Noise stochastic corrected maximum \emph{a posteriori} estimator for birefringence imaging using polarization-sensitive optical coherence tomography

DEEPA KASARAGOD, SHUICHI MAKITA, YOUNG-JOO HONG, AND YOSHIKI YASUNO∗

Computational Optics Group, University of Tsukuba, Tsukuba, Japan
∗yasuno@optlab2.bk.tsukuba.ac.jp

Abstract: This paper presents a noise-stochastic corrected maximum \emph{a posteriori} estimator for birefringence imaging using Jones matrix optical coherence tomography. The estimator described in this paper is based on the relationship between probability distribution functions of the measured birefringence and the effective signal to noise ratio (ESNR) as well as the true birefringence and the true ESNR. The Monte Carlo method is used to numerically describe this relationship and adaptive 2D kernel density estimation provides the likelihood for \emph{a posteriori} estimation of the true birefringence. Improved estimation is shown for the new estimator with stochastic model of ESNR in comparison to the old estimator, both based on the Jones matrix noise model. A comparison with the mean estimator is also done. Numerical simulation validates the superiority of the new estimator. The superior performance of the new estimator was also shown by \emph{in vivo} measurement of optic nerve head.

© 2017 Optical Society of America

OCIS codes: (170.4500) Optical coherence tomography; (110.5405) Polarimetric imaging; (170.4460) Ophthalmic optics and device; (170.4470) Ophthalmology; (170.3890) Medical optics instrumentation.

References and links
1. M. R. Hee, D. Huang, E. A. Swanson, and J. G. Fujimoto, “Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging,” J. Opt. Soc. Am. B 9, 903–908 (1992).
2. J. F. De Boer and T. E. Milner, “Review of polarization sensitive optical coherence tomography and Stokes vector determination,” J. Biomed. Opt. 7, 359–371 (2002).
3. M. Pircher, C. K. Hitzenberger, and U. Schmidt-Erfurth, “Polarization sensitive optical coherence tomography in the human eye,” Progress in Retinal and Eye Research 30, 431–451 (2011).
4. S. K. Nadkarni, M. C. Pierce, B. H. Park, J. F. de Boer, P. Whittaker, B. E. Bouna, J. E. Bressner, E. Halpern, S. L. Hoover, and G. J. Tearney, “Measurement of Collagen and Smooth Muscle Cell Content in Atherosclerotic Plaques Using Polarization-Sensitive Optical Coherence Tomography,” J. Am. College Cardiol. 49, 1474–1481 (2007).
5. M. Villiger, E. Z. Zhang, S. K. Nadkarni, W.-Y. Oh, B. J. Vakoc, and B. E. Bouna, “Spectral binning for mitigation of polarization mode dispersion artifacts in catheter-based optical frequency domain imaging,” Opt. Express 21, 16353 (2013).
6. V.-H. Le, S. Lee, B. Kim, Y. Yoon, C. J. Yoon, W. K. Chung, and K. H. Kim, “Correlation between polarization sensitive optical coherence tomography and second harmonic generation microscopy in skin,” Biomed. Opt. Express 6, 2542–2551 (2015).
7. S. Sakai, M. Yamanari, Y. Lim, N. Nakagawa, and Y. Yasuno, “In vivo evaluation of human skin anisotropy by polarization-sensitive optical coherence tomography,” Biomed. Opt. Express 2, 2623–2631 (2011).
8. K. Wiesauer, M. Pircher, E. Goetzinger, C. K. Hitzenberger, R. Engelke, G. Ahrens, G. Gruetzner, and D. Stifter, “Transversal ultrahigh-resolution polarization sensitive optical coherence tomography for strain mapping in materials,” Opt. Express 14, 5945–5953 (2006).
9. M. Yamani, M. Miura, S. Makita, T. Yatagai, and Y. Yasuno, “Phase retardation measurement of retinal nerve fiber layer by polarization-sensitive spectral-domain optical coherence tomography and scanning laser polarimetry,” J. Biomed. Opt. 13, 014013 (2008).
10. B. Cense, Q. Wang, S. Lee, L. Zhao, A. E. Elsner, C. K. Hitzenberger, and D. T. Miller, “Henle fiber layer phase retardation measured with polarization-sensitive optical coherence tomography,” Biomed. Opt. Express 4, 2296 (2013).
11. S. Fukuda, M. Yamani, Y. Lim, S. Hoshi, S. Beheregayr, T. Oshika, and Y. Yasuno, “Keratoconus diagnosis using anterior segment polarization-sensitive optical coherence tomography,” Investigative Ophthalmol. Vis. Sci.
1. Introduction

Polarization-sensitive optical coherence tomography (PS-OCT) [1, 2] is a functional extension of OCT that provides enhanced tissue contrast through the polarization properties of certain
fibrous and structurally oriented tissues, such as nerve fibers, muscle fibers, and collagen fibers. There have been widespread applications of OCT in ophthalmology [3], cardiology [4, 5], dermatology [6, 7] and also in the non-medical fields [8].

Phase retardation is one of common properties measured by PS-OCT. Conventionally, phase retardation was obtained cumulatively along the depth and is called as cumulative phase retardation (CPR). CPR imaging has been applied for the qualitative and quantitative study of ocular tissues, including the retinal nerve fiber layer [9], Henle’s fiber layer [10], cornea [11], and filtration bleb [12].

However, CPR imaging yields correct birefringence information only for certain types of tissues that demonstrate uniformity of the optic axis along their depth. This assumption is easily violated in many types of tissues that have a complicated structure across their depth [13, 14]. This shortcoming of CPR imaging has been observed over the years and some approaches have been proposed to resolve it. This includes the slope-based local phase retardation (LPR) measurement [15] and Stokes parameter-based quaternion approach [16], differential Muller matrix method [17], in addition to local Jones matrix methods [18]. In LPR analysis, optic axis uniformity is required over only a small depth separation, for example, approximately 30 to 50 μm.

Recently, local analysis based on the Jones matrix has been applied to posterior eye segment imaging [19], anterior eye segment imaging [20,21], and coronary artery imaging [5]. Sugiyama et al. provided a detailed theoretical description of the failure of CPR imaging and utility of LPR imaging, especially for the posterior eye [19].

Although LPR imaging is a progressive step toward the widespread utility of PS-OCT, there are some other factors that hinder the true quantitative evaluation polarization property of tissue using PS-OCT. This includes the entangled effect of signal intensity and birefringence to the measured phase retardation values, that is, the measured phase retardation is significantly positively or negatively biased if the signal-to-noise ratio (SNR) of OCT is low.

With recent advances in OCT technology, faster and higher data rates have enabled the acquisition of multiple B-scans at the same location. Multiple measurements allow for SNR improvement through a conventional intensity averaging technique [22] also in Jones matrix OCT (JM-OCT) data processing. However, the averaging of phase retardation does not improve its accuracy. This is because the distribution of the measured phase retardation (and birefringence) under noise is skewed, and the skewness increases as the SNR of the OCT signal becomes lower [18, 23]. Hence, the mean estimator (averaging) fails to provide a true estimation of the true birefringence.

This nonlinear dependency between phase retardation and the SNR has been modeled for a JM-OCT system, and estimators for the true birefringence have been designed. Duan et al. described a CPR-based empirical method to remove the effect of bias at a low SNR [23]. A more theoretically relevant approach based on a combination of coherent averaging and maximum a posteriori (MAP) estimation has also been presented by the authors [20]. In this method, the relationship between the measured and true birefringence was established using Monte Carlo (MC) simulation for a set effective SNR (ESNR). This was used to solve the inverse problem to estimate the true birefringence from multiple measured birefringence and SNR values by using a Bayesian statistical approach. However, in this method, the stochastic nature of the SNR was not taken into account.

In this paper, a noise stochastic corrected birefringence MAP estimator is implemented based on a noise model of JM-OCT. This method addresses the shortcomings of our previous MAP estimator [20], and presents an improved estimation of birefringence.

The superior performance of the new estimator when compared with previous estimators and the mean estimator is demonstrated using numerical simulation and in vivo experiments.
2. Theory

A brief theory of JM-OCT for LPR imaging is described in this section. The analysis that provides birefringence tomography is detailed together with a noise model on which the MAP estimator details are theoretically based. The notation used in this section is summarized in Appendix A.

2.1. Jones matrix OCT and birefringence

JM-OCT provides multifunctional contrasts, including scattering, angiography [24], polarization uniformity [25], and birefringence [18], over a single measurement. However, in this paper, we focus on birefringence contrast. Birefringence contrast is obtained through LPR imaging over a small depth domain. Because the details of how CPR and LPR are obtained from the JM-OCT signals was described in our previous publications [18, 19, 26], here we provide only a brief summary.

JM-OCT provides four OCT signals that correspond to the combination of two input polarization states and two polarization diversity detection channels. These four OCT signals correspond to four entries in a Jones matrix \( J_s(z) \), which is mathematically similar to the cumulative round-trip Jones matrix of a sample, where \( z \) is the depth position. For LPR imaging, a similar matrix to the round-trip Jones matrix of the sample within a small depth domain is obtained first as

\[
J_l(z; Z_d) \equiv J_s(z_2) J_s(z_1)^{-1},
\]

where \( J_l(z; Z_d) \) represents a Jones matrix that possesses mathematical similarity to the local Jones matrix for the small depth domain. \( Z_d \) is the thickness of the depth domain, \( z_1 = z - Z_d/2 \), and \( z_2 = z + Z_d/2 \). Although it is omitted for simplicity, \( J_l(z; Z_d) \) is a function of the lateral positions.

The LPR value is obtained as the phase difference of the eigenvalues of \( J_l(z; Z_d) \). The eigenvalues \( \lambda_1, \lambda_2 \) and phase retardation \( \delta(z) \) can be obtained as described in Section 35.2.3 in Ref. [27].

Birefringence has a linear relationship with LPR and is given by

\[
b(z) = \frac{\delta(z)}{2k_0 Z_d},
\]

where \( k_0 \) is the wavenumber that corresponds to the center wavelength of OCT. Note that this relationship requires following two conditions in order to be accurate [19]. The first condition is the local linearity (LL-) condition; the eigen-polarizations of the small depth domain must be linear polarizations. The second condition is the uniform orientation (UO-) condition; the optic axis orientations must be uniform within the depth domain. The detailed theory of local birefringence analysis for JM-OCT is provided in [18, 19].

2.2. Noise model of JM-OCT

For a JM-OCT measurement under stochastic noise, Eq. (1) is modified as

\[
J_l(z; Z_d) = [J_s(z_2) + N(z_2)] [J_s(z_1) + N(z_1)]^{-1},
\]

where \( N(z_{1,2}) \) is the \( 2 \times 2 \) noise matrix in which each entry consists of random complex Gaussian noise with a variance of \( \sigma_n^2(z_{1,2}) \), that is, each real and imaginary component has a variance of \( \sigma_n^2(z_{1,2})/2 \).

The ESNR is introduced to represent the overall SNR of the local Jones matrix that is derived from two cumulative measured Jones matrices that consist of eight OCT images as [18]

\[
ESNR = \left[ \frac{1}{4} \sum_{j=1}^{2} \sum_{i=1}^{2} \left( \frac{I_{h,j}^{(f)}(z_i) + I_{v,j}^{(f)}(z_i)}{\sigma_n^2(z_i)} \right)^{-1} \right]^{-1},
\]

where \( I_{h,j}^{(f)}(z_i) \) and \( I_{v,j}^{(f)}(z_i) \) represent the horizontal and vertical images, respectively, and \( \sigma_n^2(z_i) \) is the variance of the noise.
where $i = 1, 2$ and $j = 1, 2$ are the indices of incident polarization and depth, respectively. $I_{h}^{(j)}(z)$ and $I_{v}^{(j)}(z)$ are the signal intensities of the horizontal and vertical polarization detection channels, respectively.

The LPR obtained from this noisy local Jones matrix is also affected by noise. Because the derivation of phase retardation from the Jones matrix is nonlinear, the distribution of the measured phase retardation is not Gaussian or symmetric. It is known that both the mean and mode of the measured phase retardation do not approach the true phase retardation [18, 23]. This implies that averaging the multiple measured phase retardation does not provide the true phase retardation even with a large number of measurements. This asymmetric distribution of phase retardation is more profound with a low ESNR and with the true phase retardation close to the edge of the measurable retardation range $[0, \pi]$. For instance, if the true phase retardation is zero, the mean of the measured phase retardation is significantly positively biased, even with an ESNR as high as 25 dB [20, 23].

A derivation of the analytical relationship between the measured and true phase retardation distributions is complex and has not been successfully achieved yet. However, Makita et al. have shown that the distribution of the measured phase retardation, and hence the measured birefringence, is parameterized only by three independent factors: the ESNR, true phase retardation (or true birefringence), and mutual angle between two incident polarization states [18]. Because the last factor is a constant in practice, the distribution of the measured birefringence under stochastic noise is characterized only by two parameters: the true birefringence $\beta$ and true ESNR.

It should be noted that the ESNR cannot be accurately computed from the measured OCT signals. This is because $I_{h,v}^{(j)}(z)$ in Eq. (4) is the intensity of the true OCT signal, whereas the measured signal is the summation of the true signal and a realization of noise. Thus, we need to distinguish two quantities: the true ESNR $\gamma$ and measured ESNR $g$. The measured ESNR is the ESNR obtained using the measured signal intensity as $I_{h,v}^{(j)}(z)$. In the next sections, the relationship between the distributions of these two ESNRs is discussed in addition to how it affects the distributions of the measured and true birefringence.

### 2.3. Probability distribution functions and likelihood functions

The MAP birefringence estimator proposed in this paper is based on the relationship between the PDFs of the measured birefringence ($b$), true birefringence ($\beta$), measured ESNR ($g$), and true ESNR ($\gamma$).

Our first concern is the PDF of the measured birefringence when specific values of the true birefringence and true ESNR are given: $p(b|\beta, \gamma)$. The second concern is the PDF of the measured ESNR when specific values of the true birefringence and true ESNR are given: $p(g|\beta, \gamma)$. These two PDFs represent the distributions of the measured values.

As our goal is to estimate the true birefringence from the measured values, the PDF of our real interest is the PDF of the true birefringence under specific values of the measured birefringence and measured ESNR. This PDF is denoted as $p(\beta|b, g)$ and represented as the posterior distribution because it is specified after the measurement values have been obtained.

According to Bayes’ rule [28], the posterior distribution is given as

$$p(\beta|b, g) \propto p(b, g|\beta)\pi(\beta), \quad (5)$$

where $p(b, g|\beta)$ is the joint probability distribution of $b$ and $g$ under a given true birefringence $\beta$. $\pi(\beta)$ is the distribution of the true birefringence. $\pi(\beta)$ is known as a prior distribution because it is the subjectively expected distribution before the measurement is performed. It is frequently assumed to be a uniform distribution, which is referred to as a non-informative prior. It should be noted that $p(\beta|b, g)$ is normalized to make its integration along $\beta$ unity. Because of this normalization condition, Eq. (5) sufficiently specifies $p(\beta|b, g)$. 

This equation implies that if we can compute \( p(b, g|\beta) \) and assume a specific \( \pi(\beta) \), such as a uniform prior, we can obtain the PDF of the true birefringence. Additionally, this PDF can provide a maximum likelihood estimation of the true birefringence, as discussed in Section 2.4.

To compute its shape, \( p(b, g|\beta) \) is further specified as

\[
p(b, g|\beta) = \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma)p(\gamma|\beta) \, d\gamma,
\]

where \( p(b, g|\beta, \gamma) \) is a joint conditional PDF of the measured birefringence and measured ESNR under a specific true birefringence and true ESNR. \( p(\gamma|\beta) \) is the conditional PDF of the true ESNR under a specific true birefringence. By assuming \( \gamma \) and \( \beta \) are mutually independent, Eq. (6) is simplified as

\[
p(b, g|\beta) = \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma)p(\gamma) \, d\gamma,
\]

where \( p(\gamma) \) is the PDF of the true ESNR, which is further discussed later in this section.

\( p(b, g|\beta, \gamma) \) can be numerically obtained using the MC method, as described in Section 3.1. Thus, by substituting the numerically obtained \( p(b, g|\beta, \gamma) \) into Eq. (7) and then into Eq. (5), \( p(\beta|b, g) \) is obtained.

Then, a likelihood function \( f(\beta; b, g) \) is defined as

\[
\begin{align*}
  f(\beta; b, g) & \equiv p(b, g|\beta) \\
  & \equiv \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma)p(\gamma) \, d\gamma.
\end{align*}
\]

This likelihood function is used in the estimation process described in the following section.

\( p(\gamma) \) must be defined to obtain \( p(b, g|\beta) \) and hence \( f(\beta; b, g) \). We recall that our purpose is to estimate \( \beta \) from the measured values of \( b \) and \( g \), thus it is rational to use \( p(\gamma|g) \) as \( p(\gamma) \), where \( p(\gamma|g) \) is a PDF of the true ESNR under a specific measured ESNR. By applying Bayes’ rule, \( p(\gamma|g) \) is specified as

\[
p(\gamma|g) \propto p(g|\gamma)\pi(\gamma),
\]

where \( \pi(\gamma) \) is a prior distribution of the true ESNR, that is, the subjectively expected distribution of ESNR before the measurement was performed. By manipulating Eq. (9), as described in Appendix B, it can be shown that

\[
p(\gamma|g) \propto \pi(\gamma) \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma) \, d\beta \right] \, d\gamma \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma) \, d\beta,
\]

where \( \pi(\beta) \) is the prior distribution of the true birefringence.

By substituting \( p(\gamma|g) \) into \( p(\gamma) \) in Eq. (8), the likelihood \( f(\beta; b, g) \) is finally defined as

\[
f(\beta; b, g) \propto \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma) \left[ \pi(\gamma) \int_{-\infty}^{+\infty} p(b, g|\beta, \gamma) \, d\beta \right] \, d\gamma.
\]

This equation can be interpreted as the likelihood obtained using a two-step inferring process. The first step is to infer the PDF of the true ESNR under a specific measured ESNR, which
is represented by the part within the square brackets. The second step is represented by the outermost integration, which infers the posterior distribution of the true birefringence using the inferred PDF of the true ESNR.

A method to numerically compute this likelihood function is described in Section 3.1 and a true birefringence estimation method using this likelihood function is described in Section 2.4.

2.4. Maximum a posteriori estimation of the true birefringence

After the JM-OCT measurements have been performed, a set of specific values of the measured birefringence and ESNR are obtained. Using these measured values, the numerically defined likelihood function and a properly predefined prior distribution \( \pi(\beta) \), the posterior distribution \( p(\beta|b,g) \) is obtained. We first describe how the posterior distribution is obtained after single and multiple measurements. Additionally, we describe a method to obtain the MAP estimation from the posterior distribution.

Let \( b_1 \) and \( g_1 \) be the measured birefringence and measured ESNR from a single JM-OCT measurement. Using these values, the likelihood function is specified as \( f(\beta; b_1, g_1) \).

By substituting this likelihood into \( p(\beta|b,g) \) in Eq. (5) and assuming a uniform prior of \( \pi(\beta) \), the posterior distribution after a single measurement is obtained as

\[
p_1(\beta) \equiv p(\beta|b = b_1, g = g_1) \propto f(\beta; b_1, g_1),
\]

where subscript 1 indicates that \( p_1(\beta) \) is the posterior distribution of \( \beta \) after the first measurement.

Assume that a set of measured values \( b_2 \) and \( g_2 \) were obtained using the second measurement. By substituting a likelihood \( f(\beta; b_2, g_2) \) into \( p(\beta|b,g) \) in Eq. (5) and using \( p_1(\beta) \) as a prior distribution, the posterior distribution after the second measurement is obtained as

\[
p_2(\beta) \propto f(\beta; b_2, g_2)p_1(\beta) = f(\beta; b_2, g_2)f(\beta; b_1, g_1).
\]

This equation is generalized for the case after \( N \) measurements by iterating these processes as

\[
p_N(\beta) \propto \prod_{i=1}^{N} f(\beta; b_i, g_i).
\]

The MAP estimation of the true birefringence is finally obtained as

\[
\hat{\beta} = \arg \max_{\beta} p_N(\beta) = \arg \max_{\beta} \prod_{i=1}^{N} f(\beta; b_i, g_i).
\]

Recall that this equation is the main equation of the MAP estimation. The likelihood \( f(\beta, b, g) \) is numerically computed, as described in Section 3.1.

2.5. Reliability of the MAP estimation

The estimation process provides not only the estimation value but also the posterior distribution of the true birefringence. The reliability of the estimation is defined using the posterior distribution as

\[
\mathcal{L}(\hat{\beta}; \Delta \beta) = \int_{\hat{\beta} - \Delta \beta/2}^{\hat{\beta} + \Delta \beta/2} p_N(\beta) \, d\beta.
\]

This is the probability that the true birefringence is within a small range of birefringence (\( \delta \beta \)) centered on the MAP estimation \( \hat{\beta} \).

This reliability is used to form a pseudo-color birefringence OCT image, as described in Section 3.4.
3. Implementation

3.1. Numerical generation of the likelihood function

As described in Sections 2.3 and 2.4, it is crucial to obtain the likelihood function for the MAP estimation. The first step to obtain this likelihood function is an MC simulation of JM-OCT based on the stochastic noise model described in Section 2.2. The set parameters of the MC simulation is the true birefringence $\beta$ and true ESNR $\gamma$, whereas each trial of MC provides a measured birefringence value $b$ and measured ESNR value $g$. The second step is to obtain the PDF $p(b, g|\beta, \gamma)$ from the MC simulated dataset. The final step is to obtain the likelihood function $f(\beta; b, g)$ from the PDF.

For the first step, MC simulation with the set parameters $\beta$ and $\gamma$ is performed. Equally spaced 1,024 set values of $\beta$ are used, where the set values are distributed in a $[0, \pi]$ range of the corresponding phase retardation. $\gamma$ is set over 0 to 40 dB with a 1-dB interval. For each true birefringence, a Jones matrix with the set true phase retardation, which is defined by the true birefringence, is generated. Then, independent and identically distributed complex zero-mean Gaussian noise is added to each entry in the Jones matrix, where each realization of noise is computed using a Gauss-distributed random number generator with a set variance corresponding to the set true ESNR. Each trial of MC corresponds to a single point in the four-dimensional (4-D) space of $(\beta, \gamma, b, g)$. A total of 65,536 trials of MC are performed for each set of true birefringence and true ESNR.

In the second step, $p(b, g|\beta, \gamma)$ is obtained by applying kernel density estimation (KDE) to the MC results. The KDE is a two-dimensional (2-D) KDE with adaptive bandwidth based on the diffusion process [29].

The measured ESNRs are ranged over -23 to 40 dB with a 1-dB sampling interval.

The numerical representation of $p(b, g|\beta, \gamma)$ is a 4-D array and each array entry possesses the probability value.

The final step is to obtain the likelihood $f(\beta; b, g)$ by substituting the numerically obtained $p(b, g|\beta, \gamma)$ into Eq. (11). The prior distribution of the true birefringence is $\pi(\beta)$, which is assumed to be constant (non-informative). The prior distribution of the true ESNR $[\pi(\gamma)]$ is assumed to be constant in the dB scale. The integration in Eq. (11) is replaced by summation. Note that the selection of optimal prior of ESNR is an open question. Although we used non-informative prior in dB scale, a linear non-informative prior can also be considered.

For the practical implementation, the likelihood is computed before the measurement and stored in a logarithmic scale. Because of this logarithmic conversion, the product operation in Eq. (14) is numerically performed by a summation operation. This accelerates the estimation, while not altering the estimation result.

The entire flow of the likelihood generation and estimation is summarized in Fig. 1.

3.2. Jones matrix OCT system

A JM-OCT system was used for the validation of the noise stochastic corrected MAP estimator. This JM-OCT used a MEMS-based wavelength sweeping laser at 1.06-µm wavelength (Axsun Technology Inc., MA) with a sweep rate of 100 kHz. Two incident polarization states were multiplexed by the passive mutual delay between them [30,31]. The depth resolution was measured as 6.2 µm in tissue. Interference signals were detected with a polarization diversity detection module. Because of this diversity detection and the incident polarization multiplexing, the system provided four OCT images for a single scan, and it formed $J_s(z)$ in Eq. (1). More details on the system are provided in [19,26].
3.3. Protocol for in vivo validation

A total of 128 B-scans (frames) were obtained at the same location on the sample, and each A-line consists of 1,024 A-lines. In order to minimize the effect of sample motion in the subsequent evaluation process, sixty-four highly structurally correlated frames were selected as a dataset. Another measurement was performed to obtain depth-dependent noise as the probe beam was blocked.

The LPR was computed for a local depth size of 6 pixels, which corresponds to a depth separation of 24 μm. The birefringence was computed from the LPR and depth separation using Eq. (2).

The dataset was then processed using three estimators and two configurations of coherent averaging of Jones matrices (adaptive Jones averaging (AJA), Section 2.1.2 in Ref. [32]). Thus, the data was processed using six approaches.

The first estimator was a mean estimator, which was an average of the birefringence values. The second estimator was the previous version of the MAP birefringence estimator [20], thus referred to as the “old MAP estimator.” The old MAP estimator is similar to the estimator presented in this paper, but it assumes that the measured ESNR equals the true ESNR; that is, it does not take into account the stochastic nature of the ESNR. The third estimator was the MAP birefringence estimator described in previous sections, which is referred to as the “new MAP estimator.”

These three estimators were applied with two AJA approaches, that is, with or without AJA. For the estimation without AJA, the first 16 frames of the dataset were used. The estimation was performed over 16 pixels at the same location in all 16 frames. For the estimation with AJA, the first four frames of the dataset were used. AJA combined the four frames and provided a single frame. The estimators were then applied with a 2 × 2-pixel spatial kernel.

Here we used a non-spatial kernel for the estimation without AJA, because the new MAP estimator is primarily designed for non-spatial kernel. On the other hand, the estimation with AJA is performed with a combination of a non-spatial kernel for AJA and a spatial kernel for estimators. This kernel configuration is according to our previous study [19]. Since, the intention
Fig. 2. (a) Estimated phase retardation from 1,024 sets of the numerically simulated measured phase retardation and ESNR. The red, blue, and black curves represent the estimation results of the new MAP, old MAP, and mean estimator. The dashed horizontal lines represent 10 true phase retardation values. (b) Root mean square error of the estimations. The mean errors were computed over the 10 true phase retardations. The type of estimator is indicated by the same color in (a).

of the present study is to compare the new MAP estimator with our old method, we used these two different kernel configurations. For both configurations, the total number of pixels used for a single estimation is 16.

To examine the effect of the number of frames with respect to image quality, all 64 frames of the posterior eye were used without AJA.

3.4. Pseudo-color birefringence image

For better readability of the estimated birefringence image, a pseudo-color composition was created. In this composition, pixel brightness was defined by the OCT intensity in the dB scale, where the intensity was defined as the intensity average of four Jones matrix entries. The pixel hue was defined by the MAP estimated birefringence value. The saturation was set as the sigmoid-converted estimation reliability

\[ S = 2 \text{Sigmoid}(aL) - 1, \quad (17) \]

where \( a \) is an empirically defined constant for the best subjective quality of the image and \( L \) is the reliability of the MAP estimation defined by Eq. (16). Because \( L \) is always greater than or equal to zero, the value of the sigmoid function ranges from 0.5 to 1.0. Because of the scaling factor of 2 and offset correction (-1), the range of the saturation \( S \) was from 0.0 to 1.0. In this composition image, the pixel with higher birefringence estimation reliability appeared as a clear color, whereas that with a lower reliability appeared more monochromatic.

The reliability of the new MAP estimator images were used to form pseudo-color mean estimator images to highlight the difference between the estimated birefringence values in the reliable regions in the results section.

4. Results

4.1. Numerical validation

The performance of the new MAP, old MAP, and mean estimator were numerically validated. In this validation, the phase retardation representation was used rather than birefringence because it is a more primary quantity of the birefringence measurement. A set of measured phase retardations and ESNRs were generated by the MC method based on the JM-OCT noise model. A total of 1,024 realizations of the measured phase retardation and measured ESNR were obtained
for each of the 10 true phase retardation values and for each set of true ESNR values. The true phase retardation ranges from 0 to $\pi$ and the 10 values are equally spaced in this range. The set ESNR varies from 5 to 40 dB with a 1-dB interval.

The estimated true phase retardations from the two MAP estimators and mean estimator are plotted over the set of (true) ESNR values in Fig. 2(a). The results of the new MAP estimator, old MAP estimator, and mean estimator are presented in red, blue, and black, respectively. The ten curves of each color correspond to 10 sets of true phase retardations, which are indicated by horizontal dashed lines. All estimators provide a correct estimation at a high ESNR. By contrast, as ESNR decreases, the estimations are more significantly biased from the true value. The mean estimator (black) provides the worst estimation with the largest offset, whereas the new MAP estimator (red) provides the best estimation with the least offset.

Figure 2(b) shows the root mean square errors over the 10 sets of true phase retardation at each set of true ESNR. The mean estimator (black) shows a significantly larger error than the MAP estimators. Among the MAP estimators, the new MAP estimator (red) shows a smaller error than the old MAP estimator (blue). To summarize, the new MAP estimator has the lowest offset and the lowest mean error.

4.2. In vivo measurements

For the in vivo validation of the estimators, a healthy human posterior eye was scanned using the JM-OCT system. The scan area was horizontal 6 mm around the optic nerve head. The obtained datasets were processed by the mean, new MAP, and old MAP estimators. The protocol adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of the University of Tsukuba.

The birefringence images obtained using the estimators are shown in Fig. 3. The left column shows the results without AJA and the right column shows the results with AJA. Each row contains images of the mean, old MAP, and new MAP estimators from top to bottom.

Among the images without AJA, the mean estimator [Fig. 3(a)] shows erroneously upshifted birefringence, especially in neural retinal layers. The two MAP estimators provided reasonable birefringence in most of the tissue layers. However, certain gray areas are observed for the old MAP estimator [Fig. 3(b)], which are not observed for the new MAP estimator [Fig. 3(c)]. The gray area represents a very low reliability of estimation. This low reliability is caused by a competition of likelihoods, and it occurs frequently with the old MAP estimator at a high ESNR region of in vivo measurement. The details of the competition mechanism are as follows: A high measured ESNR results in a narrow and sharp likelihood: $f$ in Eq. (14). If the likelihoods in a set of measurements were sharp and their positions in $\beta$ varied more than their width among the measurements, then the likelihoods can cancel each other out by the product operation in Eq. (14). This likelihood competition would not have occurred if the stochastic nature of the ESNR was properly accounted for in estimation theory, which is the case of the new MAP estimator and is evident in Fig. 3(c).

The results with AJA (right column) show lower birefringence values than those without AJA (left column). This is because the coherent averaging of the Jones matrix (AJA) reduces the phase retardation if the Jones matrices in the averaging kernel are not uniform. This occurs especially for in vivo measurement, where the Jones matrix varies by the sample motion. Spatially extending the averaging kernel also causes this lower shift of birefringence. This effect is dominantly seen in both MAP estimator images. Notice that the gray pixel area in Fig. 3(e) (old MAP with AJA) is noticeably smaller than Fig. 3(b) (old MAP without AJA). This is because AJA (coherent averaging) reduces the variation of the Jones matrices, and hence reduces the variations of the measured ESNRs and measured birefringence used for MAP estimation. Fewer variations of the measured values reduce the likelihood competition, thus the gray area was reduced. By contrast, the same effect erroneously reduced birefringence.
Fig. 3. Comparison of the mean, old MAP, and new MAP estimators for *in vivo* posterior eye imaging. The right and left column represent the birefringence estimation without and with adaptive Jones matrix averaging (AJA), respectively. Each row represents the mean, old MAP, and new MAP estimations, from top to bottom.

To summarize, the new MAP estimator without AJA [Fig. 3(c)] is an optimal option because of its lower likelihood competition and lower birefringence washout.

To further highlight the properties of the MAP and mean estimators, the histograms of the images in Fig. 3 is shown in Fig. 4. In the histograms, the birefringence are displayed in the form of phase retardation because it is more primary quantity than the birefringence. Low intensity pixels that have lower intensity than +10 dB above the noise floor were not included in the histograms. The subfigure labels in Fig. 4 are corresponding to those in Fig. 3. The dotted lines indicate $\frac{2\pi}{3}$ radians. It is known that the mean estimation erroneously converges into around this value as the ESNR decreases [18]. By comparing Fig. 4(a) with other histograms, it can be found that the mean estimations are up- or down-shifted towards $2\pi/3$. Between MAP estimators, more pixels show very low birefringence with AJA (peak close to 0 radian). It would be because of the washout effect. By analyzing the histograms, we conclude that the old and new MAP estimators without AJA have superior performance to the other combinations of estimators and AJA. Since the old MAP estimator provides significant amount of low reliability pixels as shown in Fig. 3, we reemphasize the original conclusion; the new MAP without AJA provides superior performance to the other estimation configurations.
Fig. 4. The histograms of Fig. 3. The horizontal axis represents phase retardation. The subfigure labels correspond to those of Fig. 3. The dotted line represents the $2\pi/3$, to which the mean estimation erroneously converges as the ESNR decreases.

Fig. 5 compares the estimations obtained with 4, 16, and 64 frames. The first and second rows represent the mean and new MAP estimations, respectively. The third and fourth rows represent line profiles along A1-A2 and B1-B2, respectively. The lines A1-A2 and B1-B2 are indicated in Fig. 5(d). The black and red curves correspond to the mean and new MAP estimators. By observing the line profiles, we found that the variations of both estimations reduce as the number of frames increases. By contrast, the estimation offsets show different properties between two estimators. The estimation offset of the mean estimator is nearly constant among the numbers of frames, whereas the estimation offset of the new MAP estimator decreases as the number of frames increases.

In Figs 3 and 5, the posterior part of retinal nerve fiber layer shows significantly higher birefringence than its anterior part. It may be an artifact caused by inclusion of tissue boundary into the depth domain of local phase retardation (birefringence) computation. In the local phase retardation imaging, this local depth domain was assumed to be homogeneous. In other words, the inclusion of tissue boundary can violate UO- and LL-conditions. The accuracy of the birefringence measurement can be improved by segmenting the nerve fiber layer and compute the local phase retardation within the nerve fiber layer.
Fig. 5. Comparison of mean (a)-(c) and new MAP (d)-(f) estimators for 4, 16, and 64 repeated measurements. The first and second rows show pseudo-color birefringence images obtained using the mean and new MAP estimators, respectively. The third and fourth rows show the line profiles of birefringence along A1-A2 (third row) and B1-B2 (fourth row) lines. The black and red curves represent the mean and new MAP estimators, respectively. The positions of the profiles are indicated in (d).

5. Discussion

5.1. Computation time

The MAP estimator involves two steps. The first step is the generation of a likelihood function. The second step consists of computing the posterior distribution using the measured values and likelihood function, and obtaining the maximum likelihood value.

The first step is computationally intense. It took approximately 2 hours for 1,721,344 MC trials with a custom-made Python program, and approximately 3.5 hours for KDE in MATLAB. Both programs were run on a Windows 10 PC with Intel Core i7-4910MQ CPU and 16 GB RAM. Although this step takes a significant time, it is required only once for each JM-OCT system design. Thus, it does not prohibit the practical application of the MAP estimator.

The second step for the image shown in Fig. 5 (1024 × 500 pixels) took approximately 31, 72, and 232 s with 4, 16, and 64 frames, respectively. The second step was implemented in Python on the same PC as the first step.
5.2. Effect of coherent averaging

While AJA upshifts the ESNR, it also downshifts birefringence, as shown in the results. This effect may be ubiquitous among several coherent averaging techniques. Villiger et al. recently described how coherent averaging results in erroneous birefringence values [33]. Coherent averaging also reduces the image resolution if it is applied with a spatial kernel, thus it results in larger speckles. Thus, it is a substantial advantage of the new MAP estimator that it does not require AJA.

5.3. Local phase retardation and birefringence

Technically, birefringence and LPR do not have a one-to-one relationship. First, LPR is a property of a reflected probe beam, whereas birefringence is a property of a tissue. Additionally, LPR is not only defined by birefringence but also by polarization dependent scattering, the variation of the optic axis in a small region, and several other factors. Thus, the true LPR is not exactly analogous to the true birefringence.

The quantity estimated by the estimators is the true LPR rather than birefringence. The measured LPR is derived from the true LPR, but it is affected by the stochastic noise of OCT. The presented MAP estimators correct the effect of stochastic noise, thus they provide estimates of the true LPR. However, this true LPR does not always represent the true birefringence. The difference between the true LPR and true birefringence should be carefully considered, especially for the interpretation of tissue images because tissue has complicated microscopic structures and the divergence between LRP and birefringence is larger than simple samples, such as wave plates.

6. Conclusion

We have presented a MAP birefringence estimator with stochastic ESNR correction. The superiority of this MAP estimator compared with a simple mean estimator and a previously demonstrated MAP estimator, which does not take into account the stochastic nature of ESNR, was shown using numerical validation and in vivo measurement. This new MAP estimator was found to provide the most accurate estimation among these three estimators, and found to be useful, especially for highly scattering tissue samples.
Appendices

A. List of notations and abbreviations

| Notation | Description |
|----------|-------------|
| $\beta$ | True birefringence. |
| $b$     | Measured birefringence. |
| $\gamma$ | True ESNR. |
| $g$     | Measured ESNR. |
| $z$     | Depth position. |
| $\delta$ | Phase retardation. |
| $k_0$   | Wavenumber corresponding to the central wavelength of OCT. |
| $J_s(z)$ | A similar matrix to cumulative round trip Jones matrix of a sample. |
| $J_l(z; Z_d)$ | A similar matrix to a Jones matrix of a local depth region centered at $z$ and with a width of $Z_d$. |
| $Z_d$   | Thickness of a small depth region in which local phase retardation is computed. |
| $I_{h,v}^{(i)}$ | OCT signal intensity corresponding $i$-th input polarization state and horizontal ($h$) or vertical ($v$) detection polarization state. |
| $N$     | Noise matrix. |
| $\sigma_n^2$ | Variance of real and imaginary parts of noise. |
| $p(\cdots)$ | Probability density function. |
| $\pi(\cdots)$ | Prior distribution. |
| $f(\cdots)$ | Likelihood function. |
| CPR     | Cumulative phase retardation. |
| LPR     | Local phase retardation. |
| ESNR    | Effective signal-to-noise ratio. |
| MAP     | Maximum a posteriori. |
| PDF     | Probability density function. |
| MC      | Monte Carlo. |

B. Derivation of Eq. (10)

Equation (10) is derived from Eq. (9) as follows

$$p(\gamma) \equiv p(\gamma|g) \propto p(g|\gamma)\pi(\gamma)$$
$$\propto \pi(\gamma) \int_{-\infty}^{\infty} p(g|\beta, \gamma)p(\beta|\gamma)d\beta$$
$$\propto \pi(\gamma) \int_{-\infty}^{\infty} p(g|\beta, \gamma)p(\beta)d\beta.$$  \hspace{1cm} (18)

We have used the fact that the true birefringence ($\beta$) and true ESNR ($\gamma$) are mutually independent to derive the third line from the second line. Using $\int_{-\infty}^{\infty} p(h, g|\beta, \gamma)db = p(g|\beta, \gamma)$, this equation becomes

$$p(\gamma|g) \propto \pi(\gamma) \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{+\infty} p(h, g|\beta, \gamma)db \right] p(\beta)d\beta.$$  \hspace{1cm} (19)

For Bayesian estimation, $p(\beta)$ is the prior distribution of $\beta$, thus it is denoted as $\pi(\beta)$. Thus, Eq. (10) has been derived.
Funding
This research was supported in part by the Japan Society for the Promotion of Science (JSPS, KAKENHI 15K13371) and the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) through a contract of Regional Innovation Ecosystem Development Program.

Disclosure
DK, SM, YJH, YY: Topcon (F), Tomey Corp (F), Nidek (F), Canon (F)