

Model-based sensor-less wavefront aberration correction in optical coherence tomography

HANS R. G. W. VERSTRAETE,^{1,*†} SANDER WAHLS,^{1,†} JEROEN KALKMAN,^{2,‡} AND MICHEL VERHAEGEN^{1,‡}

¹Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 DC Delft, The Netherlands

²Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

*Corresponding author: h.r.g.w.verstraete@tudelft.nl

Received 24 September 2015; revised 4 November 2015; accepted 5 November 2015; posted 6 November 2015 (Doc. ID 250676); published 9 December 2015

Several sensor-less wavefront aberration correction methods that correct nonlinear wavefront aberrations by maximizing the optical coherence tomography (OCT) signal are tested on an OCT setup. A conventional coordinate search method is compared to two model-based optimization methods. The first model-based method takes advantage of the well-known optimization algorithm (NEWUOA) and utilizes a quadratic model. The second model-based method (DONE) is new and utilizes a random multidimensional Fourier-basis expansion. The model-based algorithms achieve lower wavefront errors with up to ten times fewer measurements. Furthermore, the newly proposed DONE method outperforms the NEWUOA method significantly. The DONE algorithm is tested on OCT images and shows a significantly improved image quality. © 2015 Optical Society of America

OCIS codes: (110.4500) Optical coherence tomography; (110.1080) Active or adaptive optics; (150.1135) Algorithms.

<http://dx.doi.org/10.1364/OL.40.005722>

Non-invasive three-dimensional imaging of the retina is one of the main applications of optical coherence tomography (OCT) [1]. The lateral resolution of the OCT system can be improved by increasing the pupil size. In general, this leads to increased optical wavefront aberrations that limit the resolution. Adaptive optics (AO) has been successfully used to correct these optical wavefront aberrations on large pupils (>2 mm), leading to an improved image quality. Ultra-high lateral and axial resolutions up to 3 and 2–3 μm , respectively, have been obtained by using adaptive optics in OCT [2,3]. In general, the optical aberrations in AO-OCT setups are determined by wavefront sensors such as the Shack–Hartmann (SH) wavefront sensor. Imaging systems with SH wavefront sensors suffer from several disadvantages, such as an increased cost and non-common path wavefront errors. These drawbacks can be avoided if wavefront sensor-less adaptive optics (WFSL-AO) methods are used. WFSL methods optimize image quality metrics (e.g., the strength of the OCT signal) in order to minimize the aberrations and create a sharper image.

The ideal WFSL-AO algorithm for OCT is robust with respect to noise and converges to its optimum in a small number of measurements. Here, noise includes shot noise, speckle, and variations of the sample structure in the B-scans with different lateral positions. Finite difference approximations for the explicit calculation of derivatives, such as a forward difference in the Newton–Raphson method, require extra measurements to determine the individual partial derivatives and are not robust with respect to noise. Therefore, derivative-free optimization algorithms are preferred. It has been shown that various derivative-free optimization algorithms can successfully improve the quality and signal-to-noise ratio of OCT images [4–7]. We demonstrated that the coordinate search (CS) algorithm [4] reaches the maximum OCT signal if the aberrations are not too large [7]. Additionally, simulated annealing [6] and the stochastic parallel gradient descent algorithm [8] have been successfully used for WFSL-AO in OCT and scanning laser ophthalmoscopy, respectively. However, the final obtained root mean square (RMS) wavefront error in these algorithms is very susceptible to noise, because past measurements are not exploited. It was shown that the use of models can improve the convergence rate of WFSL algorithms [9,10]. In this case, the model, which is fitted to the measurements, is used to estimate the derivatives. The final result is much less susceptible to noise, because past measurement information is used in the fitting of the model. In [9], a quadratic model was fit to a set of prior measurements of an image metric, so that the model could be used to correct the aberrations. However, the aberration correction is limited to the relatively small region where the quadratic model is a good approximation of the metric.

Recently, we developed and validated a transfer function for optical wavefront aberrations in OCT [7]. This model predicts the loss of the OCT signal caused by optical wavefront aberrations. Based on the shape of the transfer functions and simulations done using this model, we propose two advanced model-based optimization algorithms for WFSL-AO in OCT, the NEWUOA and the DONE algorithm, and compare these two algorithms with the CS algorithm. The NEWUOA algorithm [11], which is based on an adaptive quadratic model, is chosen because it is one of the most well-known and best-performing derivative-free algorithms in optimization. Our

DONE algorithm, *Data-based Online Nonlinear Extremum-seeker*, is based on a random Fourier basis. It is called an online method, because the aberrations of each measurement are chosen based on the outcome of previous measurements. We demonstrate experimentally that both algorithms outperform the CS algorithm.

The experimental results are obtained using the AO-OCT setup described in [7], except the spectrometer is a Cobra VHR (Wasatch Photonics). The three investigated wavefront correction algorithms maximize a metric based on the magnitude of the OCT signal. The OCT signal is obtained by sampling the spectrometer signal equidistantly in the k domain using interpolation, taking the absolute value of the Fourier transform of the spectrometer signal with the reference arm intensity subtracted. The time and space averaged signal from a selected depth segment of the complete OCT signal is used as the image quality metric. The metric is a real valued function, $f(\alpha)$, of the optical wavefront aberrations represented in the vector α . The aberration coefficients α_i in α are the coefficients for the normalized Zernike polynomials in micrometers based on our model [7].

The first algorithm we consider is the coordinate search algorithm from [4]. The CS algorithm successively scans in S steps through a predefined domain of each aberration α_i with a step size s . It retains the value of α_i of the scan that maximized the metric before going to the next aberration, α_{i+1} . A performance trade off exists between the number of measurements and the size of the scanning domain. The step size s has to be small enough to ensure a small final error and large enough to have fast convergence to the maximum. The coordinate search optimization method is simple and converges slowly.

The second algorithm is the NEWUOA optimization algorithm [11]. The NEWUOA algorithm is a computationally efficient, derivative-free optimization method based on a quadratic model. The m variables of the initial multidimensional quadratic model are fitted to the first m measurements of the metric based on the OCT signal. When a new measurement is taken to update the quadratic model, an old measurement is thrown away. An optimization routine minimizes the quadratic model of $-f(\alpha)$ within a bounded region, in which the quadratic model is considered accurate, leading to a new estimate of the metric's maximum. This process is iterated until some stopping criterion is fulfilled. The NEWUOA algorithm has three parameters. The first parameter, m , is the number of variables the quadratic model is based on; the default value is $(2d + 1)$. Here, d is the number of aberration coefficients that are optimized. The second parameter is the initial step size, ρ_α . This parameter should not be chosen to be too large, so that the fundamental features of the function are not skipped, and also should not be too small, such that larger variations in $f(\alpha)$ are probed. The third parameter ρ_Ω determines the final step size, which should be small. It should be smaller than the RMS wavefront error corresponding to $\frac{\lambda}{14}$, the Maréchal criterion for the diffraction limit.

The third algorithm, which we call DONE, fits a multidimensional random Fourier basis to the measurements [12]. With every new measurement taken, the function *FourierRegression*, shown below in pseudo-code with MATLAB-like notation, computes a new model $\hat{f}(\alpha)$ of the metric function $f(\alpha)$. After the model is obtained, a well-known optimization routine, *fmincon* (MATLAB R2012b) (see [13,14] for the theoretical background on *fmincon*), is run on the model.

Since we are looking for a maximum of $f(\alpha)$, we minimize $-f(\alpha)$. The initial vector of the optimization algorithm is α_{init} , which is bounded element-wise by the lower bound lb and upper bound ub . These bounds are chosen such that the aberration, which is to be corrected, is never out of the bounded region. Before the bounding, a random perturbation is added to avoid the algorithm getting stuck in an insignificant local minimum. The coefficients of the next measurement α_{i+1} are determined by adding a small, normally distributed perturbation with standard deviation σ_1 to the last found minimizer α_{min} and enforcing the bounds as before. This second perturbation is added to keep the algorithm from concentrating on a too-narrow part of the search space.

Algorithm 1. DONE Algorithm

```

1: Procedure DONE  $\alpha_0, N, lb, ub$ 
2:    $d = \text{length}(\alpha_0)$ 
3:   for  $i = 0$  to  $N - 1$  do
4:      $f_i = f(\alpha_i) + \text{measurement noise}$ 
5:      $\hat{f}(\alpha) = \text{FourierRegression}([\alpha_0 \dots \alpha_i], [f_0; \dots; f_i], d)$ 
6:      $\alpha_{\text{init}} = \max(\min(\alpha_i + \sigma_1 \text{randn}(d, 1), ub), lb)$ 
7:      $\alpha_{\text{min}} = \text{fmincon}(-f(\alpha), \alpha_{\text{init}}, lb, ub)$ 
8:      $\alpha_{i+1} = \max(\min(\alpha_{\text{min}} + \sigma_1 \text{randn}(d, 1), ub), lb)$ 
9:   return  $\alpha_{\text{min}}$ 

```

Algorithm 2. Fourier Regression

```

1: Procedure Fourier Regression  $A, F, d$ 
2:    $\omega = \sqrt{2\sigma} \text{randn}(D, d)$ 
3:    $b = 2\pi \text{rand}(D, 1)$ 
4:    $Z = \sqrt{\frac{2}{D}} \cos(\omega A + [b \dots b])$ 
5:    $w = (ZZ^T + \lambda \text{length}(F) \text{eye}(D))^{-1} (ZF)$ 
6:    $\hat{f}(\alpha) = \sqrt{\frac{2}{D}} w^T \cos(\omega \alpha + b)$ 
7:   return  $\hat{f}(\alpha)$   ▷  $\hat{f}(\alpha)$  is a function

```

In the function *FourierRegression*, Line 5 solves the least squares problem, $\|Z^T w - F\|_2^2 + \lambda \text{length}(F) \|w\|_2^2$. The constant D specifies the number of random Fourier basis functions. More basis functions will lead to a better representation of the original function, but also to an increased computational load. The parameter λ is a regularization parameter used to avoid over-fitting the model to the measurements. The parameter σ sets the standard deviation of the frequencies of the cosine bases. It should be chosen such that the higher frequencies of the unknown function $f(\alpha)$ are still captured; however, the bandwidth should not be too high such that noise in higher frequencies is also filtered. The MATLAB commands *randn* and *rand* create normally and uniformly distributed matrices, respectively. The command *eye* creates an identity matrix.

Table 1 shows the proposed parameter settings for the three algorithms. The parameters of the CS and NEWUOA algorithms are chosen such that the diffraction limit can be reached, the maximum value of the added aberrations is covered, and fundamental features of the metric function $f(\alpha)$ are not missed. In order to find suitable values for the parameters of the DONE algorithm, it was simulated using the OCT model [7] for random aberrations between $-0.45 \mu\text{m}$ and $+0.45 \mu\text{m}$.

We perform two experiments with three and seven aberrations applied to the deformable mirror (DM) shown in

Table 1. Parameter Values for the Three Wavefront Correction Algorithms (Wavefront Aberrations are Defined in μm)

	CS	NEWUOA	DONE
S	50	ρ_α	1
s	0.01	ρ_Ω	$1e-8$
		m	$2d + 1$
			σ_1
			σ
			λ
			1000
			$0.1\sqrt{\frac{3}{d}}$
			1
			0.001

Figs. 1 and 2, respectively. Each trial, we add a random combination of these wavefront aberrations with a total maximum RMS wavefront aberration of $0.45 \mu\text{m}$ to the DM. Subsequently, the metric calculated from one B-scan of a Scotch tape sample is optimized using the three algorithms. The Scotch tape sample is shown in Fig. 3. In each iteration, only one B-scan with 512 A-scans is taken. The remaining RMS error after the optimization is measured with the Shack–Hartmann wavefront sensor. In the first experiment, shown in Fig. 1, a combination of Zernike modes three to five, i.e., a defocus and two astigmatisms, is applied to the DM and corrected. In Fig. 1(a), the maximum achieved values of the OCT signal with respect to the number of iterations are averaged over 100 trials. In Fig. 1(b), the final RMS wavefront errors of the 100 trials are shown in a box plot together with a black line indicating the Maréchal criterion for the diffraction limit based on the OCT center wavelength of 850 nm . In each box, the central red line is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points not considered outliers. Outliers are plotted individually with red markers. The coordinate search algorithm converges slowly and does not have a good final RMS wavefront error after it was stopped at 1050 measurements. However, it manages to get a strong OCT signal. Before elaborating on this observation, we continue with the description of the results. The NEWUOA algorithm converges faster than CS. It obtains smaller final RMS wavefront errors than CS, even though the OCT signals are weaker. The DONE algorithm outperforms

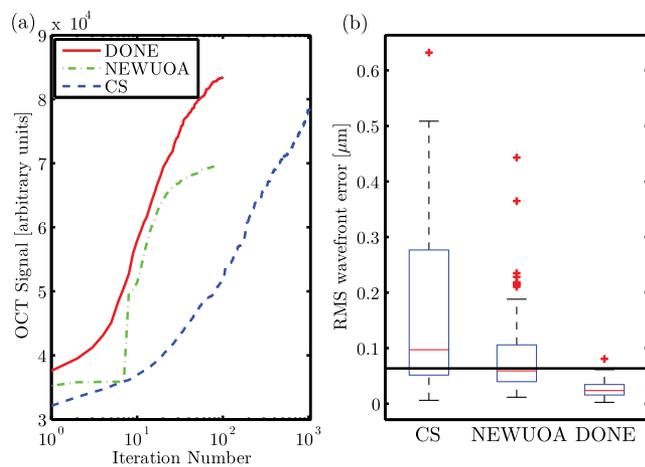


Fig. 1. AO-OCT wavefront correction for Zernike modes three to five. (a) OCT signal averaged over 100 trials versus iteration number. (b) Box plots of the 100 final RMS wavefront errors measured with the SH wavefront sensor. The black line indicates the Maréchal criterion for the diffraction limit.

both the other algorithms in final RMS error, with ten times fewer measurements than the CS algorithm. Figure 2 shows a similar experiment; however, in this case, a combination of Zernike modes three to nine are applied and corrected by the DM. The results for three and seven aberrations have similar characteristics. However, the performance of the DONE algorithm in contrast with the other algorithms in terms of the final RMS error for seven aberrations has become even better.

Since the NEWUOA algorithm has its own stopping criterion, the longest runs of NEWUOA took 87 and 205 measurements for three and seven aberrations, respectively. The DONE algorithm was set to stop after $N = 100$ and $N = 250$ measurements for three and seven aberrations, respectively. For seven aberrations, the CS and the NEWUOA algorithm often do not reach the diffraction limit, which is clear from the box plots of the final RMS wavefront errors. The DONE algorithm clearly outperforms both the CS and the NEWUOA algorithm in terms of the final RMS error. In three dimensions, DONE converges 99% of the starting aberrations below the Maréchal criterion, compared with 65% for seven dimensions.

To demonstrate the feasibility of our approach to actual OCT imaging, we image the Scotch tape sample with a defocus of $-0.54 \mu\text{m}$, a vertical astigmatism of $0.02 \mu\text{m}$, and an oblique astigmatism of $0.30 \mu\text{m}$ in Fig. 3(a). Figure 3(b) shows an OCT image where the RMS wavefront error measured by the SH wavefront sensor has been reduced to $0.03 \mu\text{m}$ after 100 iterations of the DONE algorithm. Using the same linear intensity scale, the

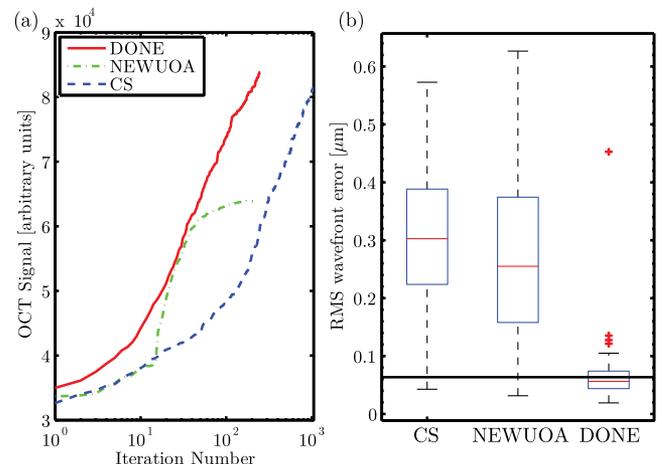


Fig. 2. Same as Fig. 1, but for wavefront aberrations consisting of Zernike modes three to nine.

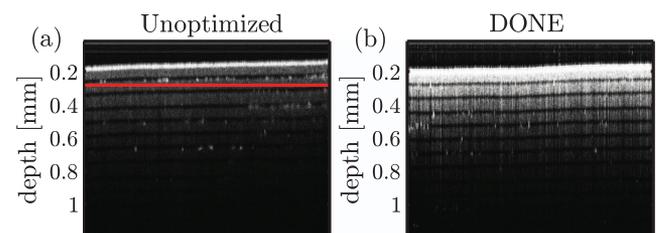


Fig. 3. OCT B-scans of Scotch tape (a) before aberration correction and (b) after 100 iterations of the DONE algorithm. The red line indicates the depth of the OCT signal optimization.

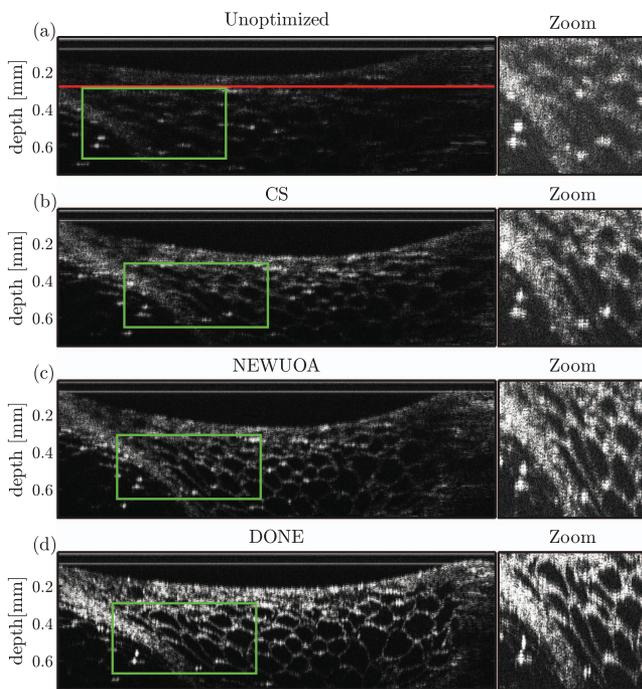


Fig. 4. OCT B-scans of a lemon slice (a) before aberration correction and (b) after 100 iterations of the CS algorithm, (c) of the NEWUOA algorithm, and (d) of the DONE algorithm. The red line indicates the depth of the OCT signal optimization, and the green rectangle indicates the zoomed-in area. An enhanced contrast scaling is applied to the zoomed-in areas.

corrected OCT image has a much stronger OCT signal. The outcomes of the other algorithms suffered from a weaker OCT signal, but had a similar structure. Thus, they are omitted.

In Fig. 4(a), we show a B-scan of a lemon slice with an added defocus of $-0.49\ \mu\text{m}$, a vertical astigmatism of $0.18\ \mu\text{m}$, and an oblique astigmatism of $0.32\ \mu\text{m}$. Figures 4(b)–4(d) show the lemon slice image after 100 iterations of the CS, NEWUOA, and DONE algorithms, respectively. A zoomed-in area for each image is added to demonstrate the clear difference in the visible structural features, in addition to the strong difference in OCT signal strength. For the large-area images, the same linear intensity scale is used; however, the zoomed-in areas all have an enhanced contrast scaling to show the features more clearly. The remaining RMS wavefront errors read from the coefficients applied to the DM for Figs. 4(b)–4(d) are 0.36 , 0.23 , and $0.06\ \mu\text{m}$, respectively.

In our experiments, we observed that even though the CS algorithm obtained a similarly strong OCT signal strength in the end, the resulting RMS error is much lower for the DONE algorithm. We attribute this phenomenon to the noise in the OCT signal measurements. The CS algorithm does not take the noise into account and can therefore yield OCT signals that are strong only due to noise. In contrast, the DONE algorithm builds its model based on all known measurements, making it less susceptible to noise as the number of measurements increases. The NEWUOA algorithm can keep track of a fixed number of measurements, but this number is inherently limited by the complexity of the quadratic model it uses. It is therefore less susceptible to noise as well, but, as can be seen in the

experimental results, the restriction to a fixed number of measurements results in a considerable increase in the final RMS wavefront error. The superior performance of the DONE algorithm, however, comes at the price of an increased computational complexity. While the computational complexity for CS and NEWUOA is independent of the number of measurements, the computational complexity of a naive implementation of DONE increases with every iteration. The CS and NEWUOA algorithms have a computational time below $1\ \text{ms}$. However, the computational time of the DONE algorithm is in the order of $60\ \text{ms}$ during the experiments. This computational time is equal to the acquisition and processing time of a B-scan in our OCT system; therefore, it is not a limiting factor. We expect that the computational complexity can be reduced below $1\ \text{ms}$ by using a recursive least squares algorithm in the FourierRegression function, optimized compiled code, and parallel processing on a graphics processing unit.

In short, the performance of three methods for WFSL-AO in OCT has been investigated experimentally. The DONE algorithm outperforms both NEWUOA and CS in terms of final RMS error, and it converges up to ten times faster than the CS algorithm. We implemented the DONE algorithm in OCT imaging and showed a significantly improved image quality. We showed that for large aberrations, the DONE algorithm succeeds in correcting a noisy signal with high accuracy.

Funding. Netherlands Enterprise Agency (RVO), Research Program for Innovation in Photonic Devices (IPD12020).

Acknowledgment. The authors would like to thank W. J. M. van Geest and C. J. Slinkman.

[†]These authors contributed equally to this paper.

[‡]These authors contributed equally to this paper.

REFERENCES

- H. David, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafito, *Science* **254**, 1178 (1991).
- Y. Zhang, J. Rha, R. Jonnal, and D. Miller, *Opt. Express* **13**, 4792 (2005).
- R. Zawadzki, S. Jones, S. Olivier, M. Zhao, B. Bower, J. Izatt, S. Choi, S. Laut, and J. Werner, *Opt. Express* **13**, 8532 (2005).
- S. Bonora and R. J. Zawadzki, *Opt. Lett.* **38**, 4801 (2013).
- Y. Jian, J. Xu, M. A. Gradowski, S. Bonora, R. J. Zawadzki, and M. V. Sarunic, *Biomed. Opt. Express* **5**, 547 (2014).
- M.-R. Nasiri-Avanaki, S. Hojjatoleslami, H. Paun, S. Tuohy, A. Meadway, G. Dobre, and A. Podoleanu, in *Proceedings of Mathematical Methods and Applied Computing*, WSEAS (2009), pp. 669–674.
- H. R. G. W. Verstraete, B. Cense, R. Bilderbeek, M. Verhaegen, and J. Kalkman, *Opt. Express* **22**, 32406 (2014).
- J. Porter, N. Sredar, H. Queener, C. Li, and H. Hofer, in *Imaging and Applied Optics* (Optical Society of America, 2013), paper OTu1A.1.
- J. Antonello, M. Verhaegen, R. Fraanje, T. van Werkhoven, H. C. Gerritsen, and C. U. Keller, *J. Opt. Soc. Am. A* **29**, 2428 (2012).
- M. Booth, *Opt. Express* **14**, 1339 (2006).
- M. Powell, in *Large-Scale Nonlinear Optimization*, G. Di Pillo and M. Roma, eds., Nonconvex Optimization and Its Applications (Springer, 2006), Vol. 83, pp. 255–297.
- A. Rahimi and B. Recht, in *Advances in Neural Information Processing Systems 20*, J. Platt, D. Koller, Y. Singer, and S. Roweis, eds. (Curran Associates, 2008), pp. 1177–1184.
- R. H. Byrd, M. E. Hribar, and J. Nocedal, *SIAM J. Optim.* **9**, 877 (1999).
- R. Waltz, J. Morales, J. Nocedal, and D. Orban, *Math. Program.* **107**, 391 (2006).