. Then, change detection results on simulated and semi-simulated datasets are presented in sub-section 4.2. Finally, the classification results on clinical datasets are presented in sub-section 4.3.

### 4.1 Datasets

The proposed framework was experimentally validated on simulated, semi-simulated, and clinical datasets. Our clinical datasets consist of 117 eyes of 75 participants. Table 1 presents a description of the clinical dataset. Specifically, the first clinical dataset consists of 27 eyes from 27 participants with glaucoma progression. The glaucomatous progression was defined based on: (1) likely progression by visual field<sup>35</sup> or (2) progression by stereophotographic assessment of the optic disk.<sup>36</sup> A third observer adjudicated any differences in assessment between these two observers. The second clinical dataset consists of 40 eyes from 23 healthy participants. The third clinical dataset consists of 50 eyes from 26 participants considered stable (each participant was imaged once a week for five consecutive weeks). All eligible participants were recruited from the University of California, San Diego Diagnostic Innovations in Glaucoma Study (DIGS) with at least five good quality visual field exams to ensure an accurate diagnosis. Note that for each exam, the baseline image and the follow-up image are automatically co-registered with built in instrument software. Specifically, the Spectralis SD-OCT (Heidelberg Engineering GmbH, Heidelberg, Germany) instrument features two different options to enhance reproducibility. A real-time eye-tracking device (eye tracker) compensates for involuntary eye

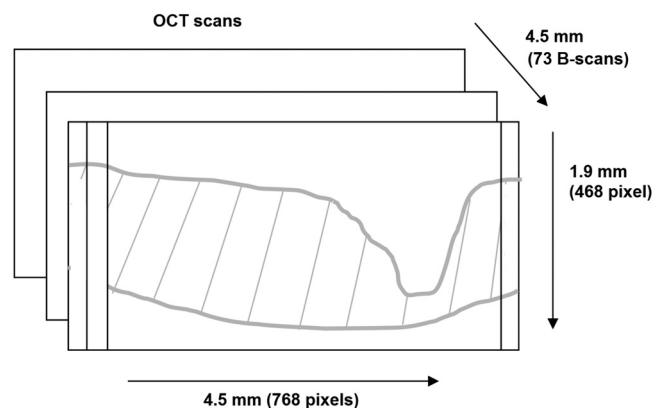
**Table 1** Description of the clinical dataset used to assess the diagnosis accuracy of the proposed framework.

	Number of patients	Number of eyes	Follow-up duration (median/interquartile range)
Normal group	23	40	2.03 (1.8) years
Stable glaucoma group	50	26	5 (0) weeks
Progressing glaucoma group	27	27	2.4 (1.6) years

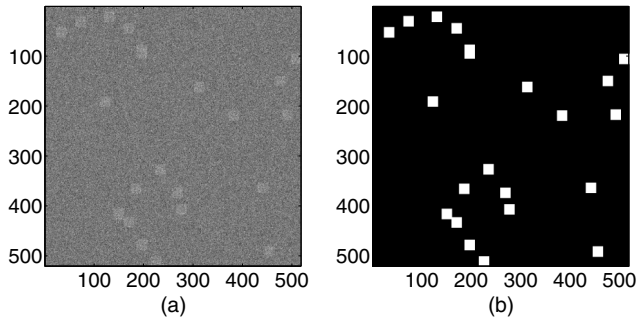
movements during the scanning process, and a retest function assures that follow-up measurements are taken from the same area of the retina as the baseline examination. Many studies demonstrate that the retest recognition and eye tracker options led to high reproducibility and accurate and repeatable alignment of the baseline and the follow-up images.<sup>36-38</sup>

Because the performance of any change detection algorithm can be affected by the OCT instrument specifications<sup>39</sup> (axial/lateral resolution, voxel sizes, etc.), a test with low measurement variability is important for detecting change during longitudinal analysis. This is of particular importance in the evaluation of glaucoma progression where visit-to-visit changes may be very small. Several studies have shown that Spectralis SD-OCT can reliably measure very small changes.<sup>37,40,41</sup> Using the high resolution mode and a  $(4.5 \times 1.9 \times 4.5 \text{ mm})$  pattern size centered on the ONH, we obtained an axial and lateral resolution of 3.87 and 5.84  $\mu\text{m}$ , respectively, where the distance between two consecutive B-scans is 62  $\mu\text{m}$  (c.f. Fig. 2). Therefore, each 3-D OCT images consists of  $(768 \times 468 \times 73)$  voxels. All experiments were performed on a workstation with Intel Core i7, 1.6 GHz CPU and 16 GB RAM memory. MATLAB® 7.14.0 (R2012a) and C++ were used for processing. The processing time is approximately 5 min per image.

Because no ground truth for glaucomatous change in the 3-D SD-OCT images is available, we generated simulated and semi-simulated datasets to assess the proposed change detection method. The simulated dataset consists of 100 images generated according to the proposed model with different values of peak signal to noise ratio (PSNR) (30, 25, 20, and 15 dB) ( $\text{PSNR} = 10 \log_{10}\{\max(\text{Image})^2/E[(\text{Noise})^2]\}$ ). Figure 3 presents an example of a simulated image. The semi-simulated dataset was constructed from 100 normal 3-D SD-OCT images. Changes were simulated (1) by the permuting 1%, 2%, 3%, and 5% of image regions in the images, respectively, and (2) by modifying the intensities of each  $15 \times 15 \times 3$  voxel-sized region in the images randomly with 0.5%, 0.75%, 1%, and 1.25% probabilities, respectively. The intensities were randomly modified by multiplying the real intensities by 0.5, 0.75, 1.5, or 2 with a 0.25% probability. In order to emphasize the robustness of the proposed approach, an additive Gaussian noise with different values of PSNR (30, 25, 20, and 15 dB) was added to both the original image and the semi-simulated images. Figure 4 presents an example of a semi-simulated image. Table 2 summarizes the simulated and semi-simulated datasets.



**Fig. 2** Dimension of the three-dimensional (3-D) optical coherence tomography (OCT) images.



**Fig. 3** Simulated image: (a) the simulated image (PSNR = 20 dB) and (b) the ground truth; the white boxes represent the changes.

## 4.2 Change Detection Algorithm Assessment

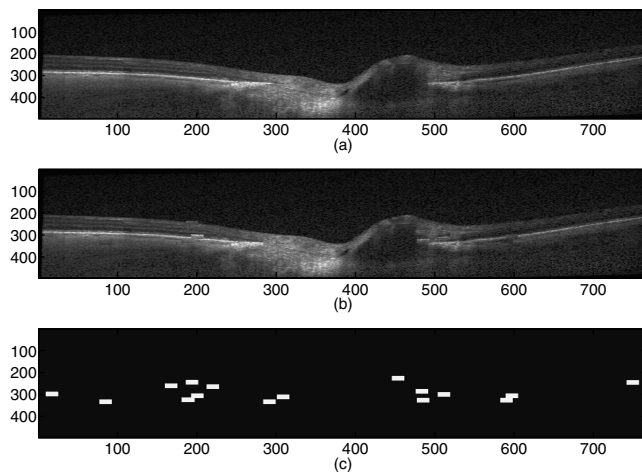
In order to emphasize the benefit of the proposed change detection approach called the nonlocal Bayesian Markovian approach and particularly the use of the nonlocal approach to define the MRF energy function, we compared the proposed nonlocal Bayesian Markovian approach with three change detection methods:

1. The proposed method with a classic MRF neighboring system
2. The kernel  $k$ -means algorithm<sup>6</sup> with the RBF Gaussian kernel
3. A Bayesian threshold-based method.<sup>42</sup>

To perform the change detection evaluation, we use the percentage of false alarm (PFA), the percentage of missed detection (PMD) and the percentage of total error (PTE) measurements defined by

$$\text{PFA} = \frac{\text{FA}}{N_F} \times 100\%; \quad \text{PMD} = \frac{\text{MD}}{N_M} \times 100\%;$$

$$\text{PTE} = \frac{\text{MD} + \text{FA}}{N_M + N_F} \times 100\%,$$



**Fig. 4** Semi-simulated image by the permutation of image regions: (a) the original B-scan image, (b) the simulated B-scan image, (c) the ground truth: the white boxes represent the changes.

**Table 2** Description of the simulated and semi-simulated dataset used to evaluate the proposed nonlocal Bayesian Markovian approach for change detection.

	Number of images
Simulated dataset	100
Semi-simulated dataset by region permutation	100
Semi-simulated dataset by intensity modification	100

where FA stands for the number of unchanged voxels that were incorrectly determined as changed ones,  $N_F$  stands for the total number of unchanged voxels, MD stands for the number of changed voxels that were mistakenly detected as unchanged ones, and  $N_M$  stands for the total number of changed voxels. The kappa coefficient of agreement between the ground truth and the change detection map was also calculated. The kappa statistic was originally developed by Cohen<sup>43</sup> in order to discern the amount of agreement that could be expected to occur through chance. A detailed description of the kappa coefficient of agreement can be found.<sup>44</sup> Strength of agreement can be categorized according to the method proposed in Ref. 45: 0 = poor, 0 to 0.20 = slight, 0.21 to 0.40 = fair, 0.41 to 0.60 = moderate, 0.61 to 0.80 = substantial, and 0.81 to 1.00 = almost perfect. Table 3 shows the results of the change detection on simulated and semi-simulated datasets. The nonlocal Bayesian Markovian approach (kappa = 0.84) tends to perform better than the kernel  $k$ -means (kappa = 0.71), classic MRF approach (kappa = 0.78), and the Bayesian threshold-based (kappa = 0.64) methods. This means that the proposed nonlocal Markov *priori* we considered improves the change detection results by modeling the spatial dependency of voxels as well as the redundancy presented in the images by extending the neighboring system.

## 4.3 Classification Results

This section describes the glaucoma progression detection results obtained with the proposed scheme. In order to perform the classification evaluation, we have retained the area under the receiver operating characteristic area under the receiver operating characteristic curve (AUROC), the sensitivity and the specificity measurements:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}; \quad \text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

where TP stands for the number of true positive identifications, FN stands for the number of false-negative identifications, and FP stands for the number of false-positive identifications.

The proposed framework was experimentally validated with clinical datasets. It is important to note that we used independent training and test sets to estimate the diagnostic accuracy of the methods. Specifically, we trained our classifiers, as well as the other classifiers used for the comparison, on 10 normal eyes, 5 stable glaucoma eyes, and 10 progressor eyes. In order to emphasize the benefit of the proposed glaucoma progression detection scheme, we have compared the proposed framework



**Table 3** Mean and standard deviation (SD) of false detection, missed detection, total errors, and the kappa coefficient resulting from the proposed nonlocal Bayesian Markovian approach, the classic Markov random field (MRF) approach, the kernel  $k$ -means method and the threshold method using simulated and semi-simulated.

Semi-simulated datasets	False detection ( $\pm$ SD)	Missed detection ( $\pm$ SD)	Total errors ( $\pm$ SD)	Kappa
Nonlocal Bayesian Markovian approach	<b>0.37</b> ( $\pm$ 0.14) %	<b>3.12</b> ( $\pm$ 0.65) %	<b>0.61</b> ( $\pm$ 0.13) %	<b>0.84</b>
Classic MRF approach	0.54 ( $\pm$ 0.21) %	5.62 ( $\pm$ 1.03) %	0.91 ( $\pm$ 0.20) %	0.78
Kernel $k$ -means	0.89 ( $\pm$ 0.41) %	8.03( $\pm$ 2.75) %	2.03( $\pm$ 0.51) %	0.71
Threshold	2.14( $\pm$ 0.71) %	10.52( $\pm$ 3.14) %	3.45( $\pm$ 1.24) %	0.64
Simulated dataset	False detection ( $\pm$ SD)	Missed detection ( $\pm$ SD)	Total errors ( $\pm$ SD)	Kappa
Nonlocal Bayesian Markovian approach	<b>0.84</b> ( $\pm$ 0.24) %	<b>5.17</b> ( $\pm$ 1.05) %	<b>0.97</b> ( $\pm$ 0.17) %	<b>0.82</b>
Classic MRF approach	1.24 ( $\pm$ 0.35) %	7.05 ( $\pm$ 1.95) %	1.51 ( $\pm$ 0.34) %	0.72
Kernel $k$ -means	1.65 ( $\pm$ 0.63) %	8.89 ( $\pm$ 2.87) %	2.38 ( $\pm$ 2.44) %	0.69
Threshold	3.78 ( $\pm$ 1.47) %	15.25 ( $\pm$ SD4.17) %	6.41 ( $\pm$ 2.74) %	0.53

Note: Bold values correspond to the best values.

called as nonlocal Bayesian-fuzzy detection scheme (BFDS) to six other methods:

1. The support vector machine (SVM) classifier of the RNFL thickness. As in Ref. 46, we used the radial basis function as a kernel. The SVM was trained by a variation of Platt's sequential minimal optimization algorithm.<sup>47</sup> The SVM hyperparameters were determined by a global optimization technique based on simulated annealing<sup>48</sup> (RNFL-SVM).
2. The ANN classifier of the RNFL thickness. As in Ref. 46, we used the multiayer perceptrons version of the ANN<sup>49</sup> (RNFL-ANN).
3. The proposed method with the classic MRF *a priori* knowledge on the change detection map standard BFDS.

Results are presented in Table 4. The nonlocal BFDS method with the use of the whole 3-D SD-OCT volume instead of the

**Table 4** Diagnostic accuracy of the nonlocal BFDS, the standard BFDS, the RNFL-ANN, and the RNFL-SVM methods.

	Progressing glaucoma group sensitivity (%)	Normal group specificity (%)	Stable group specificity (%)	AUROC
Nonlocal BFDS	<b>70</b>	<b>94</b>	<b>94</b>	<b>0.87</b>
standard BFDS	64	90	92	0.83
RNFL-ANN	69	71	78	0.69
RNFL-SVM	52	68	79	0.60

Note: Bold values correspond to the best values.

RNFL measurements results in high specificity in both normal and stable glaucoma eyes (94% and 94% respectively) while maintaining good sensitivity (70%). This can be explained by the fact that we take the noise for change detection instead of the machine measurements which strongly depend on the image quality. Moreover, the nonlocal approach led to a better sensitivity compared to the classic approach. This is due to the accuracy of the changes detection of the nonlocal approach. It is important to note that the specificity is generally considered more important in glaucoma progression detection than the sensitivity since glaucoma is a slow disease<sup>50</sup> and false negatives (people falsely identified as nonprogressing) can be eliminated in the next follow-up visits. However, if the specificity is low, this would lead to substantial follow-up costs, unnecessary treatment (e.g., surgery) and may cause much distress for people who are not progressing.

## 5 Conclusion

A new framework for glaucoma progression detection has been proposed. We particularly focus on the formulation of the change problem as a missing data problem. The task of inferring the glaucomatous changes is tackled with a hierarchical MCMC algorithm that is used for the first time to our knowledge in the glaucoma diagnosis framework. The new nonlocal approach we proposed to define the MRF energy function increased the robustness of the proposed change detection scheme compared to the classic method. The validation using clinical data of the proposed approach has shown that it has a high diagnostic accuracy for detecting glaucoma progression compared with other methods.

## Acknowledgments

The authors acknowledge the funding and support of the National Eye Institute (Grant Nos. U10EY14267, EY019869, EY021818, EY022039 and EY08208, EY11008, P30EY022589 and EY13959).

## References

1. I. Dielemans et al., "Primary open-angle glaucoma, intraocular pressure, and diabetes mellitus in the general elderly population. The Rotterdam study," *Ophthalmology* **103**(8), 1271 (1996).
2. G. Jaffe and J. Caprioli, "Optical coherence tomography to detect and manage retinal disease and glaucoma," *Am. J. Ophthalmol.* **137**(1), 156–169 (2004).
3. A. Mistlberger et al., "Heidelberg retina tomography and optical coherence tomography in normal, ocular-hypertensive, and glaucomatous eyes," *Ophthalmol.* **106**(10), 2027–2032 (1999).
4. M. Balasubramanian et al., "Localized glaucomatous change detection within the proper orthogonal decomposition framework," *Invest. Ophthalmol. Vis. Sci.* **53**(7), 3615–3628 (2012).
5. A. Belghith et al., "A Bayesian framework for glaucoma progression detection using Heidelberg retina tomograph images," *Int. J. Adv. Comput. Sci. Appl.* **4**(9), 223–229 (2013).
6. T. C. Chen, "Spectral domain optical coherence tomography in glaucoma: qualitative and quantitative analysis of the optic nerve head and retinal nerve fiber layer (an AOS thesis)," *Trans. Am. Ophthalmol. Soc.* **107**, 254–281 (2009).
7. J.-C. Mwanza et al., "Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes," *Ophthalmology* **118**(2), 241–248 (2011).
8. J. W. Cho et al., "Detection of glaucoma by spectral domain-scanning laser ophthalmoscopy/optical coherence tomography (SD-SLO/OCT) and time domain optical coherence tomography," *J. Glaucoma* **20**(1), 15–20 (2011).
9. M. Seong et al., "Macular and peripapillary retinal nerve fiber layer measurements by spectral domain optical coherence tomography in normal-tension glaucoma," *Invest. Ophthalmol. Vis. Sci.* **51**(3), 1446–1452 (2010).
10. M. T. Leite et al., "Comparison of the diagnostic accuracies of the spectralis, cirrus, and rtvue optical coherence tomography devices in glaucoma," *Ophthalmology* **118**(7), 1334–1339 (2011).
11. J. Kotowski et al., "Glaucoma discrimination of segmented cirrus spectral domain optical coherence tomography (SD-OCT) macular scans," *Br. J. Ophthalmol.* **96**(11), 1420–1425 (2012).
12. S. K. Gardiner et al., "A method to estimate the amount of neuroretinal rim tissue in glaucoma: comparison with current methods for measuring rim area," *Am. J. Ophthalmol.* **157**(3), 540–549 (2014).
13. B. C. Chauhan et al., "Enhanced detection of open-angle glaucoma with an anatomically accurate optical coherence tomography-derived neuroretinal rim parameter," *Ophthalmology* **120**(3), 535–543 (2013).
14. N. Patel, M. Sullivan-Mee, and R. S. Harwerth, "The relationship between retinal nerve fiber layer thickness and optic nerve head neuroretinal rim tissue in glaucoma," *Invest. Ophthalmol. Vis. Sci.* **55**(10), 6802–6816 (2014).
15. J. Xu et al., "Three-dimensional spectral-domain optical coherence tomography data analysis for glaucoma detection," *PLoS One* **8**(2), e55476 (2013).
16. J. W. Shin et al., "Diffuse retinal nerve fiber layer defects identification and quantification in thickness maps," *Invest. Ophthalmol. Vis. Sci.* **55**(5), 3208–3218 (2014).
17. G. Savini, M. Carbonelli, and P. Barboni, "Spectral-domain optical coherence tomography for the diagnosis and follow-up of glaucoma," *Curr. Opin. Ophthalmol.* **22**(2), 115–123 (2011).
18. N. G. Strouthidis et al., "Longitudinal change detected by spectral domain optical coherence tomography in the optic nerve head and peripapillary retina in experimental glaucoma," *Invest. Ophthalmol. Vis. Sci.* **52**(3), 1206–1219 (2011).
19. J. M. Wessel et al., "Longitudinal analysis of progression in glaucoma using spectral-domain optical coherence tomography," *Invest. Ophthalmol. Vis. Sci.* **54**(5), 3613–3620 (2013).
20. M. Balasubramanian et al., "Effect of image quality on tissue thickness measurements obtained with spectral-domain optical coherence tomography," *Opt. Express* **17**(5), 4019 (2009).
21. S. Li, *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, New York (1995).
22. A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**(2), 490–530 (2005).
23. J. Sun and M. F. Tappen, "Learning non-local range Markov random field for image restoration," in *Proc. 2011 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Providence, Rhode Island pp. 2745–2752, IEEE, Colorado Springs (2011).
24. B. Caldairou et al., "A non-local fuzzy segmentation method: application to brain mri," *Pattern Recognit.* **44**(9), 1916–1927 (2011).
25. W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, UK (1996).
26. C. Tison et al., "A new statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.* **42**(10), 2046–2057 (2004).
27. A. Belghith et al., "A unified framework for peak detection and alignment: application to HR-MAS 2D NMR spectroscopy," *Signal, Image Video Process.*, **7**(5), 833–842 (2013).
28. N. Dobigeon et al., "Bayesian separation of spectral sources under non-negativity and full additivity constraints," *Signal Process.* **89**(12), 2657–2669 (2009).
29. J. Besag, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc.* **48**(3), 259–302 (1986).
30. P. Coupé et al., "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *IEEE Trans. Med. Imaging* **27**(4), 425–441 (2008).
31. M. Hangai et al., "Ultrahigh-resolution versus speckle noise-reduction in spectral-domain optical coherence tomography," *Opt. Express* **17**(5), 4221–4235 (2009).
32. A. Gelman, G. Roberts, and W. Gilks, "Efficient metropolis jumping rules," *Bayesian Stat.* **5**, 599–608 (1996).
33. W. R. Gilks, G. O. Roberts, and S. K. Sahu, "Adaptive Markov chain Monte Carlo through regeneration," *J. Am. Stat. Assoc.* **93**(443), 1045–1054 (1998).
34. D. Goldberg, *Genetic Algorithms in Search and Optimization*, Addison-Wesley, Boston, Massachusetts (1989).
35. M. C. Leske et al., "Early manifest glaucoma trial: design and baseline data," *Ophthalmology* **106**(11), 2144–2153 (1999).
36. N. O'Leary et al., "Glaucomatous progression in series of stereoscopic photographs and Heidelberg retina tomograph images," *Arch. Ophthalmol.* **128**(5), 560–568 (2010).
37. S. J. Langenegger, J. Funk, and M. Töteberg-Harms, "Reproducibility of retinal nerve fiber layer thickness measurements using the eye tracker and the retest function of spectralis SD-OCT in glaucomatous and healthy control eyes," *Invest. Ophthalmol. Vis. Sci.* **52**(6), 3338–3344 (2011).
38. B. Pemp et al., "Effectiveness of averaging strategies to reduce variance in retinal nerve fibre layer thickness measurements using spectral-domain optical coherence tomography," *Graefes Arch. Clin. Exp. Ophthalmol.* **251**(7), 1841–1848 (2013).
39. J. U. Kang et al., "Real-time three-dimensional fourier-domain optical coherence tomography video image guided microsurgies," *J. Biomed. Opt.* **17**(8), 081403 (2012).
40. H. Wu, J. F. de Boer, and T. C. Chen, "Reproducibility of retinal nerve fiber layer thickness measurements using spectral domain optical coherence tomography," *J. Glaucoma* **20**(8), 470–476 (2011).
41. N. Serbecic et al., "Reproducibility of high-resolution optical coherence tomography measurements of the nerve fibre layer with the new Heidelberg spectralis optical coherence tomography," *Br. J. Ophthalmol.* (2010).
42. Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.* **43**(4), 874–887 (2005).
43. J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.* **20**, 37–46 (1960).
44. R. L. Brennan and D. J. Prediger, "Coefficient kappa: some uses, misuses, and alternatives," *Educ. Psychol. Meas.* **41**(3), 687–699 (1981).
45. J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics* 159–174 (1977).
46. D. Bizios et al., "Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by stratus OCT," *Acta Ophthalmol.* **88**(1), 44–52 (2010).
47. R. Fan, P. Chen, and C. Lin, "Working set selection using second order information for training support vector machines," *J. Mach. Learn. Res.* **6**, 1889–1918 (2005).

48. F. Imbault and K. Lebart, "A stochastic optimization approach for parameter tuning of support vector machines," in *Proc. 17th Int. Conf. on Pattern Recognition, (ICPR)*, Vol. 4, pp. 597–600, IEEE, Cambridge, UK (2004).
49. M. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmos. Environ.* **32**(14–15), 2627–2636 (1998).
50. R. N. Weinreb and P. T. Khaw, "Primary open-angle glaucoma," *The Lancet* **363**(9422), 1711–1720 (2004).

**Akram Belghith** received his PhD degree in computer science from University of Strasbourg. He is an assistant scientist at the Hamilton Glaucoma Center, University of California, San Diego, California, USA. His current research interests include machine learning, image processing, pattern recognition, and data analysis.

**Christopher Bowd** received his PhD degree from Washington State University, Pullman, Washington, USA. Currently, he is a research scientist at the Hamilton Glaucoma Center, University of California, San Diego, California, USA. His current work involves early detection and monitoring of glaucoma with structural imaging of the optic nerve, visual function, and electrophysiological testing using standard and machine learning classifier-based analyses.

**Felipe A. Medeiros** received the graduate degree and residency from the University of Sao Paulo. He is a professor of ophthalmology and the medical director of the Hamilton Glaucoma Center, University of

California San Diego, California, USA. He is also the director of vision function research at the same institution.

**Madhusudhanan Balasubramanian** received his PhD degree from Louisiana State University, USA. Currently, he is an assistant professor in the Department of Electrical and Computer Engineering at the University of Memphis, Tennessee, USA. His research interests include the intersection of computational science and engineering, biosolid mechanics and biofluid dynamics with emphasis in studying ocular structures, and dynamics, and the mechanism of vision loss in glaucoma.

**Robert N. Weinreb** received his electrical engineering degree from the Massachusetts Institute of Technology, USA, and his MD degree from Harvard Medical School, Massachusetts. He is a clinician, a surgeon, an educator, and a scientist. He is the distinguished professor of ophthalmology and the chairman of the Department of Ophthalmology at the University of California, San Diego, California, USA.

**Linda M. Zangwill** received her MS degree from the Harvard School of Public Health and her PhD degree from Ben-Gurion University of the Negev. She is a professor of ophthalmology at the University of California, San Diego, California, USA. Her research interests include studying the relationship between structural and functional changes in the aging and glaucoma eye, and developing computational techniques to improve glaucomatous change detection.