

Master slave *en-face* OCT/SLO

Adrian Bradu^{1*}, Konstantin Kapinchev², Frederick Barnes² and Adrian Podoleanu¹

¹Applied Optics Group, School of Physical Sciences, University of Kent, CT2 7NH, Canterbury, United Kingdom

²School of Computing, University of Kent, CT2 7NF, Canterbury, United Kingdom

*a.bradu@kent.ac.uk

Abstract: Master Slave optical coherence tomography (MS-OCT) is an OCT method that does not require resampling of data and can be used to deliver *en-face* images from several depths simultaneously. As the MS-OCT method requires important computational resources, the number of multiple depth *en-face* images that can be produced in real-time is limited. Here, we demonstrate progress in taking advantage of the parallel processing feature of the MS-OCT technology. Harnessing the capabilities of graphics processing units (GPU)s, information from 384 depth positions is acquired in one raster with real time display of up to 40 *en-face* OCT images. These exhibit comparable resolution and sensitivity to the images produced using the conventional Fourier domain based method. The GPU facilitates versatile real time selection of parameters, such as the depth positions of the 40 images out of the set of 384 depth locations, as well as their axial resolution. In each updated displayed frame, in parallel with the 40 *en-face* OCT images, a scanning laser ophthalmoscopy (SLO) lookalike image is presented together with two B-scan OCT images oriented along rectangular directions. The thickness of the SLO lookalike image is dynamically determined by the choice of number of *en-face* OCT images displayed in the frame and the choice of differential axial distance between them.

©2015 Optical Society of America

OCIS codes: (110.4500) Optical coherence tomography; (170.0110) Imaging systems; (330.4460) Ophthalmic optics and devices; (120.3890) Medical optics instrumentation; (200.4960) Parallel processing.

References and links

1. S. Duke-Elder, “*System of Ophthalmology*” (Henry Kimpton Eds, 1867).
2. R. H. Webb, G. W. Hughes, and F. C. Delori, “Confocal scanning laser ophthalmoscope,” *Appl. Opt.* **26**, 1492-1499 (1987).
3. A. Gh. Podoleanu, George M. Dobre, David J. Webb, and David A. Jackson, “Simultaneous *en-face* imaging of two layers in the human retina by low-coherence reflectometry,” *Opt. Lett.* **22**, 1039-1041 (1997).
4. A. Gh. Podoleanu, M. Seeger, G. M. Dobre, D. J. Webb, D. A. Jackson, and F. W. Fitzke, “Transversal and longitudinal images from the retina of the living eye using low coherence reflectometry,” *J. Biomed. Opt.* **3**, 12-20 (1998).
5. A. Gh. Podoleanu and R. B. Rosen, “Combinations of techniques in imaging the retina with high resolution,” *Prog. Ret. Eye Res.* **27**, 464-499 (2008).
6. C. Hitzenberger, P. Trost, P.-W. Lo, and Q. Zhou, “Three-dimensional imaging of the human retina by high-speed optical coherence tomography,” *Opt. Express* **11**, 2753-2761 (2003).
7. L. Neagu, A. Bradu, L. Ma, J. W. Bloor, and A. Gh. Podoleanu, “Multiple-depth *en face* optical coherence tomography using active recirculation loops,” *Opt. Lett.* **35**, 2296-2298 (2010).
8. V. J. Srinivasan, D. C. Adler, Y. Chen, I. Gorczynska, R. Huber, J. S. Duker, J. S. Schuman, and J. D. Fujimoto, “Ultrahigh-speed optical coherence tomography for three-dimensional and *en face* imaging of the retina and optic nerve head,” *Invest Ophthalmol. Vis. Sci.* **49**, 5103-5110 (2008).
9. A. Gh. Podoleanu, “*Principles of en-face optical coherence tomography: real time and post-processing en-face imaging in ophthalmology*,” (in *Clinical en-face OCT atlas*, B. Lambruso, D. Huang, A. Romano, M. Rispoli, G. Coscas eds., J.P. Medical Ltd 2013), Chap. 1.
10. T. Klein, W. Wieser, C. M. Eigenwillig, B. R. Biedermann, and R. Huber, “Megahertz OCT for ultrawide-field retinal imaging with a 1050 nm Fourier domain mode-locked laser,” *Opt. Express* **19**, 3044-3062 (2011).
11. Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. J. Liu, M. F. Kraus, H. Subhash, J. G. Fujimoto, J. Hornegger, and D. Huang, “Split-spectrum amplitude-decorrelation angiography with optical coherence tomography,” *Opt. Express* **20**, 4710-4725 (2012).

12. B. Braaf, K. V. Vienola, C. K. Sheehy, Q. Yang, K. A. Vermeer, P. Tiruveedhula, D. W. Arathorn, A. Roorda, and J. F. de Boer, "Real-time eye motion correction in phase-resolved OCT angiography with tracking SLO," *Biomed. Opt. Express* **4**, 51-65 (2013).
13. B. Lambruso, D. Huang, A. Romano, M. Rispoli, and G. Coscas, *Clinical en-face OCT atlas*, (J.P. Medical Ltd. 2013)
14. First international congress of *en-face* OCT, http://www.maculasociety.org/files/FIRST_ANNOUNCEMENT_INGLESE.pdf, Rome 2013.
15. Second International Congress on "En Face" OCT imaging New Developments in OCT, OCT Angiography, <http://www.brunolumbroso.it/pdf/second-international-congress-en-face.pdf>, Rome, 2014.
16. 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**, 026002 (2013).
17. J. Probst, D. Hillmann, E. Lankenau, C. Winter, S. Oelckers, P. Koch, and G. Hüttmann, "Optical coherence tomography with online visualization of more than seven rendered volumes per second," *J. Biomed. Opt.* **15**(2), 026014 (2010).
18. K. Zhang and J. U. Kang, "Graphics processing unit accelerated non-uniform fast Fourier transform for ultrahigh-speed, real-time Fourier-domain OCT," *Opt. Express* **18**(22), 23472-23487 (2010).
19. 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).
20. J. U. Kang, Y. Huang, K. Zhang, Z. Ibrahim, J. Cha, W. P. A. Lee, G. Brandacher, and P. L. Gehlbach, "Realtime three-dimensional Fourier-domain optical coherence tomography video image guided microsurgies," *J. Biomed. Opt.* **17**(8), 081403 (2012).
21. 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), 100502 (2012).
22. R. Leitgeb, W. Drexler, A. Unterhuber, B. Hermann, T. Bajraszewski, T. Le, A. Stingl, and A. Fercher, "Ultrahigh resolution Fourier domain optical coherence tomography," *Opt. Express* **12**, 2156-2165 (2004).
23. 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**, 457-463 (2002).
24. Z. Hu and A. M. Rollins, "Fourier domain optical coherence tomography with a linear-in-wavenumber spectrometer," *Opt. Lett.* **32**, 3525-3527 (2007).
25. B. Potsaid, B. Baumann, D. Huang, S. Barry, A. E. Cable, J. S. Schuman, J. S. Duker, and James G. Fujimoto, "Ultrahigh speed 1050nm swept source / Fourier domain OCT retinal and anterior segment imaging at 100,000 to 400,000 axial scans per second," *Opt. Express* **18**, 20029-20048 (2010).
26. B. Liu, E. Azimi, and M. E. Brezinski, "True logarithmic amplification of frequency clock in SS-OCT for calibration," *Biomed. Opt. Express* **2**, 1769-1777 (2011).
27. B. Park, Mark C. Pierce, B. Cense, S-H. Yun, M. Mujat, G. Tearney, B. Bouma, and Johannes de Boer, "Real-time fiber-based multi-functional spectral-domain optical coherence tomography at 1.3 μm ," *Opt. Express* **13**, 3931-3944 (2005).
28. K. Wang, G. Huang, Z. Ding and L. Wang, "High-speed spectral-domain optical coherence tomography at 830 nm wavelength," *Proc. SPIE* **6826**, 68260A (2008).
29. Y. Yasuno, V. D. Madjarova, S. Makita, M. Akiba, A. Morosawa, C. Chong, T. Sakai, K-P. Chan, M. Itoh, and T. Yatagai, "Three-dimensional and high-speed swept-source optical coherence tomography for in vivo investigation of human anterior eye segments," *Opt. Express* **13**, 10652-10664 (2005).
30. M. Mujat, B. Park, B. Cense, T.C. Chen, and J.F. de Boer, "Autocalibration of spectral-domain optical coherence tomography spectrometers for in vivo quantitative retinal nerve fiber layer birefringence determination," *J. Biomed. Opt.* **12**, 041205-041205-6 (2007).
31. B. R. Biedermann, W. Wieser, C. M. Eigenwillig, G. Palte, 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**, 2556-2558 (2008).
32. A. Gh. Podoleanu and D. A. Jackson, "Combined optical coherence tomograph and scanning laser ophthalmoscope," *Electron. Lett.* **34**, 1088-1090 (1998).
33. A. Gh. Podoleanu, G. M. Dobre, R. G. Cucu, R. Rosen, P. Garcia, J. Nieto, D. Will, R. Gentile, T. Muldoon, J. Walsh, L. A. Yannuzzi, Y. Fisher, D. Orlock, R. Weitz, J. A. Rogers, S. Dune, and A. Boxer, "Combined multiplanar optical coherence tomography and confocal scanning ophthalmoscopy," *J. Biomed. Opt.* **9**, 86-93 (2004).
34. 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**, 444-452 (2005).
35. A. Gh. Podoleanu and A. Bradu, "Master-slave interferometry for parallel spectral domain interferometry sensing and versatile 3D optical coherence tomography," *Opt. Express* **21**, 19324-19338 (2013).
36. A. Bradu and Adrian Gh. Podoleanu, "Imaging the eye fundus with real-time en-face spectral domain optical coherence tomography," *Biomed. Opt. Express* **5**, 1233-1249 (2014).
37. A. Bradu and A. Gh. Podoleanu, "Calibration-free B-scan images produced by master/slave optical coherence tomography," *Opt. Lett.* **39**, 450-453 (2014).
38. W. J. Donnelly, F. Romero-Borja, and A. Roorda, "Optimal pupil size for axial resolution in the human eye," *Invest. Ophthalmol. Vis. Sci.* **42**, 161 (2001).

39. A. Bradu, M. Maria, and A. Podoleanu, "Demonstration of tolerance to dispersion of master/slave interferometry," *Opt. Express* **23**, 14148-14161 (2015).
 40. 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), 030511 (2010).
 41. Y. Huang, X. Liu, and J. U. Kang, "Real-time 3D and 4D Fourier domain Doppler optical coherence tomography based on dual graphics processing units," *Biomed. Opt. Express* **3**, 2162-2174 (2012).
 42. X. Li, G. Shi, and Y. Zhang, "High-speed optical coherence tomography signal processing on GPU," *J. Phys. Conf. Ser.* **277**, 012019 (2011).
 43. A. Bradu, K. Kapinchev, F. Barnes, and A. Podoleanu, "On the possibility of producing true real-time retinal cross-sectional images using a graphics processing unit enhanced master-slave optical coherence tomography system," *J. Biomed. Opt.* **20**, 076008 (2015).
-

1. Introduction

Before the optical coherence tomography (OCT) advent, the most commonly used imaging modalities in ophthalmology based on retinal fundus imaging via a camera [1] and on laser scanning ophthalmoscopy (SLO) [2], delivered images with *en-face* orientation. These two traditional retinal imaging modalities can generate excellent quality images but with limited axial resolution.

In longitudinal time-domain (TD)-OCT, the generation of an *en-face* image is possible by extracting it from a 3D dataset of axial scans (A-scans). As only a few hundred axial scans per second can be acquired, the 3D dataset cannot be generated at sufficiently high speeds to produce high quality images. However, transversal time domain OCT technology allows a real time, *en-face* image to be generated without the need of acquiring an entire 3D dataset [3-5]. By detecting the back-scattered light from a specific depth while scanning the light in a raster pattern on the retina, a real time *en-face* image is produced. Although *en-face* images can be produced in real-time up to a few tens of Hz [6], the production of the *en-face* views is limited to a specific axial plane, hence simultaneous monitoring of different retinal layers is not possible. Attempts were made to increase the number of *en-face* images, such as using a low-coherence interferometer configuration, equipped in each arm with an adjustable optical path length ring [7]. Using this idea, simultaneous *en-face* OCT imaging with sufficient good signal to noise ratio could be performed from 5 different depths only in a *Drosophila melanogaster* fly.

Conventional spectral (Fourier) domain (SD)-OCT performs cross-sectional imaging by stitching together axial reflectivity profiles (A-scans) of the back-scattered light [8]. Therefore, like in longitudinal TD-OCT, before any *en-face* image can be produced, a whole volume of A-scans needs to be acquired. Only after that, an *en-face* image can be sliced at the depth of interest. Therefore, *en-face* imaging via SD-OCT became attractive only after the speed of A-scan acquisition improved to 50 - 100 kHz and a sufficiently dense volume could be acquired in a few seconds [9]. The recent progress in tuning swept sources at over 1 MHz [10], together with the progress in no dye angiography [11, 12] that requires display of vessel structure in *en-face* views, revived the interest in *en-face* OCT [13-15]. To alleviate the problem of data processing and display, graphic processing unit (GPU) cards [16-21] on Compute Unified Device Architecture (CUDA) parallel computing platforms are used. However, even if the processing time is drastically reduced, such methods are not genuinely real time, as the brightness of each pixel in the transversal section is not established simultaneously with the beam being incident on that pixel. Such association of signal to depth can only be done after all volume of data was acquired and the depth of interest was selected. In order to produce a 3D volumetric image in SD-OCT (where the *en-face* view is extracted from), each channeled spectrum acquired while scanning the probing beam over the sample is subject to a fast Fourier transform (FFT). However, before the FFT, several preparatory signal processing steps are necessary, such as zero padding, spectral shaping, apodization, dispersion compensation or data re-sampling. Without these preparatory steps, axial resolution and sensitivity suffer [22-23]. As all these steps can only be sequentially executed, the production of the 3D (*en-face*) images is slowed down. So far, several techniques involving both hardware and/or software solutions have been demonstrated to successfully eliminate or diminish the execution time of the preparatory steps for both

implementations of SD-OCT, spectrometer based (Sp)-OCT and swept source (SS)-OCT [24-31]. All the hardware methods demonstrated so far complicate the optics hardware, add to the cost of the system and do not yield to a perfect preparation of the signal. All the proposed software techniques are normally computationally expensive and limit the real-time operation of the OCT systems.

Dual imaging systems able to provide pixel-to-pixel correspondence between the SLO and the *en-face* OCT images have proven their utility in navigation and in providing complementary information. The first systems were based on early TD-OCT technology [3-4, 6, 32-33]. Later, with the advancement of SD-OCT methods, an equivalent SLO image was reported by averaging the intensity of all peaks in the A-scans, along their depths. The results per each A-scan is then mapped into an SLO image [34]. The SLO image can be produced in real-time if such average is calculated in the time interval the beam sits on each pixel. However, linearization procedures required by the conventional FFT based SD-OCT technology slow down the process of obtaining an average for all modulation components in the A-scan.

An alternative technology to the conventional FFT based SD-OCT technology has been proposed in [35], that can directly provide an *en-face* OCT image without the need for volumetric assembly of data. This method, termed as master slave (MS), is based on comparison of shapes of the channeled spectrum at the interferometer output with stored channeled spectrum shapes, comparison implemented by cross-correlation. An immediate advantage is that the correlation operation can be performed with the raw data, without the need of organizing the data in linear frequency slots, as required by FFT. Another advantage is that the MS method allows measurement of interference strength for signal reflected from any chosen depth, in real-time by employing a correlator for each depth. Therefore the MS method is ideally suited for parallel processing, with a processor required for each depth. In opposition, the conventional FFT based SD-OCT technology provides an A-scan. Therefore, if information from a selected depth is needed, such as that to assemble an *en-face* image, then information relating to that depth has to be sampled from the A-scan, i.e. two operations are required, FFT and selective sampling. In MS-OCT, the selective sampling is decided prior to measurement/imaging, with a correlator allocated to each depth of interest.

A fast algorithm for the comparison operation required by the MS method, based on a simplified correlation calculation was introduced in [36]. This has allowed production of up to four *en-face* images in real time at a rate of 0.6 Hz or of a single image at 1 Hz.

In this paper, in order to increase the number of images produced and the frame rate, a robust GPU enhanced real-time MS-based combined OCT/SLO system is presented able to provide 40 *en-face* OCT images at a rate of 1.25 Hz. At the same frame rate, a lookalike image together with two cross-section OCT images are also displayed. As all these images are produced at the frame acquisition rate, such a performance can be labeled as quasi real time.

2. Methods

2.1 Experimental set-up

The MS principle can be applied to any SD-OCT technology. This is illustrated here on a conventional swept-source OCT setup. The experimental set-up employed for this paper is identical to that used in Ref. [36]. It uses 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, a balance detection receiver (Thorlabs, Newton, New Jersey, model PDB460C) and a fast digitizer (Alazartech, Quebec, Canada, model ATS9350).

MS operates in two stages, as described in [36]. In the Master stage (preparation mode), the object is replaced by a model eye, including an achromatic lens and a mirror and channeled spectra are acquired for a set of optical path difference (OPD) values. In this regime, the interferometer operates as a Master interferometer. For this study, a number of 384 channeled spectra are deposited in the Memory block in Fig. 1. In the Slave stage (live imaging mode), the model eye is removed and the system is oriented

towards the human eye to be imaged. In this stage, the current channeled spectrum captured while scanning the probing beam over the sample is compared with p recorded channeled spectra (masks) in the Master stage and p reflectivity values from p depths in the eye are delivered.

2.2. Methodology of producing en-face MS-based OCT images

For a start, let us refer to the conventional FFT based SD-OCT method, where a 3D volumetric image has to be produced first, followed by rendering the plane desired. A 3D volume is produced by assembling A-scans, each A-scan being the result of a single FFT operation. For simplicity, let us consider that no preparatory steps are required before FFT. In this case, mathematically, we can describe the volumetric acquired content as:

$$V_{\text{SD-OCT}} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1q} \\ A_{21} & A_{22} & \dots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \dots & A_{rq} \end{bmatrix} = \begin{bmatrix} |FFT(CS_{11})| & |FFT(CS_{12})| & \dots & |FFT(CS_{1q})| \\ |FFT(CS_{21})| & |FFT(CS_{22})| & \dots & |FFT(CS_{2q})| \\ \vdots & \vdots & \ddots & \vdots \\ |FFT(CS_{r1})| & |FFT(CS_{r2})| & \dots & |FFT(CS_{rq})| \end{bmatrix} \quad (1)$$

In Eq. 1, CS_{ij} ($i = 1 \dots r, j = 1 \dots q$) are channeled spectra collected at $r \times q$ lateral positions of the scanning beam on the sample. All the p components (for $l = 1, 2, \dots, p$ depths) of a particular A-scan, A_{ij} , are obtained simultaneously in a single FFT operation (p is in principle equal to half the number of spectral sampling points of each channeled spectrum CS_{ij}). To obtain an *en-face* image of size $r \times q$, a number of $r \times q$ independent FFT operations need to be performed followed by rendering the *en-face* plane at the desired axial position, l . As the FFT operations can be executed in parallel, theoretically the time to produce a 3D volume can be reduced to the time to perform a single FFT operation, t_{FFT} . In the same way, the time to render p *en-face* planes can be reduced to the time to render a single *en-face* plane, t_{render} . In this case, the shortest time to deliver p *en-face* images, $t_{\text{en-face}}(\text{FFT})$, is:

$$t_{\text{en-face}}(\text{FFT}) = t_{\text{acq}} + t_{\text{FFT}} + t_{\text{render}} \quad (2)$$

where t_{acq} is the time to acquire the 3D dataset. To keep the comparison of different regimes of operation simple, let us consider that the acquisition and processing are not interleaved.

In MS-OCT, to create an *en-face* image there is no need to create a 3D volumetric image, so there is no need to render the *en-face* planes. An *en-face* image, produced from the axial position l can be represented as:

$$I_{\text{MS-OCT}} = \begin{bmatrix} T_{11}^1 & T_{12}^1 & \dots & T_{1l}^1 \\ T_{21}^1 & T_{22}^1 & \dots & T_{2q}^1 \\ \vdots & \vdots & \ddots & \vdots \\ T_{r1}^1 & T_{r2}^1 & \dots & T_{rq}^1 \end{bmatrix} \quad (3)$$

where each line (or column) in the above equation is now a T-scan (transversal reflectivity profile) from the axial position l . Each T_{ij}^1 component of the T-scan is the result of a Master/Slave operation with a reference channeled spectrum (mask), M_l . We denote by the Master/Slave operation, the comparison operation of channeled spectra shapes that leads to the strength of interference signal from the depth used to generate the mask. This operation requires two operations: (i) correlation and (ii) integration.

- (i) The correlation operation involves calculation of the cross-correlation between the signal proportional to the shape of the currently acquired channeled spectrum CS_{ij} ($i = 1 \dots r, j = 1 \dots q$) with the signal proportional to the shape of the mask, M_l . The p masks are recorded in the Master stage, for different optical path differences (OPD_i) between the sample and reference arm lengths of the interferometer when a highly reflective mirror is used behind a lens in an eye model [32-33].

- (ii) The integration operation is performed over a window of size $W=2w-1$ points around the maximum value of the correlation result. Thus, each component of the T-scans can be expressed as:

$$T_{ij}^l = \sum_{k=-(w-1)}^{k=+(w-1)} |CS_{ij} \otimes M_l|_k \quad (4)$$

The masks M_l should be recorded at OPD values in the set $OPD_1, OPD_2, \dots, OPD_p$, separated by half the coherence length of the optical spectrum used (broadband optical source, when using Sp-OCT or tuning bandwidth when using SS-OCT) or denser. As all the elements T_{ij}^l can be computed independently, the time to produce an *en-face* image can in principle be reduced to the time to perform a single cross-correlation:

$$T_{ij}^1 = iFFT \left[FFT(CS_{ij}) \cdot FFT(M_1)^* \right] \quad (5)$$

In the equation above, $iFFT$ denotes the inverse fast Fourier transform, while the "*" operator is the complex conjugate. As suggested in a previous report [37], to speed up the process, we can pre-compute and record the $FFT(M_l)^*$ values. This reduces the time to work out the reflectivity T_{ij}^1 of a point in the *en-face* image to the time to sequentially execute the two remaining FFTs operations in Eq. (5) only. To obtain an *en-face* image of size $r \times q$, a number of $r \times q$ independent cross-correlations have to be performed. To produce p *en-face* images, a number of $r \times q \times p$ independent cross-correlations are needed. As all these operations can be executed in parallel, the time to produce p *en-face* images $t_{en-face}(MS)$ can, in theory, be reduced to the time to perform a single cross-correlation:

$$t_{en-face}(MS) = t_{acq} + 2t_{FFT} \quad (6)$$

As a consequence of the above analysis, it appears that it is possible to produce MS based *en-face* images at a rate competitive to that of the conventional FFT based SD-OCT technology. For systems where scanning is slow and t_{acq} large, there would be practically no difference in the time of producing *en-face* images between the SD-OCT and MS-OCT. If the time of producing an FFT, t_{FFT} is comparable to the time of extracting an *en-face* plane from the 3D volume, t_{render} , then even for fast acquisition systems, there would not be any time difference between the MS and the FFT based SD-OCT technologies in rendering *en-face* images. However, the MS-OCT gains in this comparison in those cases where data in SD-OCT need to be prepared before FFT. In such cases, the processing time exceeds the time for an FFT operation. Depending on the processing involved, which also depends on the nonlinearities to be corrected, it is therefore plausible to assume that $t_{en-face}(MS) < t_{en-face}(SD)$, i.e. the MS technique presents the potential of delivering *en-face* images faster than the FFT technique.

2.3. LabVIEW programs

Two LabVIEW projects were created to deal with the live operation of the imaging instrument assembled (LVApp1) and for post-processing of data (LVApp2). Both projects (description of their user interfaces is presented in the Appendix) include Dynamic Link Libraries (DDL1,2) to communicate with a GPU application (GPUApp). These programs control interfaces via three digital switches, K1, K2 and K3 (Fig. 1) to achieve three modes of operation: (i) Mask acquisition mode (MA), Live imaging mode (LI) and display of images in a post processing mode of operation (PP).

The first LabVIEW project, LVApp1, performs most of the tasks in the MA and LI modes.

In the Mask acquisition mode, MA, a flat mirror is used behind a lens (as an eye model). Electrical signal delivered by the balanced photo-detector, BPD, proportional to the shape of the channeled spectrum, is digitized and placed in the 'Mask Storage Block'. The block 'Control mask recording' sends commands to a translation stage that moves the support of the reference optical fiber, to alter the OPD, measured between the reference and sample arm lengths of the interferometer. For p values chosen for the OPD, the stored versions of channeled spectra shapes represent the p reference signals, or p masks.

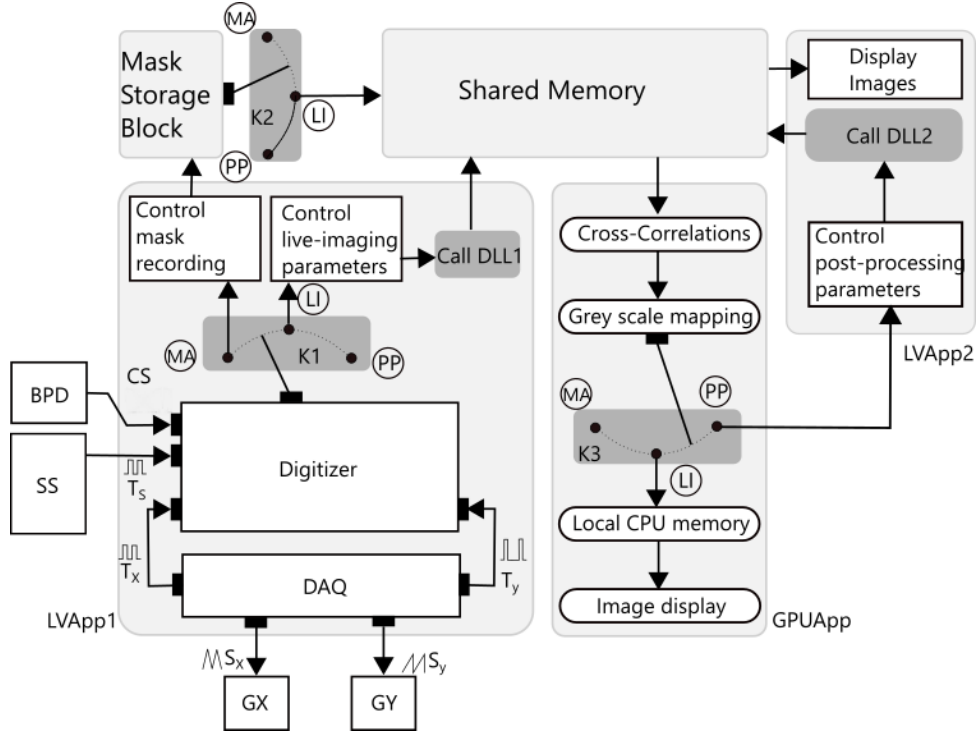


Figure 1: Schematic diagram demonstrating the operation of the MS-OCT instrument.

In the live imaging mode, LI, LVApp1 generates a triangular waveform via a data acquisition board (DAQ, NI PCI 6110) to drive the two galvo-scanners (GX and GY) performing the lateral scanning, with triangular (S_x) and sawtooth (S_y) waveforms respectively. LVApp1 acquires and stores the data using a two buffer mechanism via the digitizer according to timings set by the DAQ (TTL signals T_x and T_y associated to S_x and S_y respectively) and the trigger signal (T_s) provided by the swept source. LVApp1 communicates with the GPU application via the dynamic link library DLL1. Via switch K1, the LVApp1 exchanges buffered data with the GPUApp via DLL1, with data placed in the 'Shared Memory'. Along with the buffered data, a number of control values required to manipulate the contrast and brightness of the image are exchanged between LVApp1 and the GPUApp, which performs the signal processing and generates the resulting image using OpenGL drawing primitives. An additional important parameter is the value of the window W over which the correlation signal is integrated in Eq. (4). The control values to manipulate the contrast and brightness in the image and the parameter W are modified via the 'Control live-imaging parameters' block. LVApp1 is also used to produce an axial position guiding audio signal. From the buffered data to be transferred to the Shared Memory, an A-scan corresponding to a single channeled spectrum is produced, via a FFT. A sinusoidal audio signal is generated by the sound card of the computer with its amplitude proportional to the amplitude of the peak with the largest strength within the A-scan and with a frequency proportional to the OPD of such peak. This helps the user to position the imaged eye axially without the need to see the computer monitor.

In the post-processing mode, PP, the second LabVIEW application (LVApp2) is used. In this case, there is no data transfer between the digitizer and the Shared Memory, so the latest data buffer transferred to the shared memory is continuously called by the GPUApp via the second dynamic link library (DLL2). In this case, to easier manipulate the images, instead of using the OpenGL primitives, output data are routed back to LVApp2 for display, analysis and eventually recording. LVApp2 is used to explore the whole volume of 384 *en-face* OCT images captured, allowing similar versatility in 3D

presentation, reported so far for volumetric assemblies of A-scans obtained using the conventional FFT based SD-OCT.

A commonly available NVIDIA GeForce GTX 780 Ti (3 GiB on-board memory, 2880 CUDA cores) is employed, installed in a Dell Precision 7500 equipped with a dual CPUs (E5503, 4 cores @ 2.0 GHz).

SD-OCT (including MS-OCT) requires fast swept sources or linear cameras to generate sufficiently high resolution *en-face* images (in terms of the number of digital pixels) at decent rates (over 1 Hz). The swept source employed in this study has a tuning rate of 100 kHz. As a result, a 3D dataset required to build an *en-face* image of size $r \times q = 200 \times 200 \text{ pixels}^2$ can be acquired in 0.4 s. However, due to the limitations on the shape and frequency of the signal that can be sent to the galvo-scanners, the time to acquire data is $t_{\text{acq}} = 0.8 \text{ s}$. This was achieved in our case by driving the fast line galvo-scanner G_x with a triangular waveform of frequency $F_x = 250 \text{ Hz}$ and the slow frame scanner (G_y) with a sawtooth waveform at $F_y = 1.25 \text{ Hz}$. As a consequence, to ensure a real-time operation of the system, once data are acquired, the set of multiple images targeted have to be produced within $T_y = t_{\text{acq}} = 0.8 \text{ s}$.

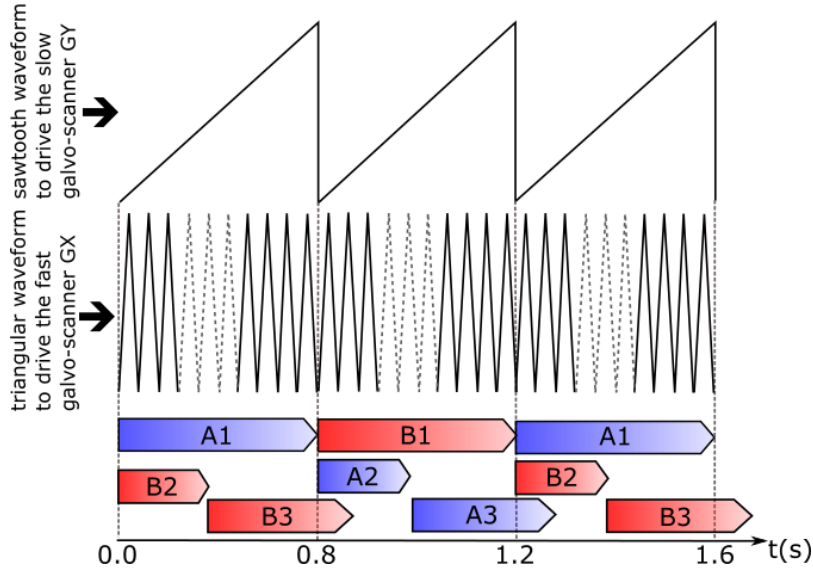


Figure 2: Flowchart showing the timing of data acquisition, transfer and processing (A3 and B3) in a GPU enhanced MS-OCT system. The timings shown on the figure ensure a real-time operation of the system for $r = q = 200$ pixels and $p = 40$ depths.

As G_y is driven with a sawtooth waveform in LVApp1, a two stage buffer mechanism was implemented to exchange data with GPUApp: while the buffer "A" is populated with new data (process A1 in Fig. 2), buffer "B" is transferred to the shared memory (process B2 in Fig. 2), emptied and ready to be populated with data. Then, while data are stored in "B" (process B1), the content of the other buffer, "A", is transferred to the shared memory (process A2) and once this operation is done, emptied.

To ensure a real-time operation two conditions have to be fulfilled simultaneously:

- (i) The transfer of data to the shared memory to be completed faster than T_y . For buffers of size $r \times q \times p = 200 \times 200 \times 1024 \text{ pixels}^3$ (163.84 MiB, as 4 bytes for each sampling point were allocated to avoid any eventual need of data format change and keeping maximum computational performance) the time to transfer data to the shared memory is evaluated as lower than 0.4 s.
- (ii) The time to process data on the GPU to be shorter than T_y . Obviously the data processing time depends on the number p of images to be generated.

In Fig. 3, the time to process data for different number p of *en-face* images to be produced is presented. As we aim to display both sets of *en-face* images (including an SLO image) (C-scans) and cross section images (B-scans), two cases are shown in Fig. 3:

B-scans computed (in red) or not (in blue). As it can be seen, the production of the B-scans alongside the *en-face* images has an important effect on the real-time operation of the system. This is expected: to produce the B-scan images, a number of $p = 384$ masks are used, hence a number of $r \times p = 200 \times 384 = 76,800$ cross-correlation operations are required to produce each cross-sectional image against only $r \times q = 200 \times 200 = 40,000$ operations to produce a single *en-face* view.

As presented in Fig. 3, when all the capabilities are enabled, i.e. producing SLO and B-scan images alongside *en-face* OCT images, real-time operation is possible when displaying up to 56 *en-face* OCT images.

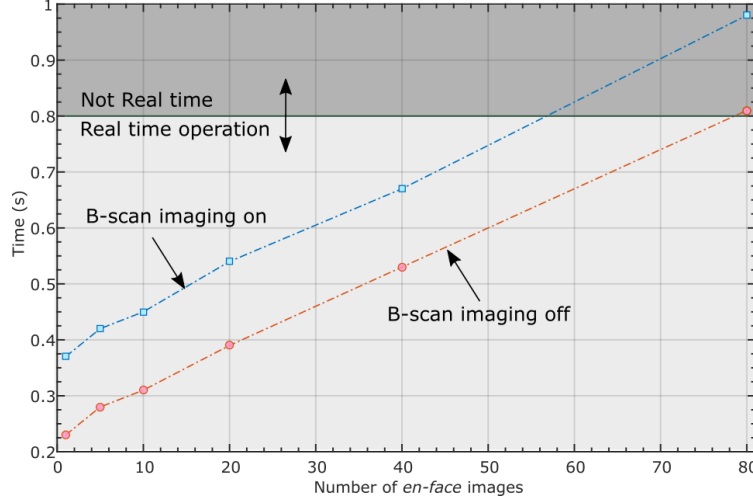


Figure 3: Processing time required vs. the number of *en-face* images to be produced. Red points: no B-scan images are computed; Blue points: B-scans computation turned on.

5. Results

A number of “ p ” channelled spectra are recorded initially in the MA mode by actuating on the position of a miniature linear translation stage TS (Model MFA-CC, Newport, Irvine, CA, Fig. 1 in [36]) and using a flat high reflective mirror in the eye model. The recording of masks was automatized through a LabVIEW software program using NI-VISA drivers to control the position of TS via the RS-232 serial communication port of the computer and the TS motion controller (Model MM4600, Newport, Irvine, CA). The process of recording masks is not time consuming (about 1s/mask) and is necessary to be performed once only.

While LVApp1 runs, the GPUApp produces using OpenGL primitives images such as those presented in Fig. 4. On the left, 40 *en-face* MS-OCT images (organized in 8 columns and 5 rows) are shown. The image at the left top corner is the closest image to OPD = 0. In the case presented in Fig. 4, its position was set at an axial depth $z = 1.575$ mm. The other *en-face* images are collected from depths separated by 21 μm , hence the last *en-face* image is collected from $z = 2.415$ mm (all distances measured in air). At the top right of Fig. 4 a SLO lookalike image is produced by averaging the 40 *en-face* MS-OCT images displayed. Its thickness depends on the number of OCT images superposed and the differential distance between them. To help positioning the retina in the correct position, B-scans in two orthogonal planes (XZ and YZ) are also generated. Their positions, controlled via the block CXYZ of LVApp1, is marked over the SLO image by green and red lines. Also, to conveniently locate the axial positions of the *en-face* images, short green and red lines placed on the right side of the B-scans are displayed. The B-scan images are formed from $p = 384$ T-scans (p masks) separated by 7 μm , hence the B-scan axial range is about 2.7 mm (in air).

The screen presented in Fig. 4 was extracted from a movie (Media 1) that demonstrates the versatility in producing *en-face* images by modifying the axial positions where the images originate from, as well as the differential axial distance between them. At the bottom of the frames that Media 1 is built from, the axial position of the first *en-face* image with respect to OPD = 0 is also specified.

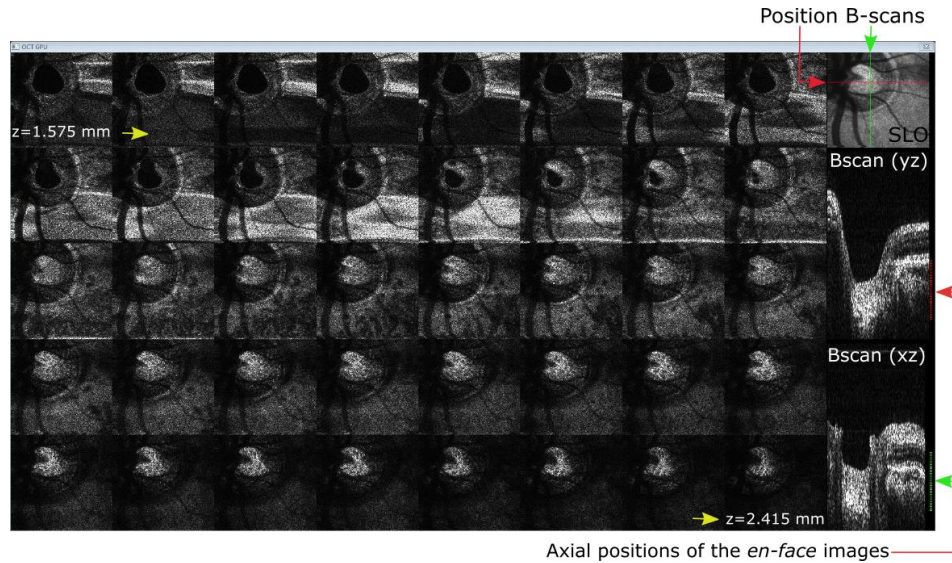


Figure 4: The output of GPUapp, in live mode showing 40 *en-face* images (left) separated axially by 21 μm (measured in air), a SLO looklike image (top right) as well as two orthogonal MS based B-scan images used for eye guidance and depth localization of the *en-face* planes. All images are updated at the frame rate. The images were obtained by imaging the optical nerve area of a healthy volunteer. The size of the *en-face* images is $3.0 \times 3.0 \text{ mm}^2$, while the size of the B-scans is 3.0 (lateral) \times 2.7 (depth) mm^2 .

Media 1 is made by repeating the display for the same data collected in one scanned frame, while changing the different settings and cursors for the same data, to illustrate the capabilities of the system. The 25 frames of the movie are repeated at a frame rate close to the driving frequency of the slow galvo-scanner (1.25 Hz), as in LI mode. No fixation aids were used during data acquisition apart from the audio guidance, and therefore it is expected for images to be distorted by eye movement. This is why in Media 1, we repeated the same frame, to avoid the eye movements confuse the effect of changes of different cursors and settings. In a real-time case scenario, the second movie (Media 2) was generated, where axial and lateral movements of the eye are obvious, but still, at 1.25 Hz decent quality images are produced. Media 2 was made for an axial differential distance between adjacent *en-face* images of 14 μm .

Once the operator is satisfied with images displayed per each frame, the switches in Fig. 1 are set to the Post processing mode, PP, in which case the flow of data to the 'Shared Memory' is interrupted, and the second application, LVApp2 is used now to send queries to the GPU application.

In a third multimedia file (Media 3) the utilization of LVApp2 is demonstrated: the desired positions of the B-scans is first set by adjusting the positions of the sliders 'CBX' and 'CBY', then scrolling through the *en-face* views is demonstrated by acting on 'SEF' (slider to control the depth position of the *en-face* image to be displayed in the bottom right corner). A display similar to that shown in Fig. 8 (Appendix) is also achievable by the conventional SD-OCT, however here, the volume is made of *en-face* OCT images and not of cross section OCT images. The SLO looklike image is made from superposition of *en-face* OCT images whilst the cross section OCT images are made from T-scans, like in time domain *en-face* OCT [4-6], i.e. all images, the *en-face* OCT, the SLO and the cross sections are obtained by applying the MS algorithm. This confer all images, independence to resampling accuracy, that affect the conventional FFT based SD-OCT algorithm.

In all multimedia files, the *en-face* images were collected from equally spaced axial positions, and have all the same axial resolution and sensitivity, while the SLO lookalike images were created from the 40 *en-face* slices. However, if required, these parameters can be easily modified via the two LabVIEW programs.

The individual integration window W employed in the correlation calculation can be adjusted, that affects the individual axial resolution and sensitivity of the MS-OCT *en-face* view [35]. *En-face* images from specific axial positions can be chosen, out of the depths of the 384 masks recorded in the preparation stage, as indicated by the two arrows (Axial positions of the *en-face* images).

A useful capability of the instrument is producing the SLO lookalike image simultaneously with the OCT images. This image is not as fragmented as the *en-face* OCT images, due to its larger thickness. SLO lookalike images with adjustable axial thickness are obtained by modifying the number of *en-face* OCT images averaged. Figure 5 depicts 6 such images obtained by averaging 9, 18, 35, 70, 140 and 280 *en-face* images centered at an axial position ($z = 1.645$ mm). Images of thickness 0.0625, 0.125, 0.25, 0.5, 1.0 and 2.0 mm are so produced. As a SLO system generates a SLO image with an axial resolution of 100-200 μm [38], only the images in Fig. (a), (b) and (c) can be correctly labeled as equivalent SLO images.

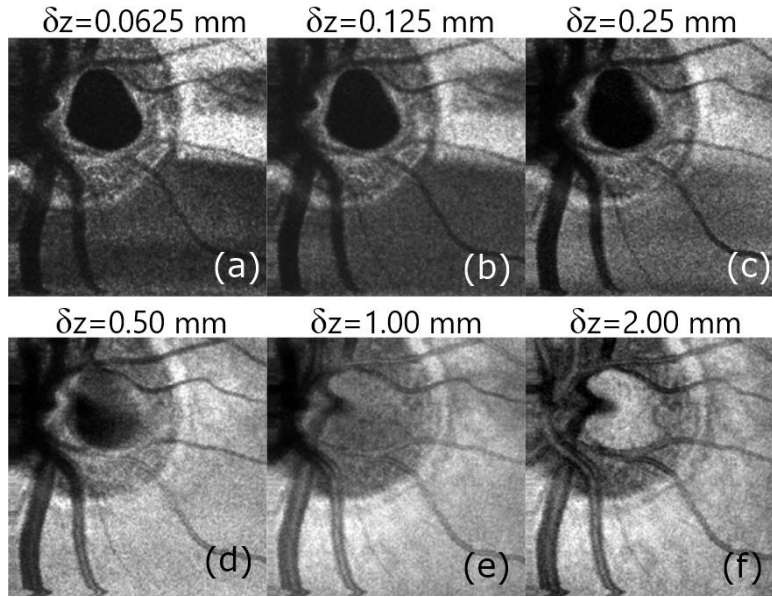


Figure 5: Averages of 9 (a), 18 (b), 35 (c), 70 (d), 140 (e) and 280 (f) *en-face* images (media 1, $p = 235$) with an equivalent thickness of 0.0625, 0.0125, 0.25, 0.5, 1.0, and 2.0 mm respectively at $z = 1.645$ mm. Size of the images is 3.0×3.0 mm².

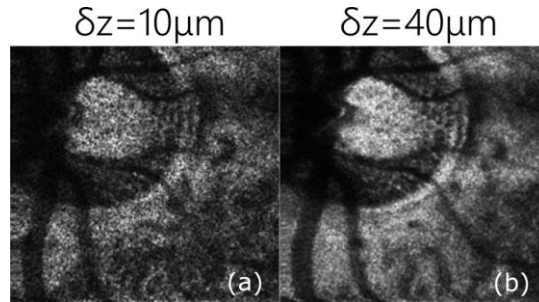


Fig 6: *En-face* MS-OCT based images of the optic nerve area of a healthy volunteer obtained by adjusting the size of the integration window of the correlation signal, from $W = 4$, which determines a depth resolution of ~ 10 μm (a) to $W = 2047$, which determines an axial resolution of ~ 40 μm (b). Size of the images is 3.0×3.0 mm².

As demonstrated in Refs. [35] and [37], the axial resolution in MS-OCT is adjustable by tweaking the value of the integration window W used to evaluate the cross-correlation of the two channeled spectra in Eq. (4). This is a distinctive MS-OCT capability. Increasing W is not equivalent to an average of images, i.e. this procedure does not slow down the acquisition or processing. Simply, the correlation calculation is averaged over more or less lag steps along the wavenumber coordinate. The larger the W , less is the axial resolution and better the sensitivity. According to Ref. [35], the axial resolution of the system can be tweaked from around 10 to about 40 μm by adjusting W . In Fig. 6, two images collected from the same axial position but showing different axial resolutions depending on the value of the integration window of the correlation signal are shown. To produce Fig. 6(a), the value of W was set at 4 (depth resolution around 10 μm) while to generate Fig. 6(b), the value of W was increased to the full range of the correlation signal (2047), which gives an axial resolution of about 40 μm when using the clock from the Axsun swept source. The clock is not necessary for the MS method, however to ensure (verify) that images of the same axial resolution and sensitivity are produced by both MS and FFT based SD-OCT method using the clock, we kept the k-clock enabled. This had some effect on the axial resolution range achievable by controlling W . When not using the clock, for the same W value, the axial resolution achieved was larger than 40 μm . Obviously, thicker *en-face* OCT images can be generated by averaging over a number of neighboring C-scans as illustrated in Fig. 5.

6. Discussion

In general, the 384 images calculated here can be interpreted as result of 384 independent flying spot *en-face* time domain OCT systems, each of them selecting signal from a distinct depth, where each such system exhibits the high sensitivity typical to the SD-OCT principle.

As demonstrated in [39], the MS technique is immune to the dispersion in the interferometer. This allows the developer to focus on compensating the dispersion due to the sample only. Different eyes exhibit different eye lengths and therefore different dispersion. For the images presented in the manuscript, the masks (proportional to shapes of channeled spectra) were recorded using an eye-model made of a lens and a flat mirror. However, experimentally, the Mask acquisition mode can be enhanced to tailor the masks for different dispersion as determined by different eye lengths. A water cuvette can be added in the object arm of the interferometer when the masks are recorded. Then, masks can be recorded for several lengths of the water cuvette. This has to be done once only and a set of 384 masks for each cuvette length requires only 590 kB, so there is no difficulty in storing several sets of masks. As masks are recorded for the dispersion imprinted by the different length of the water paths, the dispersion introduced by the eye itself can be minimized.

More studies are required in terms of the option of producing variable thickness SLO looklike images as well as on displaying *en-face* OCT images. A limited axial range is achieved via increasing the window W (for a given mask), while a wider axial range is available by averaging *en-face* images from several depths (i.e. by using several masks).

As the speed of computing a single FFT of the channeled spectrum signal is typically faster than that of producing an *en-face* cut from a 3D volumetric image, according to Eqs. (2) and (6) the generation of the *en-face* images via the MS-OCT can be faster than via the FFT based SD-OCT.

If re-sampling of data is needed prior to FFT processing, or numerical calculations are required to eliminate the dispersion mismatch between the arms of the interferometer, then the advantage of using MS-OCT method is even more obvious. The *en-face* OCT images obtained using the MS-OCT and using the FFT based OCT with enabled clock, look exactly the same [32]. If let us say that a swept source with no clock is employed, then a software linearization calibration procedure needs to be implemented in case the FFT based method is used. The time to re-sample data via a cubic spline interpolation is longer than the time to perform a FFT operation, implemented here via the CuFFT library, by at least 50% irrespective of the performance of GPU employed or of the amount of

data to be processed [40-42]. The procedure of re-sampling data not only requires time but the correction is not perfect for all depths as imperfections in linearization/calibration are unavoidable. Images using the FFT based procedure are not shown here, as they have been already reported by many other groups.

As another difference from the conventional SD-OCT, the depth positions of the *en-face* images, as well as the values of depth along the axial coordinate in the cross section images do not come from processing of the signal. The axial coordinate in the conventional SD-OCT method is the result of signal processing (FFT and linearization) and it is known that the linearization procedure moves the A-scan peaks along the depth coordinate. The depth position of images in the Master/Slave OCT is determined by the OPD measurement when acquiring the masks. The accuracy of depth positions is determined by the encoder resolution of the translation stage used to adjust the OPD in the MA mode, when the masks are collected. To produce *en-face* images of larger size, faster SSs or faster linear camera (already available) have to be used. This requires better GPUs, with more memory to be employed as well as computers with faster interconnects that can reduce the time to transfer data between different memories, improving both latency (delay in data being available to the GPU application) and visualization throughput (at higher frame rates).

While the MS method exhibits similar sensitivity and axial resolution to that provided by the FFT based technique, for the particular GPU used in this work, the production of several *en-face* retinal images simultaneously in real-time is possible for up to 56. This number can be increased up to 80 if the B-scan cross-sectional images generation is switched off.

Such a display of many *en-face* OCT images was not possible till now, therefore we can only speculate of possible applications. We envisage possible interest in real time monitoring of ablation or surgery in cases where details are not important to follow, but to see how different contours scale with depth. In terms of fundus imaging, superficial *en-face* aspect can be compared with corresponding deeper *en-face* views, along corresponding transversal pixels. Other possible applications may arise with use.

The size of the *en-face* images (in number of lateral pixels) is limited by the sweeping speed of the SS employed in this paper, hence a real-time production of the images was done at a frame rate of 1.25 Hz.

The time to produce MS based lookalike SLO images and multiple depths *en-face* OCT images demonstrated here when considered in combination with other advantages in terms of hardware cost, makes the MS-OCT method worth considering for imaging the eye. As there is no need for data re-sampling, MS-OCT can operate with a simpler swept source, not equipped with a k-clock, or even with potentially highly nonlinear tunable lasers. A MS-OCT system can operate in terms of its axial resolution and sensitivity decay with depth at the level of a perfectly corrected (linearized and calibrated) FFT based SD-OCT set-up. Therefore, MS-OCT presents potential in providing better sensitivity, resolution and axial range than its FFT based counterpart. With speeds in displaying *en-face* views faster than those typically reported by SD-OCT systems, MS-OCT can become the technique of choice for true real-time *en-face* imaging in ophthalmology. However, if a fast B-scan only is needed, FT based method will continue to be the method of choice. Therefore, there is a need for GPU systems to be equipped with both signal procedures, FFT and MS to be applied optimally, depending on the application.

7. Conclusions

A quasi real-time MS based OCT/SLO instrument is demonstrated, by harnessing the GPUs capabilities to employ the potential of the MS technology in terms of its parallelization. This paper describes for the first time how this exquisite feature can be employed to generate multiple *en-face* views. For parallel processing of signals from 384 depths, CUDA programs interleaved with LabVIEW programs have been created to perform the entire signal processing load that the MS method requires. Another feature presented here is the capability of assembly of a SLO lookalike image, as an immediate

by-product of parallel processing that leads to real time generation of 40 images. The number of images is adjustable up to 40 as well as the differential distance between them that allows the generation of an adjustable thickness fundus image (and if the thickness of the compound image is in the range 100 – 200 μm , the image is close to what a SLO instrument would deliver). As yet another feature presented here is that of MS generation of two B-scan images simultaneously with the display of several *en-face* OCT images. This could only be made possible by speeding up the process of B-scan generation via the MS method, as described in [37], again made possible by parallel processing of signals from the different depths of the T-scans assembled together. The result is a system where each of the generated *en-face* images can be tweaked in terms of their sensitivity, axial resolution and axial position.

7. Appendix: LabVIEW user interfaces

The interface of the live LabVIEW application LVApp1 is presented in Fig. 7. As it can be noticed, LVApp1 takes care of all aspects related to the acquisition of data: driving the galvo-scanners (block SGS in Fig. 7), generating the audio signal for eye guidance (block MDT), setting the correct parameters of the digitizer (block SDP) and controlling the brightness and contrast of the OpenGL images generated (blocks CXYZ and RS). Also, by adjusting the cut-off frequency of the correlation signal and its integration range (window W) [36-37], the axial resolution and the sensitivity of the *en-face* images can be adjusted. The digital switch 'DS' waits for the last data buffer to be successfully transferred to the 'Shared Memory' and interrupts the flow of data between the digitizer and the shared memory.

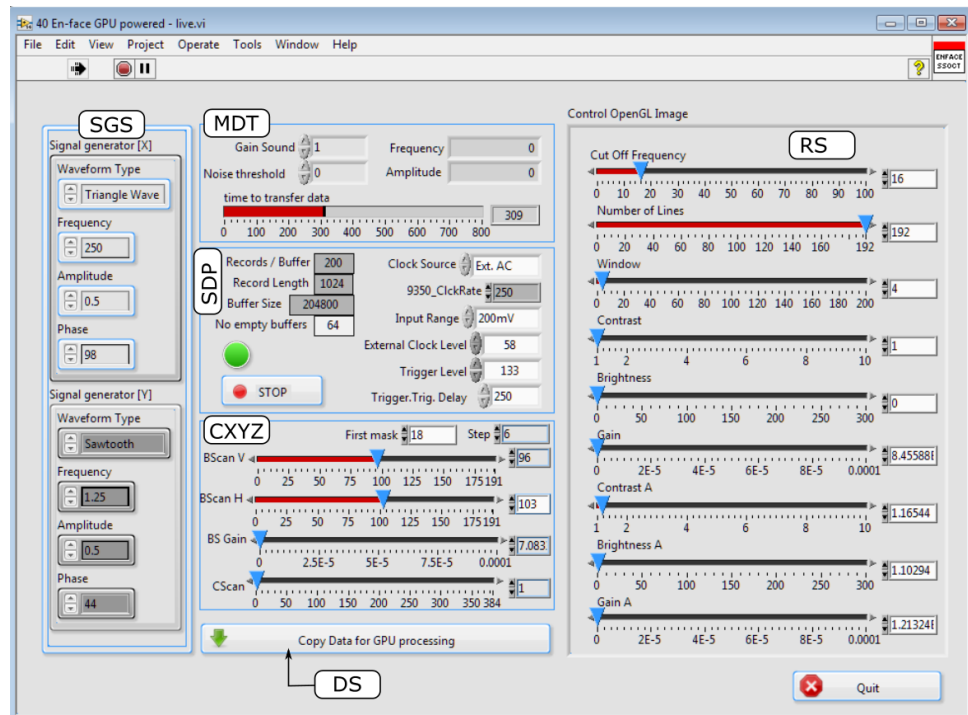


Figure 7: LVApp1 interface in LI mode.

In Fig. 8, the visual interface of LVApp2 is presented. Four images are produced: SLO lookalike (top left), orthogonal B-scans (top right and left bottom) and *en-face* (bottom right). Block 'RS' controls the parameters to be used to perform the cross-correlations, as well as the brightness/contrast of the *en-face* images, block 'SBF' controls the gain on the B-scans and the axial positions of the *en-face* views and the axial distance between them (used to produce the SLO view), 'CSLO' controls the brightness/contrast of the SLO image, 'CRS' is used to save the whole stack of *en-face* images, the sliders

‘CBX’ and ‘CBY’ are used to control the positions of the two B-scans to be displayed, while a third slider, ‘SEF’ is used to scroll through the stack of *en-face* images (the axial distance between consecutive *en-face* images is equal to the axial distance between the recorded masks, in our case 7 μm).

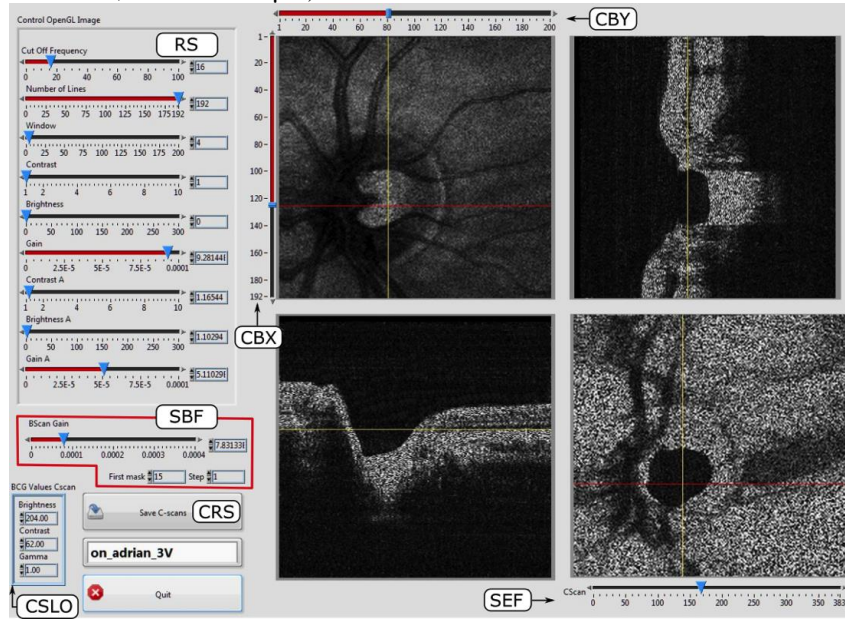


Figure 8: LVApp2 interface.

Most of the controls are not changed during the measurements. For example, in Fig. 7), there is no need to modify the block SGS that controls the galvos or SDP which sets the parameters for the digitizer. Once the brightness and the contrast in the images are satisfactory (they do not really need to be modified from a “patient” to another). Only 3 controls are essential: depth of the first *en-face* images (control “First mask”, in the middle), distance between consecutive masks (“Step”, in the middle) and eventually the size of the window (“Window”, right) if changes in axial resolution are required.

Similarly, in Fig. 8, scrolling through the stack of images is made possible by actuating on either of the bar controls CBX, CBY or SEF only. The values of the other controls are not typically modified (they are initially set in LVApp1).

In comparison to the images shown in Fig. 4, for a better visualization, in Fig. 8, the lateral size of the *en-face* and B-scans was scaled up by a factor of 2, so each displayed images has a 400×400 pixels² (*en-face*) and 400×384 pixels² (B-scans).

Therefore, the B-scans images shown in Fig. 8 (and also in Media 3) may look pixelated. The process of re-scaling is done in LabVIEW, where a tradeoff between the interpolation method used to rescale the images which provide their quality and the speed of producing them has to be established. For further analysis, the user records *en-face* images of good quality (original size, 200×200 pixels²) similar to those demonstrated in Fig. 4. The MS technique is capable of producing B-scan images of equal quality to those generated by its FFT based counterpart, not shown here but demonstrated in a recent report [43].

Acknowledgements

K. Kapinchev and F. Barnes acknowledge the support of the School of Computing at the University of Kent. K. Kapinchev also acknowledges the support of EPSRC DTA funding. A. Bradu and A. Podoleanu acknowledge the support of the European Research Council (ERC) (<http://erc.europa.eu>) COGATIMABIO 249889. A. Podoleanu is also supported by the NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and the UCL Institute of Ophthalmology.