We present a filtering procedure based on singular value decomposition to remove artifacts arising from sample motion during dynamic full field OCT acquisitions. The presented method succeeded in removing artifacts created by environmental noise from data acquired in a clinical setting, including in vivo data. Moreover, we report on a new method based on using the cumulative sum to compute dynamic images from raw signals, leading to a higher signal to noise ratio, and thus enabling dynamic imaging deeper in tissues.

1 Introduction

Optical coherence tomography (OCT) is routinely used for 3D imaging of microstructures in tissue and relies on the endogenous backscattering contrast [1, 2]. Full-field optical coherence tomography (FFOCT) is an en face, high transverse resolution version of OCT [3, 4]. Using a camera and an incoherent light source, FFOCT acquires en face virtual sections of the sample at a given depth and has been used for biology [5] and medicine [6]. Recently, a novel contrast mechanism has been exploited by measuring temporal fluctuations of the backscattered light in a technique called dynamic full field OCT (DFFOCT) [7]. In ex vivo fresh tissues, these dynamic measurements reveal subcellular structures that are very weak backscatters and provide contrast based on their intracellular motility [8, 9]. The penetration depth of DFFOCT is an order of magnitude lower than FFOCT due to the small cross-section of the scatterers, limiting its use in thick samples. Up to now, using DFFOCT for in vivo imaging has remain elusive as this technique is sensitive to nanometric axial displacements of the sample. In this paper we propose two methods to overcome the previous mentioned limitations. First, we introduce a framework based on the singular value decomposition (SVD) to filter out the axial displacement of the sample from the local fluctuations linked to cell motility, enabling in vivo use of DFFOCT. Dynamic measurements allow imaging with a different contrast and have been successfully used previously to reveal ganglion cells in the human retina [10]. SVD based algorithms have previously been applied for OCT data, e.g. for smart OCT where the SVD is applied on the reflection matrix in order to extend the penetration depth [11]. A SVD filtering method for DFFOCT has been previously proposed for simulated data [12] but does not work on our experimental data mainly because the image formation model is different. Also, similar SVD based algorithms for spatio-temporal filtering have been used effectively in acoustics for Doppler acquisitions [13, 14]. Secondly, we present a new operator to compute the local dynamics based on the cumulative sum in order to enhance the non-stationary part of the signal, leading to a strong increase of the signal to noise ratio (SNR). Finally, we report on the first DFFOCT acquisition in vivo to image the mouse liver at 80 µm depth where the two proposed algorithms greatly improved the image quality by removing motion artifacts and increasing the SNR by a factor of 3.

*https://www.jscholler.com
2 Removing artifacts using SVD

In order to construct a dynamic image, a stack of typically 512 direct images \((1440 \times 1440\) pixels) is acquired with a standard FFOCT setup, which consists of a Linnik interferometer where both arms contain identical microscope objectives conjugated to the camera plane, as shown in Fig. 1(a), and where each pixel is processed independently. In the first report on DFFOCT, the level of the dynamic signal at each pixel was computed using a running standard deviation averaged over the whole acquisition \([7]\). Calculating the standard deviation of the signal in time removes highly scattering stationary structures such as collagen or myelin fibers and reveals cells with a much better contrast. Indeed, strongly backscattering structures can dominate the signal even outside the coherence volume thereby masking weakly scattering structures such as cells. In the laboratory, we succeeded in stabilizing the setup by mounting it on a sturdy optical bench, and carried out ex vivo experiments without motion artifacts. For real life applications however, DFFOCT devices are currently being used by clinicians in hospitals for imaging biopsied tissues \([15]\) and by anatomo-pathologists in busy environments with vibrations arising from vibrational modes of the building, from people walking around the device and from air conditioning. Mechanical vibrations can lead to sample arm motion or oscillations, creating strong signal fluctuations, especially from highly reflective structures such as collagen fibers.

2.1 Motion artifact model

In the FFOCT camera plane, the resulting electric field is the sum of the backscattered light from both the sample and the reflected light from the reference arm \([16]\):

\[
I(r, t) = \eta I_0 \left( R(r, t) + R_{inc} + R_{ref} + 2\sqrt{R(r, t)}\alpha\cos\left(\Delta\phi(r, t)\right)\right)
\]

where \(I(r, t)\) is the intensity recorded at position \(r = (x, y)\) and time \(t\), \(\eta\) is the camera quantum efficiency, \(I_0\) is the power LED output impinging the interferometer considering a 50/50 beam-splitter, \(R_{ref}\) is the reference mirror reflectivity (i.e. the power reflection coefficient), \(R(r, t)\) is the sample reflectivity (i.e. the power reflection coefficient) at position \(r\) and time \(t\), \(\Delta\phi(r, t)\) is the phase difference between the reference and sample back-scattered signals at position \(r\) and time \(t\), \(I_{incoh} = R_{inc}I_0/4\) is the incoherent light back-scattered by the sample on the camera, mainly due to multiple scattering and reflections out of the coherence volume. The processed dynamic signal can then be written \([16]\):

\[
I_{dyn}(r) = \frac{1}{N} \sum_i SD \left( \frac{\eta I_0}{2} \sqrt{R_s(r, t[i,i+\tau])R_{ref}f\cos\left(\Delta\phi_s(r, t[i,i+\tau])\right)} \right)
\]

where \(SD\) is the standard deviation operator, \(N\) is the total number of sub-windows, \(\tau\) is the sub-windows length so that \(t[i,i+\tau]\) is the time corresponding to one sub-window, \(R_s\) and \(\Delta\phi_s\) are respectively the reflectivity and phase of the
local scatterers that induce the temporal fluctuations that DF-OCT aims to measure. In the event of the entire sample small displacements, on the order of the depth of field or smaller, the processed signals will be the sum of the actual local fluctuations and the modulation created by the whole sample motion creating a global phase shift. The resulting artifacts are therefore proportional to the sample reflectivity, which is orders of magnitude higher than the reflectivity of the scatterers probed by DF-OCT (e.g. mitochondria and vesicles) leading to strong artifacts on the dynamic image, which mask the signal of interest. In the presence of mechanical noise, the measured fluctuation can be written:

$$I_{\text{mes}}(r) = I_{\text{dyn}}(r) + I_{\text{art}}(r)$$

where the artifactual signal can be expressed as:

$$I_{\text{art}}(r) = \frac{1}{N} \sum_{i} SD \left( \frac{\hbar I_0}{2} \sqrt{R(r, t_{i,i+\tau})} R_{\text{ref}} \cos \left( \frac{4\pi}{\lambda} z(t_{i,i+\tau}) \right) \right)$$

where $$z(t_{i,i+\tau})$$ is the sample axial displacement on the $$i-th$$ sub-window. Here we neglected the sample deformation for the sake of clarity. Nonetheless it could be taken into account by processing the stack in spatial patches where deformations are negligible. For a highly reflective zone we have $$I_{\text{dyn}}(r) \ll I_{\text{art}}(r)$$ and the dynamic signal is completely masked by artifacts that look like the corresponding static DF-OCT image that could be obtained by randomly sampling the path difference instead of using the standard PZT modulation mentioned before. Indeed, modulating the position of the sample around the coherence volume is equivalent to modulating the piezo position, which explains why artifacts look like the standard DF-OCT image.

### 2.2 Proposed algorithm

The first step is to unfold the 3D $$M(x, y, t)$$ cube of data into a 2D matrix $$M_u(r, t)$$ to perform the decomposition. Higher dimensions of SVD do exist but are not required here as the horizontal $$x$$ and vertical $$y$$ dimension do not differ when considering axial motion artifacts. SVD is the generalization of the eigendecomposition of a positive semidefinite normal matrix, and can be thought of as decomposing a matrix into a weighted, ordered sum of separable matrices which will become handy when reconstructing the denoised signals:

$$M_u = U \Sigma V^* = \sum_{i} \sigma_i U_i \otimes V_i$$

where $$U$$ contains the spatial eigenvectors, $$V$$ contains the temporal eigenvectors and $$\Sigma$$ contains the eigenvalues associated with spatial and temporal eigenvectors. Performing the SVD on an unfolded $$M_u (1440 \times 1440, 512)$$ dynamic stack takes around 30 seconds on a workstation computer (Intel i7-7820X CPU, 128Go of DDR4 2666 MHz RAM) and requires 45Go of available RAM using LAPACK routine for SVD computation without computing full matrices. Investigating such decomposition for artifact-free data sets we found that spatial eigenvectors related to motion artifacts have particular and easily identifiable associated temporal eigenvectors. Indeed, when looking at the first temporal eigenvectors, we observed sinusoid-like patterns with increasing frequency, see Fig. 2(a). In presence of motion artifacts, temporal eigenvectors appeared with random, high-frequency components that are easy to detect with simple features. Here, the zero crossing rate (the number of time a function crosses $$y = 0$$) is used to detect temporal eigenvectors involved in motion artifacts. In presence of motion artifacts some of the firsts temporal eigenvectors present a high zero crossing rate, see Fig. 2(b)(c). In order to detect these outliers we computed the absolute value derivative of the zero crossing rate.
crossing rate (D-ZRC) and applied a threshold: if the D-ZRC is higher than three times the D-ZRC standard deviation then the corresponding eigenvalue is set to zero in $\hat{\Sigma}$ and the denoised stack $\hat{M}_u$ is reconstructed as:

$$\hat{M}_u = \sum_i \hat{\sigma}_i U_i \otimes V_i$$  \hspace{1cm} (6)

The denoised stack $\hat{M}_u$ can then be folded back to its original 3D shape $\hat{M}$ and the dynamic computation can be performed. Interestingly, the use of an automatic selection of eigenvectors allow a more reproducible analysis. For example, the SVD can be performed on spatial sub-regions without visible artifacts, something very hard to obtain with manual selection of eigenvectors. This can also improve the filtering procedure in the case of sample spatial deformation or if the computation requires too much RAM.

2.3 Results

We tested the proposed SVD filtering on different acquisitions taken on different setups (the one presented in Fig. 1(a) and a commercial system manufactured by LLTech SAS). When motion artifacts were present, image quality after denoising was greatly improved in each case. In Fig. 3 we present lung biopsy images taken with the LLTech system in a clinic, where imaged tissues were waste tissues from biopsy procedures that were destined to be destroyed, and we imaged them just before destruction. The imaging was carried out according to the tenets of the Declaration of Helsinki and followed international ethic requirements for human tissues. SVD filtering effectively removes motion artifacts from collagen fibers and reveals cells in Fig. 3(b)(e). DFFOCT images were also higher contrast after SVD denoising, allowing easier interpretation for clinical applications, e.g. lung tumor detection in the presented images. We imaged fibroblasts with the setup presented in Fig. 1(a). Cells were very flat leading to fringes pattern created by the specular reflection on their surface. These fringes were highly visible on the processed DFFOCT image preventing the visualization of subcellular structures, see Fig. 4(a). After SVD filtering it is possible to distinguish single subcellular entities and track them, see Fig. 4(b), enabling biological study without the need of a costly optical bench setup.

3 Extending penetration depth using non-stationarities

A drawback of DFFOCT compared to standard static FFOCT is the reduced penetration depth. While FFOCT can acquire images as deep as $1 \text{ mm}$, DFFOCT is limited to about $100 \mu\text{m}$ due to the weak signal level produced by the sample fluctuations we wish to measure. In order to enhance the dynamic signal strength and so improve penetration depth, we propose computation of the dynamic image from the cumulative sum of the signal, rather than the raw signal. Indeed, the model for the dynamic image formation is small scatterers moving in the coherence volume during the acquisition leading to phase and amplitude fluctuations in the conjugated camera pixel. While a pure Brownian motion is stationary, hyper-diffusive displacements are not and we therefore propose to use the cumulative sum to enhance these non-stationarities. Intuitively, summing a centered noise will give a noisy trajectory that stays close to zero whereas if there is a small bias it will be summed for every sample and the cumulative sum will therefore have a slope equal to this bias.

3.1 Theoretical considerations

In the case of pure centered Gaussian noise, taking the cumulative sum gives a Brownian bridge. Theoretically, Brownian bridges are expected to be maximal close to the edges as the probability distribution of the maximum is the third arcsin law which has a typical U-shape [17]. If we consider a Brownian bridge $W_s \forall s \in [0, 1]$:  

$$W_M = \sup \{ W_s : s \in [0, 1] \}$$

(7)

$$P[M \leq x] = \frac{2}{\pi} \arcsin(\sqrt{x}), \quad \forall x \in [0, 1]$$

(8)

where $W_M$ is the maximum, $W_s$ are the Brownian bridge samples and $M$ is the maximum position. More importantly, the Brownian bridge maximum follows a Rayleigh distribution:

$$P[W_M \leq u] = 1 - e^{-2u^2}$$

(9)

According to the Rayleigh distribution, the Brownian bridge maximum must therefore scale as $\sqrt{t}$ with $t$ scaling as the number of frames. Now, if there is a bias in the distribution, which is the case if a scatterer is moving with constant velocity in the coherence volume, the cumulative sum will scale as $\frac{t}{2}$ due to the slope introduced by the bias. It will also be either always positive or negative and the maximum will be reached around the center of the bridge. The cumulative sum will therefore exhibit a completely different behavior for centered noise than for actual motility signals, leading to
Fig. 3. (a)(d) Original DFFOCT images computed on the raw stack. (b)(e) Denoised images computed with SVD filtering. (c)(f) Sum of the spatial eigenvectors absolute value removed by the SVD filtering. See Visualization 1 for complete $1260 \times 1260 \, \mu m$ fields of view.
Fig. 4. (a) Original DFFOCT image computed on the raw stack. (b) Denoised image computed with SVD filtering. Subcellular features appear with a much better contrast enabling segmentation and tracking. (c) Sum of the spatial eigenvectors absolute value removed by the SVD filtering.

a better signal to noise ratio on dynamic images. Note that for Brownian bridges it is often observed that the function changes sign regularly (the probability of sign changes is also well established [17]), which is not the case when there is non-stationarities.

We simulated an experiment by introducing a linear bias of $\sigma/3$ on a centered Gaussian distribution which is not perceptible on the presented signals Fig. 3(a). Looking at the cumulative sum the bias is much more obvious as the maximum reached by the bridge is three times higher, hence motility signals are detected with a higher sensitivity using the cumulative sum.

### 3.2 Results

The dynamic image is computed by taking the average of the maxima of the absolute values of the running cumulative sum:

$$I_{dyn}'(r) = \frac{1}{N} \sum_i \max \left(|\text{CumSum}(I(r, t[i,i+\tau]) - \bar{I}(r, t[i,i+\tau])|\right)$$  \hspace{1cm} (10)

where $\text{CumSum}$ is the cumulative sum operator, $N$ is the total number of sub-windows, $\tau$ is the sub-windows length so that $t[i,i+\tau]$ is the time corresponding to one sub-window and $\bar{I}(r, t[i,i+\tau])$ is the signal mean on the sub-window.

We tested the proposed method with $\tau = 50$ on the photoreceptor layer of an explanted macaque retina at 85 $\mu m$ depth, see Fig. 5(b)(c). In this case the SNR is twice better with the proposed method and the camera column noise is almost completely removed. We tested the proposed algorithm on several acquisitions on tissue and cell cultures and the average SNR improvement factor was 1.9, allowing to image deeper into tissues with DFFOCT.
Fig. 5. (a) Simulation of Gaussian noise with and without bias and their cumulative sum. The maximum is 3 times higher with the bias. (b) Dynamic image of a macaque photoreceptor layer using standard deviation. (c) Dynamic image of a macaque photoreceptor layer using the proposed method based on cumulative sum.

4 Applying both methods for in vivo dynamic imaging

The proposed SVD filtering procedure is of great interest for applying DFFOCT in vivo. A custom head was adapted on the sample arm of an LLTech setup combined with a pump in order to create a weak suction force to immobilize the sample. We then acquired a stack of images on a living mouse liver at 80 µm depth. The animal manipulation protocol was approved by our local animal care committee. The mouse (4 week-old C57BL/6 (Janvier Lab, Le Genest Saint Isle, France)), was anesthetized by isoflurane, and sacrificed after the imaging procedure by CO2 inhalation. The standard DFFOCT images were very noisy mainly due to the heartbeat and breathing of the mouse, leading to tissue motion that creates artifacts, see Fig. 6(a). On applying the proposed SVD filtering, we were able to remove motion artifacts Fig. 6(b). Nonetheless signals are still very low due to the deep imaging in a strongly scattering organ and applying the cumulative sum algorithm dramatically increased the SNR by a factor of 3 Fig. 6(c). In the end, there are remaining artifacts produced by the coherence volume axial drift during the acquisition. Indeed, if the coherence volume shifts more than its axial extension, even if motion artifacts are perfectly removed, the probed dynamics would be averaged over several depths leading to an axial blur. To overcome this issue, the position of the coherence gate inside the sample may be compensated by monitoring the breathing and moving the reference arm with a precision corresponding to the optical sectioning (10 µm for the in vivo acquisition here) in order to compensate for the axial drift.

5 Conclusion

We proposed a filtering algorithm based on the SVD to effectively remove motion artifacts from dynamic images. The proposed method adds ~ 40 seconds for a (1440, 1440, 512) stack which will require GPU processing in order to speed up the process for real time applications. This method was applied on an in vivo data set and is promising as long as axial motion to be smaller than the coherence volume depth. Tracking and compensating methods are currently being investigated in order to acquire DFFOCT stacks in a completely artifact-free manner for cornea [18] and retina [19]. We also proposed a method based on the cumulative sum to enhance non-stationarities in temporal signals which led to an SNR factor increase of 1.9 on average for ex vivo samples and 3 on our in vivo data set. These general techniques could be applied to any other imaging modality with sub-diffraction phase sensitivity.
**Fig. 6.** (a) In vivo DFFOCT image of a mouse liver computed on the raw stack with the standard deviation. (b) In vivo DFFOCT image of a mouse liver computed on the denoised stack with the standard deviation. (c) In vivo DFFOCT image of a mouse liver computed on the denoised stack with the cumulative sum. See Visualization 2 for complete 1260 × 1260 μm fields of view.

**Funding Information**

HELMHOLTZ grant, European Research Council (ERC) (610110).

**Acknowledgments**

The author would like to thank LLTech SAS for sharing its raw data, especially Emilie Benoit and Louis Dutheil for carrying out in vivo and clinical experiments. The author is also grateful to Olivier Thouvenin, Pedro Méce, Kassandra Groux, Mathias Fink, Claude Boccara and Kate Grieve for fruitful discussions and valuable comments regarding this paper.

**References**

[1] D Huang, EA Swanson, CP Lin, JS Schuman, WG Stinson, W Chang, MR Hee, T Flotte, K Gregory, CA Puliafito, et al. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.

[2] Wolfgang Drexler and James G. Fujimoto, editors. *Optical coherence tomography: technology and applications*. Springer, 2 edition, 2015.

[3] E. Beaurepaire, A. C. Boccara, M. Lebec, L. Blanchot, and H. Saint-Jalmes. Full-field optical coherence microscopy. *Opt. Lett.*, 23(4):244–246, Feb 1998.

[4] Arnaud Dubois, Kate Grieve, Gael Moneron, Romain Lecaque, Laurent Vabre, and Claude Boccara. Ultrahigh-resolution full-field optical coherence tomography. *Appl. Opt.*, 43(14):2874–2883, May 2004.

[5] Juliette Ben Arous, Jonas Binding, Jean-Francois Léger, Mariano Casado, Piotr Topilko, Sylvain Gigan, A. Claude Boccara, and Laurent Bourdieu. Single myelin fiber imaging in living rodents without labeling by deep optical coherence microscopy. *Journal of Biomedical Optics*, 16(11):116012, November 2011.
[6] Manu. Jain, Nidhi. Shukla, Maryem. Manzoor, Sylvie. Nadolny, and Sushmita. Mukherjee. Modified full-field optical coherence tomography: A novel tool for rapid histology of tissues. *Journal of Pathology Informatics*, 2(1):28, 2011.

[7] Clement Apelian, Fabrice Harms, Olivier Thouvenin, and A. Claude Boccara. Dynamic full field optical coherence tomography: subcellular metabolic contrast revealed in tissues by interferometric signals temporal analysis. *Biomed. Opt. Express*, 7(4):1511–1524, Apr 2016.

[8] Charles-Edouard Leroux, Fabien Bertillot, Olivier Thouvenin, and Albert-Claude Boccara. Intracellular dynamics measurements with full field optical coherence tomography suggest hindering effect of actomyosin contractility on organelle transport. *Biomed. Opt. Express*, 7(11):4501–4513, Nov 2016.

[9] Olivier Thouvenin, Claude Boccara, Mathias Fink, José Sahel, Michel Pâques, and Kate Grieve. Cell motility as contrast agent in retinal explant imaging with full-field optical coherence tomography. *Investigative Ophthalmology & Visual Science*, 58:4605, 09 2017.

[10] Zhuolin Liu, Kazuhiro Kurokawa, Furu Zhang, John J. Lee, and Donald T. Miller. Imaging and quantifying ganglion cells and other transparent neurons in the living human retina. *Proceedings of the National Academy of Sciences*, 114(48):12803–12808, 2017.

[11] Amaury Badon, Dayan Li, Geoffroy Lerosey, A. Claude Boccara, Mathias Fink, and Alexandre Aubry. Smart optical coherence tomography for ultra-deep imaging through highly scattering media. *Science Advances*, 2(11), 2016.

[12] H. Ammari, F. Romero, and C. Shi. A signal separation technique for sub-cellular imaging using dynamic optical coherence tomography. *Multiscale Modeling & Simulation*, 15(3):1155–1175, 2017.

[13] C. Demené, T. Deffieux, M. Pernot, B. Osmandski, V. Biran, J. Gennisson, L. Sieu, A. Bergel, S. Franqui, J. Correas, I. Cohen, O. Baud, and M. Tanter. Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and ultrasound sensitivity. *IEEE Transactions on Medical Imaging*, 34(11):2271–2285, Nov 2015.

[14] J. Baranger, B. Arnal, F. Perren, O. Baud, M. Tanter, and C. Demené. Adaptive spatiotemporal svd clutter filtering for ultrafast doppler imaging using similarity of spatial singular vectors. *IEEE Transactions on Medical Imaging*, 37(7):1574–1586, July 2018.

[15] Emilie Benoit a la Guillaume, Clement Apelian, Eugénie Dalimier, Claude Boccara, Audrey Mansuet-Lupo, Guillaume Chassagnon, and Marie-Pierre Revel. Lung biopsy assessment with dynamic cell optical imaging, 2018.

[16] Jules Scholler, Viacheslav Mazlin, Olivier Thouvenin, Kassandra Groux, Peng Xiao, José-Alain Sahel, Mathias Fink, Claude Boccara, and Kate Grieve. Probing dynamic processes in the eye at multiple spatial and temporal scales with multimodal full field oct. *Biomed. Opt. Express*, 10(2):731–746, Feb 2019.

[17] Paul Lévy. Sur certains processus stochastiques homogènes. *Compositio Mathematica*, 7:283–339, 1940.

[18] Viacheslav Mazlin, Peng Xiao, Eugénie Dalimier, Kate Grieve, Kristina Irsch, José-Alain Sahel, Mathias Fink, and A. Claude Boccara. In vivo high resolution human corneal imaging using full-field optical coherence tomography. *Biomed. Opt. Express*, 9(2):557–568, Feb 2018.

[19] Peng Xiao, Viacheslav Mazlin, Kate Grieve, Jose-Alain Sahel, Mathias Fink, and A. Claude Boccara. In vivo high-resolution human retinal imaging with wavefront-correctionless full-field oct. *Optica*, 5(4):409–412, Apr 2018.