Imaging the eye fundus with real-time en-face spectral domain optical coherence tomography

Adrian Bradu1,* and Adrian Gh. Podoleanu1

1Applied Optics Group, School of Physical Sciences, University of Kent, Canterbury CT2 7NH, UK
a.bradu@kent.ac.uk

Abstract: Real-time display of processed en-face spectral domain optical coherence tomography (SD-OCT) images is important for diagnosis. However, due to many steps of data processing requirements, such as Fast Fourier transformation (FFT), data re-sampling, spectral shaping, apodization, zero padding, followed by software cut of the 3D volume acquired to produce an en-face slice, conventional high-speed SD-OCT cannot render an en-face OCT image in real time. Recently we demonstrated a Master/Slave (MS)-OCT method that is highly parallelizable, as it provides reflectivity values of points at depth within an A-scan in parallel. This allows direct production of en-face images. In addition, the MS-OCT method does not require data linearization, which further simplifies the processing. The computation in our previous paper was however time consuming. In this paper we present an optimized algorithm that can be used to provide en-face MS-OCT images much quicker. Using such an algorithm we demonstrate around 10 times faster production of sets of en-face OCT images than previously obtained as well as simultaneous real-time display of up to 4 en-face OCT images of 200 × 200 pixels² from the fovea and the optic nerve of a volunteer. We also demonstrate 3D and B-scan OCT images obtained from sets of MS-OCT C-scans, i.e. with no FFT and no intermediate step of generation of A-scans.

©2014 Optical Society of America

OCIS codes: (120.3180) Interferometry; (110.4500) Optical coherence tomography; (170.0110) Imaging systems; (200.4960) Interferometric imaging.

References and links

1. S. Taplin, A. Gh. Podoleanu, D. J. Webb, and D. A. Jackson, “Displacement sensor using channeled spectrum dispersed on a linear CCD array,” Electron. Lett. 29, 896–897 (1993).
2. M. Wojtkowski, R. Leitgeb, A. Kowalczyk, T. Bajraszewski, and A. F. Fercher, “In vivo human retinal imaging by Fourier domain optical coherence tomography,” J. Biomed. Opt. 7(3), 457–463 (2002).
3. S. R. Chinn, E. A. Swanson, and J. G. Fujimoto, “Optical coherence tomography using a frequency-tunable optical source,” Opt. Lett. 22(5), 340–342 (1997).
4. H. Lim, M. Mujat, C. Kerbage, E. C. W. Lee, Y. Chen, T. C. Chen, and J. F. de Boer, “High-speed imaging of human retina in vivo with swept-source optical coherence tomography,” Opt. Express 14(26), 12902–12908 (2006).
5. R. Leitgeb, C. K. Hitzenberger, and A. F. Fercher, “Performance of fourier domain vs. time domain optical coherence tomography,” Opt. Express 11(8), 889–894 (2003).
6. K. Zhang and J. U. Kang, “Real-time 4D signal processing and visualization using graphics processing unit on a regular nonlinear-k Fourier-domain OCT system,” Opt. Express 18(11), 11772–11784 (2010).
7. J. A. Izatt, M. R. Hee, G. M. Owen, E. A. Swanson, and J. G. Fujimoto, “Optical coherence microscopy in scattering media,” Opt. Lett. 19(8), 590–592 (1994).
8. A. G. Podoleanu, G. M. Dobre, D. J. Webb, and D. A. Jackson, “Coherence imaging by use of a Newton rings sampling function,” Opt. Lett. 21(21), 1789–1791 (1996).
9. A. Gh. Podoleanu, M. Seeger, G. M. Dobre, D. J. Webb, D. A. Jackson, and F. Fitzke, “Transverse and longitudinal images from the retina of the living eye using low coherence reflectometry,” J. Biomed. Opt. 3, 12–20 (1998).
10. A. G. Podoleanu and R. B. Rosen, “Combinations of techniques in imaging the retina with high resolution,” Prog. Retin. Eye Res. 27(4), 464–499 (2008).
11. R. B. Rosen, M. Mathaway, J. Rogers, J. Pedro, P. Garcia, P. Laisseau, G. M. Dobre, and A. G. Podoleanu, “Multidimensional en-face OCT imaging of the retina,” Opt. Express 17(5), 4112–4133 (2009).
12. M. Pincher, B. Baumann, E. Götzinger, and C. K. Hitzenberger, “Imaging the human retina and cone mosaic in vivo with PS-OCT,” Proc. SPIE 6429, 64290T (2007).
13. K. Wiesauer, M. Pincher, E. Götzinger, S. Bauer, R. Engelke, G. Ahrens, G. Grützner, C. Hitzenberger, and D. Stifter, “En-face scanning optical coherence tomography with ultra-high resolution for material investigation,” Opt. Express 13(3), 1015–1024 (2005).
14. H. Liang, M. G. Cid, R. Cucu, G. Dobre, A. Podoleanu, J. Pedro, and D. Saunders, “En-face optical coherence tomography - a novel application of non-invasive imaging to art conservation,” Opt. Express 13(16), 6133–6144 (2005).
15. D. C. Adler, C. Zhou, T.-H. Tsai, J. Schmitt, Q. Huang, H. Mashimo, and J. G. Fujimoto, “Three-dimensional endomicroscopy of the human colon using optical coherence tomography,” Opt. Express 17(2), 784–796 (2009).
16. S. Jiao, C. Wu, R. W. Knighton, G. Gregori, and C. A. Puliafito, “Registration of high-density cross sectional images to the fundus image in spectral-domain ophthalmic optical coherence tomography,” Opt. Express 14(8), 3368–3376 (2006).
17. B. Potsaid, I. Gorczynska, V. J. Srinivasan, Y. Chen, J. Jiang, A. Cable, and J. G. Fujimoto, “Ultrahigh speed spectral / Fourier domain OCT ophthalmic imaging at 70,000 to 312,500 axial scans per second,” Opt. Express 16(19), 15149–15169 (2008).
18. W. Wieser, B. R. Biedermann, T. Klein, C. M. Eigenwillig, and R. Huber, “Multi-megahertz OCT: High quality 3D imaging at 20 million A-scans and 4.5 GVoxels per second,” Opt. Express 18(14), 14685–14704 (2010).
19. S. Van der Jeught, A. Bradu, and A. G. Podoleanu, “Real-time resampling in Fourier domain optical coherence tomography using a graphics processing unit,” J. Biomed. Opt. 15(3), 035011 (2010).
20. T. E. Ustun, N. V. Iftimia, V. J. Srinivasan, Y. Chen, J. Jiang, A. Cable, and J. G. Fujimoto, “Ultrahigh-speed optical coherence tomography using a field programmable gate array,” Rev. Sci. Instrum. 79(11), 114301 (2008).
21. S. Jiao, R. Knighton, X. Huang, G. Gregori, and C. Puliafito, “Simultaneous acquisition of sectional and fundus ophthalmic images with spectral-domain optical coherence tomography,” Opt. Express 13(2), 444–452 (2005).
22. V. J. Srinivasan, D. C. Adler, Y. Chen, I. Gorczynska, R. Huber, J. S. Duker, J. S. Schuman, and J. G. Fujimoto, “Ultra-hight-speed optical coherence tomography for three-dimensional and en face imaging of the retina and optic nerve head,” Invest. Ophthalmol. Vis. Sci. 49(11), 5103–5110 (2008).
23. https://www.sog-sso.ch/fileadmin/SOG-Dokumente/Veranstaltungen/en-face-OCT-Dec-14-2013.pdf
24. Clinical en-face OCT atlas, B. Lambruso, D. Huang, A. Romano, M. Rispoli, and G. Coscas Ed., Jaypee Brothers, (2012).
25. B. R. Biedermann, W. Wieser, C. M. Eigenwillig, P. GJose, D. C. Adler, V. J. Srinivasan, J. G. Fujimoto, and R. Huber, “Real-time en face Fourier-domain optical coherence tomography with direct hardware frequency demodulation,” Opt. Lett. 33(21), 2556–2558 (2008).
26. Y. Jian, K. Wong, and M. V. Sarunic, “Graphics processing unit accelerated optical coherence tomography processing at megahertz axial scan rate and high resolution video rate volumetric rendering,” J. Biomed. Opt. 18(2), 026002 (2013).
27. J. Probst, D. Hillmann, E. Lankenau, C. Winter, S. Oelckers, P. Koch, and G. Hüttermann, “Optical coherence tomography with online visualization of more than seven rendered volumes per second,” J. Biomed. Opt. 15(2), 026014 (2010).
28. K. Zhang and J. U. Kang, “Graphics processing unit accelerated non-uniform fast Fourier transform for ultra-high-speed, real-time Fourier-domain OCT,” Opt. Express 18(22), 23472–23487 (2010).
29. K. Zhang and J. U. Kang, “Real-time intraoperative 4D full-range FD-OCT based on the dual graphics processing units architecture for microsurgery guidance,” Biomed. Opt. Express 2(4), 764–770 (2011).
30. J. U. Kang, Y. Huang, K. Zhang, Z. Ibrahim, J. Cha, W. P. A. Lee, G. Brandacher, and P. L. Gehlbach, “Real-time three-dimensional Fourier-domain optical coherence tomography video image guided microsurgery,” J. Biomed. Opt. 17(8), 081403 (2012).
31. M. Sylwestrzak, D. Szlag, M. Szkulmowski, I. Gorczynska, D. Bukowska, M. Wojtkowski, and P. Targowski, “Four-dimensional structural and Doppler optical coherence tomography imaging on graphics processing units,” J. Biomed. Opt. 17(10), 105020 (2012).
32. A. Gh. Podoleanu, and A. Bradu, “Master-slave interferometry for parallel spectral domain interferometry sensing and versatile 3D optical coherence tomography,” Opt. Express 21, 19324–193238 (2013).
33. A. Bradu and A. G. Podoleanu, “Calibration-free B-scan images produced by master/slave optical coherence tomography,” Opt. Lett. 39(3), 450–453 (2014).
34. H. S. Carslaw, Introduction to the Theory of Fourier's Series and Integrals (Macmillan, 1921).
35. B. P. Flannery, W. H. Press, S. A. Teukolsky, and W. Vetterling, Numerical Recipes in C (Cambridge University Press, 1992).
36. J. Lewis, “Fast template matching,” Vision Interface 95, 120–123 (1995).
37. M. Goebel, S. H. Gage, and J. F. Bartholome, “A two-stage cross correlation approach to template matching,” IEEE Trans. Pattern Anal. Mach. Intel. 6(3), 374–378 (1984).
38. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH Image to ImageJ: 25 years of image analysis,” Nat. Methods 9(7), 671–675 (2012).
1. Introduction

OCT imaging of the eye fundus is dominated by spectral domain (SD)-OCT methods. The spectrometer (Sp) based OCT method employs a broadband source and a fast linear camera in a spectrometer [1,2]. The swept source (SS) based OCT method uses a fast tunable laser and a fast photo-detector [3,4]. We will refer in what follows to both Sp based OCT and SS based OCT methods as spectral domain methods of practicing OCT, although some reports refer to both methods as Fourier domain OCT [2,5]. When producing cross-sectional OCT images (B-scans), SD-OCT methods are clearly superior to conventional time domain (TD)-OCT in terms of both sensitivity and acquisition speed [5]. However, because the conventional SD-OCT methods are based on A-scans (one dimensional reflectivity profiles), they cannot produce a real-time en-face (C-scan) image. For a frame made of L lines, each of L pixels, L^2 A-scans are acquired. Then, by post-processing, an en-face plane at any desired depth can be inferred by software cut of the volume of L^2 A-scans. The time to produce an en-face image is determined by the time required to collect all volume data plus the time needed to post-process the acquired data [6]. This involves several signal processing steps such as zero-padding, fast Fourier transformation (FFT), data re-sampling (not required when clocked swept sources are employed but compulsory in camera based imaging systems), spectral shaping, apodization as well as rendering of the en-face plane from the 3D volume of A-scans, etc. All these steps on a non-specialized computer, take more time than the time needed to execute a full frame of en-face scanning. In other words, the display rate is slower than the acquisition rate. This was not the case with the TD en-face OCT method, where an en-face image could be delivered at the rate of acquisition [7,8].

In the early years of OCT development, en-face imaging evolved as a TD method and has shown its value in ophthalmology [9] allowing association of unique en-face patterns to different pathological states of the retina [10]. Several reports have shown the value of en-face imaging in ophthalmology as well as its challenges due to the curvature of the tissue and sensitivity of the image aspect to the eye movement [11]. Some improvement in this direction was achieved by using resonant scanners [12]. Despite superiority of SD-OCT methods, TD en-face OCT imaging is preferred in material investigation [13] and art conservation [14] due to the focus control and ability to work with high numerical aperture microscope objectives. For these applications, the high-speed offered by SD-OCT methods is not necessary as the targets are usually stationary.

During the early development of SD-OCT methods in ophthalmology, the en-face view was reinstated in the form of an approximate fundus image, useful in guiding the B-scanning investigation [15,16]. Continuous increase in speed of the SD-OCT methods has led to reduction of the time to produce a software cut en-face OCT image. Line rates larger than 300 kHz [17] have been reported using fast linear cameras in spectrometers in Sp-OCT and line rates as over several MHz using fast tunable lasers in SS-OCT [18]. To speed up the process, graphics processing units (GPU) [19] and field-programmable gate arrays [20] have been reported.

The progress in acquisition speed of SD-OCT methods, reenacted the interest in en-face imaging of the eye fundus [21,22] as proven by the organization of the first congress of en-face imaging in OCT in December 2013 [23] and publication of dedicated books [24] on en-face OCT.

An attempt to directly produce en-face OCT images, using SD-OCT was proposed in [25], where the amplitude of a single frequency band is extracted from the photo-detected signal while tuning the swept source, by mixing the photo-detected signal with a reference signal of a particular chosen frequency delivered by a local oscillator. An en-face image contains points at the same axial position. This means that for these points, the same modulation of the channelled spectrum (CS) is produced. Points at the same depth value produce the same number of peaks in the CS and so when the CS is read by tuning the optical frequency, a
particular frequency is obtained for the pulsation of the photo-detected signal. However, for this method to work, the spectral scanning needs to be highly linear to produce a well-defined radio frequency component. Therefore this method is only applicable to SS-OCT systems using Fourier domain mode-locked swept laser sources, which provide a highly linear dependence between optical frequency and time sweep. The use of this method in any other SS-OCT systems requires re-sampling of data. This method presents also the disadvantage that supplementary modulation of the swept source is needed to ensure a Gaussian profile for the final coherence gate. If more en-face images are required from more depths, then more filters or mixers need to be assembled in the digital interface. To produce a new en-face image at a different depth, the volume of data needs to be read along the axial coordinate to produce the modulation corresponding to the depth wherefrom an en-face image is to be inferred from. If the calibration is not perfect, then the amplitude of the signal and the brightness in the image are lower.

Progress in providing real time en-face OCT images using FFT based SD-OCT methods has been obtained by harnessing the power of graphic processing units (GPU) cards [26–31]. By using Compute Unified Device Architecture (CUDA) parallel computing platform and GPUs, the time required by the signal processing steps mentioned above is drastically reduced and volumes can now be produced and displayed in real time. GPU can also allow quicker rendering of en-face planes from the 3D volume of A-scans.

In [32], we introduced an improved method that can directly lead to C-scan OCT images. There, we introduced a new class of spectral domain interferometry set-ups, made from two interferometers, a Master Interferometer (MI) and a Slave Interferometer (SI). These set-ups operate as time-domain interferometers in terms of their outcome, providing information from a single point in depth in the object placed in the MI, determined by the optical path difference selected in the SI. By replacing the MI with a storage bank of channeled spectra shapes (masks), where the storage delivers similar signals to those previously delivered by the MI, we regained the advantage of spectral domain interferometry consisting in simultaneous interrogation of several depths. In other words, the employment of two interferometers in the original Master/Slave (MS) configuration [32] is replaced by a two-stage process. In the first stage, “preparation”, a mirror is placed as object in the interferometer, acting as a MI, and masks are created. In the second step, “imaging”, the object replaces the mirror and the interferometer behaves like the SI, where the acquired channeled spectrum is compared with the set of channeled spectra stored as masks. The conventional spectral domain interferometry delivers an A-scan along a single line, by a single channel. In MS interferometry, points from the A-scan reflectivity profile are delivered in parallel, by multiple hardware channels, their number being equal to that of memories (masks) stored. The MS interferometry method, not being based on Fourier transformations, does not require re-sampling of data. Providing data in parallel from different depths, makes the MS method ideally suited for direct en-face imaging, like in TD-OCT, however with the sensitivity of SD-OCT methods.

In [32], it was also demonstrated that the MS method exhibits similar sensitivity and depth resolution to that provided by the FFT based technique. The MS method requires comparison operations of channeled spectra provided by the MI and by the SI. A possible comparison method consists in correlation, which makes MS method similar to an adaptive filtering method, where when the two shapes being compared are similar, a maximum is delivered. In [32], the correlation operation was implemented via three FFTs operations and processing of 200 × 200 channeled spectra to generate 64 en-face images took several tens of seconds (~60 s).

An improvement step in terms of signal processing was presented in [33], where the MS method was proven capable of generating B-scans. In this last paper [33], one of the FFT steps was transferred to the Preparation step. By storing the FFTs of the channeled spectra collected with a mirror, less FFT calculations are left for the 2nd stage, Imaging. An FFT of the current channeled spectrum is calculated and then for each depth required (i.e. for each
mask), the complex conjugate of an inverse FFT operation is calculated. Even so, the time required to produce a B-scan of \( L = 200 \) lines was 128 ms (each A-scan contains 256 points).

In this case the time required to perform a correlation to deliver the corresponding point at depth is longer than the time required for a single FFT but shorter than the time for data re-sampling followed by FFT currently used by the SD methods.

Here, we present a novel algorithm for the comparison operation required by the MS method, which is much faster than the two algorithms used in [32] and in [33]. This algorithm allows production of en-face images in real time with no need for extra hardware such as GPU cards. As the most significant result of the correlation operation, as described in [32], is the correlation for lag zero, the algorithm presented here uses a finite number of multiplication operations calculated in the vicinity of zero lag. We refer to this algorithm as a short correlation algorithm. Using this algorithm, a whole set of 36 en-face images is obtained in 3 s and real-time generation of en-face images of the eye fundus in-vivo becomes possible, being able to deliver up to four such images in real-time at a rate of 0.6 Hz, or one en-face image at 1 Hz.

2. Experimental set-up

Figure 1 presents the schematic diagram of the implemented MS-OCT system. As optical source, a swept source (SS) (Axsun Technologies, Billerica, MA), central wavelength 1060 nm, sweeping range 106 nm (quoted at 10 dB) and 100 kHz line rate is used. The interferometer configuration, I, employs two single mode directional couplers, DC1 and DC2. DC1 has a ratio of 20/80 and DC2 is a balanced splitter, 50/50. DC2 feeds a balance detection receiver from (Thorlabs, Newton, New Jersey, model PDB460C), using two photo-detectors, PhD2 and PhD2 and a differential amplifier DA, part of a Signal Acquisition Block (SAB). 20% from the SS power is launched towards the object arm, via lens L1 (focal length 15 mm), which collimates the beam towards a pair of scanners XYSH, (Cambridge Technology, Bedford, MA, model 6115) followed by an interface optics made from two lenses, L2 and L3, (both of 75 mm focal length). The power to the object O is 2.2 mW, where the object O is shown either as the retina of an eye or a flat mirror, as detailed below, in a model eye, ME, using lens L4 (focal length 22 mm).

At the other output of DC1, 80% from the SS power is directed towards the reference arm equipped with slave reference mirrors, SRM1, SRM2, placed on a translation stage, TS to adjust the optical path difference (OPD) in the interferometer. Collimating lenses L5 and L6 are similar to L1. The signal from the balanced receiver is sent to one of the two inputs of a dual input digitizer (Alazartech, Quebec, Canada, model ATS9350, 500 MB/s). A trigger signal from the SS synchronizes the acquisition (input T). The acquired channelled spectra CS are manipulated via a program implemented in Labview 2013, 64 bit, deployed on a PC equipped with an Intel Xeon processing unit, model E5646 (clock speed 2.4 GHz, 6 cores).

The MS method proceeds in two stages, depending on the position of switches K1, K2 and K3 that are used to switch the functionality of the interferometer and of the SAB. In the Preparation stage, switches K1, K2 and K3 are placed in position 1. The ME is used as object, the XYSH is on axis (at rest) and channelled spectra, \( M_p \), for \( p = 1, 2 \ldots P \), corresponding to a set of optical path differences, \( \text{OPD}_p \), measured between the reference and sample arm lengths of the interferometer are recorded and placed in the Memory block. Normally, signals \( M_p \) (masks) should be recorded at OPD values separated by half of the coherence length of the optical source or denser and stored in the Memory block part of SAB. In the Measurement stage, switches K1, K2 and K3 are all placed in position 2, the eye is placed in the object arm of the interferometer and the channelled spectrum acquired, CS, is correlated with all masks \( M_p \) in the Processing block. Correlation with each mask, \( M_p \), provides an output signal of amplitude \( A_p \), as a point in the A-scan for the OPD value used to create that mask. The processing block shown in Fig. 2(b) details how the amplitude of the signal originating from a certain depth in the sample is calculated.
Fig. 1. Schematic diagram of the Master Slave SS-OCT system. I: interferometer; SAB: signal acquisition block; K₁, K₂, K₃: switches from the position 1 (Preparation) to position 2 (Measurement); SS: swept source; DC₁, 20/80 single mode directional coupler; DC₂: 50/50 single mode directional coupler; XYSH: two-dimensional XY lateral scanning head; L₁ to L₆: lenses; O: object; ME: eye model; SAB: slave acquisition block; PhD₁, PhD₂: photo-detectors; DA: differential amplifier; CS: channelled spectrum delivered by the SAB; DOB: digital output board.

1. Correlation of the current channelled spectrum, CS, with a mask, Mₚ according to [32]. The correlation is calculated over the wavenumber axis, with variable the wavenumber, k. If the swept source is stepped in a number of M frequency steps, then the summation of products of the two terms is performed over 2M-1 points. For the data shown in this paper, 2M-1 = 1023.

2. After correlation, the signal is high-pass filtered (HPF) to remove the DC component, and rectified. The amplitude of the signal is then evaluated in k = 0. A maximum is expected to be obtained around k = 0, however its position depends on the phase between the mask and the current CS acquired, and therefore an average of values in the vicinity of zero lag is performed as detailed in the next paragraph.

The procedure of inferring the reflectivity strength from a scattering center at a depth zₓ using the MS method is shown in Fig. 2(b). For simplicity, the channelled spectrum CS and the mask, Mₓ, are shown as sinusoidal signals. In practice, CS is a superposition of many chirped sinusoidal shapes and the Mₓ is a chirped sinusoidal versus wavenumber, as no linearization of data is performed. For comparison, Fig. 2(a) shows the whole A-scan obtained via a single FFT step from the channelled spectrum CS. If all points along the A-scan are needed, then the MS method in Fig. 2(b) needs to be repeated for all P masks, Mₓ with p = 1,2…P.
3. Algorithms to implement the comparison operation

The comparison operation proposed in [32] is correlation. The correlation of two signals \( s(k) \) and \( h(k) \) measured in the wavenumber “k” space is defined by the cross-correlation integral [34]:

\[
\text{Corr}(k) = s(k) \otimes h(k) = \int_{-\infty}^{+\infty} s(u)h(u + k) \, du
\]

As it can be observed in Eq. (1), the correlation process yielding to \( \text{Corr}(k) \) involves a few steps:

1. Shift of the signal \( h(u) \) by a lag \( k \).
2. Multiply the shifted signal \( h(u + k) \) by \( s(u) \).
3. Integrate \( s(u)h(u + k) \) to obtain the value of the cross-correlation in a single lag \( k \).
4. Steps 1-3 have to be completed for all positive and negative values of the lag \( k \).

In practice, the two signals to be correlated are digitized. Considering signal \( s(k) \) represented by \( M \) elements and signal \( h(k) \) represented by \( N \) elements, the cross-correlation of the two signals is defined as [35]:

\[
\text{Corr}(k) = \sum_{m=0}^{M-1} s(m)h(k + m)
\]

The result of the cross-correlation is an array of \( N + M - 1 \) elements (\( k = -(N-1), -(N-2), \ldots, -1, 0, +1, \ldots, (M-2), + (M-1) \)).

According to Eq. (2), a number of \( M \times N \) additions and \( M \times N \) multiplications are required to perform a direct cross-correlation.

In the practice of fast digital signal processing, the cross-correlation is calculated based on Fourier transformations, which depending on the values of \( M \) and \( N \) can be performed faster than the direct method. The correlation theorem states that by multiplying the FFT of a function \( h \) by the complex conjugate FFT of the other \( s \) leads to the FFT of their cross-correlation [36]. Therefore, the cross-correlation can be obtained as:

\[
\text{Corr}(k) = i\text{FFT} \left[ \text{FFT}^* \left[ s(u) \right] \text{FFT} \left[ h(u) \right] \right]
\]
where iFFT signifies the inverse FFT. To obtain the same number of elements in the correlation result (i.e. \( N + M - 1 \) elements when using the direct method), prior to FFT, the two functions require zero padding.

In the experiments, the two functions \( s(k) \) and \( h(k) \) whose correlation is evaluated are the channeled spectrum, \( CS(k) \) and the Masks \( M_p(k) \), for \( p = 1, 2, \ldots, P \) OPD values.

The number of operations required to perform a single FFT operation is \((N + M - 1) \log_2(N + M - 1)\) complex multiplications and \((N + M - 1) \log_2(N + M - 1)\) additions [35,37]. However, in practice other efficient discrete Fourier transform (DFT) algorithms can be used. To achieve the best FFT computation time, the number of elements of \( s \) and \( h \) has to be a power of 2, in which case a radix 2 Coolidge-Tukey algorithm can be used to compute the DFT (in this case the number of operations required to perform a single FFT operation is reduced by a factor of 2). The Labview’s FFT vi used in this paper for benchmarking purposes, uses the radix 2 Coolidge-Tukey algorithm when the number of samples is a valid power of 2, but different other efficient algorithms (such us the chirp z algorithm) to calculate the discrete Fourier transform otherwise. For the sake of simplicity, in the calculations that follow, we considered that a number of \((N + M - 1) \log_2(N + M - 1)\) complex multiplications and \((N + M - 1) \log_2(N + M - 1)\) additions are required to perform a single FFT. To compute the cross-correlation according to Eq. (3), a number of three FFTs are required, hence \(3(N + M - 1) \log_2(N + M - 1)\) complex multiplications and \(3(N + M - 1) \log_2(N + M - 1)\) additions or \(12(N + M - 1) \log_2(N + M - 1)\) real multiplications and \(12(N + M - 1) \log_2(N + M - 1)\) additions.

Equation (3) can also be implemented using 2 FFTs only, when one of the two signals does not change during the acquisition of data, procedure used in [33]. In this case, \(8(N + M - 1) \log_2(N + M - 1)\) real multiplications and \(8(N + M - 1) \log_2(N + M - 1)\) additions are required to compute Eq. (3).

Simple mathematical calculations show that correlation implemented via Eq. (2) can be faster than the FFT based methods only for small values of \( M \) and \( N \). On the other side, when large values of \( M \) and \( N \) are required, the FFT can be used to speed up the production of the cross-correlated signal. For \( N = M = 512 \), the direct correlation implemented via Eq. (2) requires 262,144 operations while when implemented via FFTs (Eq. (3) only about 122,740 operations when three FFT operations are performed and 81,828 when two FFTs are used. Finally, a single FFT of a signal of 1024 elements involves only about 10,240 operations. Typically in both Sp-OCT and SS-OCT, the channeled spectra are digitized from a few hundred up to a few thousands points. Consequently, the fastest way of producing a single point in the A-scan is via the conventional FFT based SD-OCT method. As another tremendous advantage of the conventional FFT based SD-OCT method, reflectivity of all points is obtained via a single FFT operation. Therefore, the only way forward for the MS method to be competitive is to perform faster rendering of a point in depth doubled by parallel processing to obtain all points of an A-scan in a similar time or less than the time required by conventional FFT based SD-OCT.

The amplitude, \( A \), of the correlation calculation in Eq. (2) is evaluated in \( k = 0 \). However, in practice, there might be phase instabilities from the time the mask was acquired until the actual measurement. Therefore, by using the correlation result for \( k = 0 \) only, may lead to a too low strength. Therefore, an average is executed over the \( Corr(k) \) results within a window of size \( 2W-1 \) points around \( k = 0 \):

\[
A = \sum_{k=-(W-1)}^{k=(W-1)} |Corr(k)|
\] (4)

When the channeled spectrum \( CS(k) \) and the Mask \( M_p(k) \) are employed in Eq. (2), the amplitude delivered by Eq. (4) is \( A_p \), i.e. the reflectivity of the scattering center located at the depth corresponding to the OPD \( p \), used to determine the mask \( M_p \).
The window 2W-1 can be conveniently used to define the depth resolution of the system and tweak its signal-to-noise ratio (SNR) and sensitivity as demonstrated in [32]. A narrow lag window 2W-1 determines a good depth resolution but renders the sensitivity worse.

The short correlation algorithm we are proposing here is based on the fact that in order to sensitize the depth selection, the correlation signal needs to be windowed. Instead of calculating a complete correlation according to Eq. (2) followed by truncation of results around k = 0 according to Eq. (4), the algorithm proposed here consists in the direct evaluation of the truncated result only. The windowed signal can be obtained straight away by computing \( \text{Corr}(k) \) in Eq. (2) for 2W-1 points only, where k = -W, -(W-1), -1, 0, + 1, + (W-1), + W and W<<M. Consequently, the total number of operations (multiplications and additions) required to produce a single point in depth in the \textit{en-face} image is reduced to:

\[
\text{Number Total Operations} = 2[(2W+1)M - W(W+1)]
\]

when M = N. For example when M = N = 512 and W = 10, according to Eq. (5), a number of 10,642 multiplications and the same number of additions are required to produce a single point in the \textit{en-face} image, which brings the time required by the short correlation method proposed here within the same range as the conventional FFT based SD-OCT. As shown in [32], the larger W, the worse the depth resolution, therefore the value of W should be kept small. This procedure allows an accurate calculation of \( A_p \) with a reduced number of operations. For W = 5, only 5,602 multiplications and the same number of additions are needed.

4. Implementation of the method

We aimed to use the system presented in Fig. 1 to produce MS-OCT \textit{en-face} images of lateral size \( L^2 = 200 \times 200 \) pixels. Both galvo-scanners are driven with triangular ramps. For 200 pixels, each acquired in the period of the swept source of 10 \( \mu \)s, the fast galvo-scanner requires a ramp duration of 2 ms. Only one active ramp is used, so the period of the triangular signal applied to the fast galvo-scanner is 4 ms, hence the frequency of the signal is 250 Hz. \( L = 200 \) lines in the frame requires 400 ms, i.e. the period of the triangular signal applied to the frame scanner is 0.8 s therefore a new full data set can be acquired again after 1.6 s. This determines a frame rate of 0.625 Hz (the time to the galvo-scanners to return to their initial position have been taken into account). As a consequence, the time required to acquire the full data set of \( 200 \times 200 = 40,000 \) channelled spectra is 0.8 s. After acquisition, the following 0.8 s can be used for data processing.

A number of 512 points were used to sample each channelled spectrum (by digitizing the acquired signals using a sampling rate of 100 MS/s). This number of sampling points was used in all methods investigated to allow comparison of the time required. We used this number of sampling points to evaluate comparatively: (i) the traditional FFT based SD-OCT method, either with data not resampled or with data re-sampled (in which case a cubic spline interpolation has been used); (ii) the MS-OCT using 2 FFT based correlation; (iii) the MS-OCT using 3FFT based correlation and (iv) the novel algorithm presented in this paper for MS-OCT using the short correlation calculation. The interpolation and FFT steps were executed using dedicated Labview\vis.

Figure 3 displays the time, \( t_1 \), required by the short correlation method to obtain the reflectivity value of a single point, \( p \), at depth, using the mask \( M_p \), for different lag values W. Data were produced using a software code isolated from the extended software program performing not only the production of the \textit{en-face} images but also the acquisition of data. The values on the vertical axis on the right are obtained by simple multiplication of the left vertical axis by \( L^2 = 40000 \). These can be interpreted as the time required to produce an \textit{en-face} image of \( L^2 \) size, using the novel method presented in this paper. For comparison, horizontal lines mark the true time required by the previous MS methods used in [32] and [33] for generating.
an \textit{en-face} OCT image. In addition, horizontal lines showing the time required by the conventional FFT based SD-OCT methods are also shown. These represent the time required by the calculation of $L^2$ A-scans (FFT and interpolations) to which the time for cutting the volume of $L^2$ A-scans was added too.

Fig. 3. Time required to infer the reflectivity of a point at depth, $t_1$ (shown on the left vertical axis) using the short correlation method versus the window size, $W$ (circles). The red horizontal lines indicate the time (values shown on the right vertical axis) required by the conventional FFT-based SD-OCT as well as the time of two other MS-OCT methods, evaluated for $W = 255$. The values on the vertical axis on the right are obtained by simple multiplication of the left vertical axis values by $L^2 = 40000$.

According to Fig. 3, FFT operations for all $L^2 = 200 \times 200$ pixels in transversal section and volume cutting can be produced using the traditional FFT based SD-OCT method in about 94 ms (36 ms are required to compute $L^2$ FFTs, and 58 ms to render the \textit{en-face} projection from the data set). However, if re-sampling of data is needed and if a cubic spline interpolation for instance is used, then 335 ms are needed to produce a $200 \times 200$ pixels \textit{en-face} image.

As the FFT and interpolation operations have to be performed only once for a given data set, a number of 13 (no resampling required) or 9 (resampling required) images can be produced in a time less than that required for \textit{en-face} scanning, of 0.8 s.

The correlation methods can deliver a C-scan in 306 ms when evaluated via 3 FFTs according to procedure in [32] and in 204 ms when evaluated via 2 FFTs, according to procedure in [27], hence up to 2 or 3 \textit{en-face} images respectively could be produced in 0.8 s.

However, for all the three situations presented above (conventional FFT with data resampling, MS based on 3 FFTs and MS based on 2 FFTs), if faster swept sources are used, then the acquisition time to collect the $L^2$ data set becomes shorter than the time to produce an \textit{en-face} image and the real time \textit{en-face} display is no longer viable.

For the short correlation method proposed here, the larger the lag window 2W-1, the slower the method becomes. For $W > 100$, a single \textit{en-face} image requires 586 ms, hence a single image in 0.8 s. For $W < 10$ however, the production of an \textit{en-face} image via the novel algorithm of short correlation can clearly compete even with the FFT based SD-OCT method with no data re-sampling, as the time for a single point at depth becomes comparable with 94 ms.

Obviously, when CUDA implementations on graphics cards are employed, the rendering of all the \textit{en-face} projections from the data set in FFT based SD-OCT can be nearly as fast as rendering a single image. On the other hand, for each mask, the MS method produces a single \textit{en-face} image. However, the MS method is ideally suited for parallel computing algorithms on GPUs due to its parallel nature.

In practice, in parallel with the production of the \textit{en-face} images, some other tasks have to be performed, that increase the demands on the CPU activity. In Fig. 4, a flowchart summarizing the main simultaneous tasks required for the production in real-time of the images is presented. The data acquisition board (Alazartech) acquires data from the photo-
detector according to the timings set by the signal triggers generated by the SS and by a digital output board (DOB) that at the same time controls the position of the galvo-scanners. The Processing block is used to perform both the Short correlation method and FFT function. Continuously, B-scan images produced at a frame rate of 250 images per second are produced and displayed on screen. They are not to be used for investigation, but for an approximate indication of the distance between the eye fundus and the zero optical path difference, as detailed below. For each B-scan image the 200 A-scans are averaged. Then, the maximum value of each averaged signal and its position from zero optical path difference is evaluated. The sinusoidal audio signal that is generated by the sound card of the computer has its amplitude proportional to the maximum value of the averaged A-scans and a frequency proportional to the distance of the main reflecting point in the tissue from the OPD = 0 value. This helps the user as well as the volunteer (patient) to position the eye axially without the need to see the computer’s display.

Two regimes of operation are possible using a data set of L^2 channelled spectra.

1. Volume acquisition. Once all data set of L^2 channelled spectra are acquired, corresponding to a 3D volume, the en-face images are inferred by the comparison operation, here based on the new algorithm of short correlation, followed by their display and recording.

2. Real time display of en-face images. This regime is similar in outcomes to the time domain version of en-face OCT, where a real time displayed en-face image is delivered and its depth adjusted by actuating on the reference path length. Using W = 10, a number of up to 4 simultaneous en-face images from 4 different depths are simultaneously produced and displayed on screen.

Eye guidance

Accurate positioning of the eye pupil in the focal plane of lens L_3 is important to generate en-face images with maximum sensitivity. We implemented a procedure where the position of the eye can be inferred from the B-scan images produced continuously during data acquisition, at a rate of 25 Hz. These images are produced with the channelled spectra delivered by the photo-detector with no resampling (i.e. the k-clock of the swept source was not engaged). The immediate effect is that the B-scan images lack resolution. However, they can still provide sufficient axial guidance information to position the eye correctly.

The sampling rate of the digitizer limits the maximum number of cycles in the channelled spectrum to be correctly displayed. For M = 512, up to 256 cycles in the CS can be decoded into 256 values of OPD = 2z depths. The round trip axial resolution is determined by the FWHM tuning range of Axsun, \Delta \lambda = 86 \text{ nm} (measured experimentally) and the central wavelength \lambda = 1050 \text{ nm}, as \frac{0.88\lambda^2}{\Delta \lambda} = of 11.3 \text{ \mu m}. This corresponds to an OPD axial range of AR = 2.9 \text{ mm}, to a depth range of 1.45 \text{ mm} and to an axial resolution in depth of 5.65 \text{ \mu m} respectively, last two quantities being measured in air. A scattering center situated at 2AR or at 3AR will generate a CS with double and respectively three times more modulating cycles.
than the CS for a depth at the AR value. In other words, scattering centers creating a CS with 512 and 768 cycles will be placed in the B-scan image at the depth determining a CS with 256 peaks. For considerations of speed, we limited M to 512, but this presents the disadvantage that several multiples of B-scans are displayed as the head is moved axially, instead of a single B-scan moving within a 2AR or a 3AR range respectively. The correct position is where the brightest images with best delineation of edges is obtained. Then once the eye is within this useful axial range, the only other adjustment needed before acquisition of data is to avoid any mirror terms manifesting in the B-scans. This is achieved by moving the head slightly away.

For this purpose, within the Labview software, in the first step, the 200 A-scans used to build each B-scan are averaged. Then, the maximum value of the averaged A-scan is translated into the strength of an audio signal in the sound card of the PC whose frequency is made proportional to the distance of the eye fundus from OPD = 0.

**Regime 1: Volume acquisition (Non real-time operation)**

This regime is illustrated with images collected from the fovea and the optic nerve of AP. 48 masks are initially recorded in the Preparation stage using a mirror, for \( \Delta z = 25 \, \mu m \) measured in air.

![Fig. 5. 36 en-face images of the fovea area obtained using the short correlation method (W = 10) showing different retinal and choroidal layers. Each image is 200 × 200 pixels. The depth separation between consecutive images is 25 \( \mu m \) measured in air, as dictated by the masks acquired in the Preparation stage. The eye was placed in such a way to avoid mirror terms perturbing the images. Lateral image size: 3 mm × 3 mm. The first image (top, left) is recorded at an approximate depth \( z = 100 \, \mu m \).](image-url)
Fig. 6. 48 en-face images of the optic nerve area obtained using the short correlation method (using $W = 10$) showing the lamina cribrosa. Each image is $200 \times 200$ pixels. The depth separation between consecutive images is $25 \, \mu m$ measured in air, as determined by the set of masks acquired in the Preparation stage. The eye orientation and voltage applied to the galvo-scanners were adjusted to fit all the lamina in the middle of the image. The optic nerve was placed slightly away from OPD = 0 to avoid mirror terms perturbing the images. The first image (top, left) is recorded at a depth $z = 50 \, \mu m$.

In Fig. 5 and 6, 36 and 48 C-scan images from the fovea and the optic nerve areas respectively are presented. As the time required to produce a single C-scan is around 80 ms, the 36 images of the fovea presented in Fig. 6 were produced in 2.88 s, while the 48 images of the optic nerve shown in Fig. 6 were generated in 3.84 s. For this regime (not in real-time) further improvement of time is possible making use of CUDA parallel computing
architectures on graphic cards. Parallelization is possible in both directions, along the depth axis, typical for the MS method as well as along the transversal direction.

Using the stack of C-scans shown in Figs. 5 and 6, 3D reconstructions of the eye fundus can be performed, hence B-scans can be rendered from the volume thus obtained. Such images are demonstrated in Fig. 7. Figure 7(a) shows a 3D reconstruction of the fovea using the 36 C-scans presented in Fig. 5. A B-scan rendered from this volume at the position shown by the yellow line is shown in Fig. 7(b). In Fig. 7(c) a 3D reconstruction of the optic nerve using the 48 C-scans shown in Fig. 6 is demonstrated. Also a B-scan rendered from Fig. 7(c) at the position shown by the yellow line is presented in Fig. 7(d). All images demonstrated in Fig. 7, being inferred from MS-OCT C-scans images are not produced by Fourier transformation, hence all other intermediate steps before FFT such as data re-sampling, spectral shaping, apodization, zero padding are eliminated.

Fig. 7. (a) 3D reconstruction using the set of en-face images shown in Fig. 5. (b) Rendered B-scan image from (a), along the yellow line shown. (c) 3D reconstruction using the set of en-face images shown in Fig. 6. (d) Rendered B-scan image from (c), along the yellow line shown.

Regime 2: Real-time operation

For the example of image size selected as described above, up to four en-face images can be produced in real time.

To demonstrate the functionality of the system in real-time, movies showing the en-face images of different layers of the foveal and optic nerve area are demonstrated.

In Fig. 8 four single excerpts from a movie (Media 1) showing en-face images of the fovea area are shown (a-d) together with the cross sections used to compute the audio signal employed for eye guidance. The 4 simultaneous en-face images, separated by 25 μm measured in air correspond to masks recorded from the area demarcated by the yellow rectangle shown in the B-scan image (Fig. 8(e)).

Figure 9 shows four single excerpts from a movie (Media 2) showing images from the optic nerve area. To present images from a larger depth range, the four images are separated here by 50 μm measured in air (the masks were recorded approximately from the area demarcated by the yellow rectangle shown in the B-scan (Fig. 9(e))).

The images above exhibit movement disturbance. Deliberately, no bite bars and no tracking were employed in order to assess the capability of the MSOCT to deliver C-scans fast.
To record the two movies (Media 1 and Media 2), the volunteer’s head was axially moved, so different layers appear in the en-face images. Only the audio signal inferred from the B-scan was used for guidance.

5. Conclusion

A MS-OCT system able to produce images from the eye fundus in real-time is demonstrated. The production in real-time of the en-face images is possible by implementing a short
correlation algorithm that efficiently replaces the FFT based calculations of the cross-correlation required by the comparison operation of channeled spectra. In the previous reports on the MS method the computing time of a single en-face image was 580 ms or 368 ms when a 3 FFTs based correlation [32] or a 2 FFTs based correlation [33] was used respectively. The computing time of the short correlation MS method depends on the lag value W, the smaller W value, the faster the production of the en-face images is. Lag values as small as W = 5 or W = 10 can still provide excellent depth resolutions and reasonable sensitivities as demonstrated in [32]. When W = 5, a factor of about 7 or 14 shorter computing times than the values reported in [27] for 2 FFTs or in [32] for 3 FFTs based correlation method respectively are obtained. For W = 10, the short correlation method can perform faster by a factor of 5 or 10 than the 2 FFTs or 3 FFTs based correlation method respectively. Even so, the conventional FFT based SD-OCT method is still faster, for the example of 200x200 pixels used here, being capable of delivering twice faster en-face OCT images than the improved MS-OCT method presented. However, in case the spectral scanning is highly nonlinear, then higher order linearization procedures may be needed, which may require longer time for the SD-OCT method, which inclines the balance in favor to the MS-OCT. In case the nonlinear swept source is provided with a clock, no resampling is needed before FFT and the conventional FFT based method is the fastest method for delivering en-face OCT images.

With our current computing capacity (only a single CPU with six cores) we are able to generate 48 en-face images in 3.84 s when the system is not operating in real-time and up to four simultaneous images in less than 0.8 s in real-time (W = 10). To produce a larger number of images within the same time interval, faster multi-core processors or CUDA parallel computing architectures on GPU can be used. When operating in real-time, a single en-face image can be produced in less than 200 ms (which allows 4 images within the frame scanner return time). When only a single image is displayed, 0.8 s are needed to acquire the full data set and about 0.2 s to produce the image, hence a frame rate of producing en-face images in real time of 1 Hz. The rate they can be produced at can only be increased by using a faster sweeping rate. We demonstrate here a system employing a swept source laser able to perform at 100 kHz, hence 1.6 s to produce up to four en-face images in real time. Higher tuning speed swept sources exist. If a four times faster sweeping speed source is used (400 kHz), then the acquisition time to acquire the full 3D data set would become 0.2 s and producing a single real-time en-face image of 200 × 200 pixels size would only take 0.4 s, so real-time operation could be achieved at a frame rate of 2.5 Hz.

The production of en-face images via the MS-OCT method is highly parallelizable, the utilization of the CUDA parallel computing architecture on GPU cards being an interesting avenue to follow. By transferring the entire set of data to the GPU, the production of the images could be made much faster, perhaps, nearly instantaneous once the full data set is transferred from the host (PC) to the GPU. For a real-time operation, if only a single image needs to be displayed, CUDA does not help as the transfer of data to and from the device could take as long as the time to produce an image on the CPU (~0.2 s). However, as a single image can be produced on the GPU nearly instantaneous, it is possible within 0.2 s to produce, in real-time a larger number of en-face images when using the MS-OCT method in comparison with using the traditional FFT based SD-OCT method. If such avenue is used, then larger sets of channeled spectra can be acquired for better definition than used here of only 200 × 200 pixels.

We have demonstrated an architecture of signal processing which makes use of both the MS-OCT method for en-face imaging and the conventional FFT based SD-OCT method for guidance in positioning the eye. This approach changes the current practice where cross sections are produced and an average en-face image of the eye fundus is used for guidance.

We have also demonstrated production of cross sections obtained from a stack of en-face images, approach also different from current practice where volumes of B-scans are assembled first and C-scans are inferred second, from such volumes.
In all scenarios presented, *en-face* OCT images are inferred with no need of FFT and so, both C-scan and B-scan images obtained from the stack of C-scans are obtained with no need for calibration of spectral data.

More work is required to take full advantage of the parallelization offered by the Master/Slave technology. With the new algorithm demonstrated here, the time to produce *en-face* images using the MS-OCT becomes comparable with the time required by conventional SD-OCT methods, although still almost twice longer. However, the time improvement reported here when considered in combination with the other advantages in terms of hardware costs [32], makes the MS-OCT a method worth considering in imaging the eye. Not needing linearization, MS-SS/OCT can operate with a simpler swept source, with no clock, or even with potentially highly nonlinear tunable lasers. A MS-OCT set-up operates in terms of decay with depth and axial resolution at the level of an ideally corrected FFT based SD-OCT set-up. Therefore, an MS-OCT set-up achieves better axial range and better sensitivity than any improper corrected FFT based SD-OCT set-up.

**Acknowledgments**

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme, Advanced Grant agreement “COGATIMABIO”, No. 249889. A. Podoleanu is also supported by the NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology.