Dynamic optical coherence tomography for cell analysis [Invited]

**Salvatore Azzollini,**1 **Tual Monfort,**1,2 **Olivier Thouvenin,**3 and **Kate Grieve**1,2,*

1 Sorbonne Université, INSERM, CNRS, Institut de la Vision, 17 rue Moreau, F-75012 Paris, France
2 CHNO des Quinze-Vingts, INSERM-DGOS CIC 1423, 28 rue de Charenton, F-75012 Paris, France
3 Institut Langevin, ESPCI, PSL, 1 rue Jussieu 75005 Paris, France

* kate.grieve@inserm.fr

**Abstract:** Label-free live optical imaging of dynamic cellular and subcellular features has been made possible in recent years thanks to the advances made in optical imaging techniques, including dynamic optical coherence tomography (D-OCT) methods. These techniques analyze the temporal fluctuations of an optical signal associated with the active movements of intracellular organelles to obtain an ensemble metric recapitulating the motility and metabolic state of cells. They hence enable visualization of cells within compact, static environments and evaluate their physiology. These emerging microscopies show promise, in particular for the three-dimensional evaluation of live tissue samples such as freshly excised biopsies and 3D cell cultures. In this review, we compare the various techniques used for dynamic OCT. We give an overview of the range of applications currently being explored and discuss the future outlook and opportunities for the field.

© 2023 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Background

Optical coherence tomography [1] (OCT) is often described as the optical equivalent of ultrasound imaging, as it relies on the measurements of light echoes, to investigate layered tissues noninvasively. These measurements are achieved using optical interferometry since no electronic device is capable of temporally resolving echoes arriving from structures spaced only a few microns apart. For instance, a Michelson interferometer can be employed, in which the beam from a low temporal coherence light source is split in two parts: one part goes to the sample arm where it is back-scattered from the internal sample structures located at different depths, whereas the other part is reflected by a reference mirror, placed at a distance that matches the propagation times of both beams, so that they can interfere once recombined [1]. The amount of back-scattered light depends on the differences in the refractive index distribution from a confined location inside the sample. Thanks to the low coherence, OCT only detects the light backscattered from a coherence plane whose depth location is chosen by adjusting the length of the reference arm, leading to a sectioning ability which depends on the bandwidth of the source, and is hence decoupled from the lateral resolution. OCT axial resolution can therefore exceed the theoretical axial diffraction limit constraining widefield and confocal microscopies [2]. Thanks to these attributes, OCT has been extensively used since its invention, especially for medical diagnosis, where a large field of view but also high-resolution imaging is desirable. So far, OCT has mostly been used to provide morphological information based on backscattered light coming from structures with different refractive indices inside the tissues, such as the observation of retinal layers or the organization of the extracellular matrix [1]. Nonetheless, thanks to its imaging speed and sensitivity capabilities, longitudinal OCT imaging enables quantification of how a large variety of biological structures move, grow, or remodel over time, potentially adding a higher level of characterization that can be used to increase OCT specificity [3,4].
Taking advantage of this temporal behavior, at various time scales, has allowed researchers to have a deeper insight into cellular and subcellular mechanisms, progressively leading to the establishment of a functional non-invasive label-free imaging technique called dynamic OCT (D-OCT). The core idea is to collect the signal coming from the same imaging area over time in order to generate a series of frames that will contain the temporal evolution of the interferogram. Rapid temporal (0.1 - 20Hz) fluctuations of the OCT signal in live samples typically rely on active transport inside cells, hence requiring live cells [5,6] displaying ATP production [7] and functional cytoskeleton organization [8]. In some cells, including RPE cells, it was demonstrated that D-OCT was mostly driven by mitochondrial transport [9].

As confirmed by correlative histology and/or fluorescence microscopy, D-OCT enables visualization of specific cells [10,11] and their nuclei, permitting the identification of cell mitotic state [12,13], while temporal fluctuations can be boosted by enhanced cell motility or membrane fluctuations, as in leukocytes, lymphocytes, and red blood cells [10,11], hence demonstrating the potential to specifically identify some cell types. As a result, thanks to the wide variety of structures it can identify, D-OCT circumvents, in some cases, the need for fluorescent dyes; overcoming all the drawbacks related to photo-toxicity and biases introduced by the fluorescent markers [14] besides being completely independent from sample preparation procedures [15].

If temporal behavior appears to be an indispensable aspect to encompass the whole biological picture, biological time scales span fractions of milliseconds to several days, and the associated temporal analysis, as well as the technical solutions, differ greatly depending on what is imaged. As a result, over the years many groups have developed their own ways to analyze temporal trends of the interferometric signal produced by their different OCT setups, different signal processing pipelines and different samples. This review aims to put all these different approaches together and compare them.

2. Optical techniques

2.1. Basis of OCT imaging: acquisition protocols and resolution

Although they should not be seen as absolute boundaries, most of the literature converges towards frequencies related to organelle transport, as captured by D-OCT, that are deduced to range between 0.1 and 20Hz [16,17,18]. Hence, one of the main features to consider in designing a D-OCT system is the temporal resolution. Today, there are two main types of D-OCT architectures, enabling an optimal temporal analysis of 2D images: scanning Fourier domain OCT (FD-OCT) and time domain full-field OCT (TD FF-OCT). In FD-OCT, the image of a point source is scanned transversally on the sample and the interference signal from the entire depth of field is recorded by operating an inverse Fourier transform of the interference spectrum [19]. The spectrum is either obtained in a single shot of a line camera behind a diffraction grating [20], or by sweeping the wavelength of a monochromatic light source [21]. In both cases, a B-scan (an \(xz\) plane acquisition - see Fig. 1(ii)) can typically be achieved in a few dozens of milliseconds [24]. In FF-OCT, the image of an extended source is projected onto the sample and the resulting interference is collected by a two-dimensional camera, as shown in Fig. 1(i). An \(enface\) \(xy\) plane can be recorded at the camera frame rate, typically around a few hundred Hz. Of course, time series could be captured in 1D, but because of the rather slow frequencies of interest (down to 0.1 Hz), it would require a few hours to capture a full 2D D-OCT image, making this unrealistic. To our knowledge, there is no technique reported that could achieve 3D D-OCT in a single time-series acquisition (i.e. with volume rates above 20Hz with high enough sensitivity).

The temporal evolution in scanning D-OCT is then followed by realizing BM- or MB-scans: the former is most often used as it allows mapping of an entire cross-section at subsequent time points. The temporal resolution is given by the B-scanning frequency \(f_B\) (see Fig. 1(ii)), namely how many B-scans can be acquired in a predetermined time interval. Consequently, the fastest cross-sectional dynamics that can be observed will equal \(\frac{f_B}{2}\), according to the Shannon
Fig. 1. i) schematic showing the usual setups employed in scanning FD OCT vs TD FF-OCT ii) In scanning FD-OCT, one A-scan will be acquired in one shot, carrying the information of the whole depth in z, which will be then extracted in post processing by Fourier-transforming the signal. Subsequently, a B- and C-scans are needed to have the complete view of the three-dimensional sample. In the case of dynamic imaging, two different procedures can be followed: either one can acquire an entire cross-sectional plane first and then follow the evolution of the entire plane over time (BM-scan) or follow the evolution of the intensity over a fixed amount of time for one spot and then shift along the x direction to do the same with all of the other points (MB-scan) iii) For D-FF-OCT the temporal evolution of one transverse plane is recorded and then the focus is shifted in the axial direction to acquire the volume. NPBS: non polarizing beam splitter; Pos.: position; CE: Compensation element, Im. L: imaging lens; FPI: focal plane of interest.

Theorem [22]. A typical range for $f_{B-SCAN}$ is between 0.5 and 180 Hz. If one needs to follow faster dynamics, MB-scans are usually captured, as it is possible to have a time series of A-scans on the same locations with frequencies on the order of tens of kHz [11,23], namely $f_{A-SCAN} \approx 20$ kHz. Please refer to Fig. 1(ii) for the visualization of the scanning D-OCT acquisition protocol. An important aspect is also that the illumination phase and intensity should be stable over the acquisition time so that they do not mask the sample dynamics.

The case of D-FF-OCT is slightly different, given its realization. Microscope objectives are usually employed both in the reference and sample arms in a Linnik configuration, and the use of a two-dimensional camera allows the acquisition of an entire transverse plane in one shot. Differently from scanning D-OCT, the dynamics over the transverse plane is analyzed first, collecting a series of interferograms over time on the very same z position with a frame rate of 80 Hz minimum, and then the cross-sectional view is reconstructed by scanning over the axial direction [7]. Fig. 1(iii) shows how this process works.

Generally speaking, the most important parameters for dynamic imaging are the speed of acquisition, the phase sensitivity to the scatterers’ motion, and the resolution. The phase sensitivity is important to translate the smallest shifts of scatterers into measurable variations in the scattered signal. With poor phase sensitivity, only large axial motion of scatterers can be captured, hence through larger acquisition times, or smaller frequencies. The optical resolution
is important so that the movement of organelles can be detected by means of signal intensity variations coming from single voxels.

The difference in the resolution among all the techniques is mainly related to the wavelength used, the bandwidth of the source and the numerical aperture (NA) of the optical components present in the sample arm. The formula for the transverse resolution ($\Delta r$) is the same for every technique and depends on the wavelength of the source ($\lambda$) and the NA of the objective lens in the sample arm [24]:

$$ \Delta r = \frac{\sqrt{2 \ln 2} \lambda}{\pi NA} $$ (1)

As a result, short wavelengths and high NA are preferred for a higher transverse resolution. Both parameters are also important to boost the scattering cross section for small organelles. One of the constraints on the wavelength comes from the samples themselves, as most of the tissues can be photo-damaged if illuminated with relatively high power in the visible range. Besides, the penetration depth can only exceed a few hundred micrometers when imaging in the near infrared, and with relatively low NA to minimize aberrations and multiple scattering.

As far as axial resolution ($\Delta z$) is concerned, it is governed at low NA by the spectral bandwidth ($\Delta \lambda$) of the interferometric signal for scanning OCT, which, for Gaussian sources can be written as [25]:

$$ \Delta z = \frac{2 \lambda^2 \ln 2}{\pi \Delta \lambda} $$ (2)

so the broader the bandwidth, the higher the axial resolution. When high NA objectives are used, as in D-FF-OCT, or in µOCT, the axial resolution depends on a mix between the temporal coherence as written above and the spatial longitudinal coherence, which arises from the large angular bandwidth authorized by the objective NA [26]. For high NA systems with poor temporal coherence length, the axial resolution is dictated by the longitudinal coherence length as follow [26,27]:

$$ \Delta z = \frac{0.44 \lambda}{n(1 - \cos(\theta_{max}))} = \frac{0.88 \lambda}{n \left( \sin \left( \frac{\theta_{max}}{2} \right) \right)^2} $$ (3)

being independent from the bandwidth and only related to the NA ($\theta_{max}$ is the maximum acceptance angle of the objective lens) as well as to the wavelength. In general, axial and transverse resolutions used in D-OCT microscopes range between 0.5 and 20 µm, hence the possibility to look at cells, or even inside them, arises.

The first idea of analyzing how OCT signals change over time in tissues was presented by Van der Meer et al. in 2010 [28] where they studied how the attenuation coefficient of the scattered signal, namely the inverse of the scattering mean free path, recorded in the measurement is related to either apoptosis or necrosis of human fibroblasts. The type of cellular dynamics extracted in this study is very different from the one that can be found in the state-of-the-art studies nowadays, as single B-scans (See Fig. 1) were taken every few minutes, but still enabled differentiation of apoptotic and necrotic cells [28]. Technological improvements, namely the increased frame rate of the detectors for scanning Fourier domain OCT and dynamic full-field OCT, and the increased wavelength sweeping speed for swept source OCT, permitted the increase of temporal resolution to a point where active movement of organelles became accessible, making live D-OCT imaging and the study of ATP-driven processes possible [29].

Analyzing temporal fluctuations associated with subcellular movements is not restricted to OCT-based measurements. For instance, dynamic holography (DH) is another interferometric technique based on a Mach-Zender-like interferometer setup [30]. A series of subsequent holograms is captured and statistical fluctuations of dynamic speckle are recorded to have an insight into cell motility as an integrated modulation of the path of multiply scattered photons. The evolution of the main fluctuation frequency over time is used to visualize the response of
different parts of the samples to different stimuli \[31,32,33,34\]. Although leading to a similar type of contrast in the sample, DH exploits multiply scattered photons, and hence probes fluctuations related to the combination of the dynamics of many cells together. On the other hand, D-OCT techniques are, to our knowledge, quite unique in the label-free dynamic microscopy field as they can access subcellular resolution and study metabolic information at the single cell level. Henceforth, given the fact that we want to focus this review on the inspection of cell dynamics, we choose to exclusively focus on D-OCT. We will first present the different setups used to achieve D-OCT, then we will show an overview of the metrics that are extracted in order to better understand and explain what is observed in all the applications explored by the different scientists.

Most of the results that will be mentioned in this article are obtained using custom-made OCT systems, such as spectral domain OCT (SD-OCT) \[23\], scanning microscopic OCT (m-OCT) \[35\], micro spectral domain OCT (µOCT) \[11\], polarization-sensitive SD \[36\] or swept-source \[37,38\] OCT (PS-SD-OCT, PS-SS-OCT) and full-field OCT (FF-OCT) \[12,13\], and distinct ways of analyzing the data, and extracting metrics from the recorded signals. A summary of the technical parameters can be found in Table 1; however, the different analyses and modalities will be further presented in detail in the following paragraphs.

### Table 1. summary of the technical parameters of the setups being considered. SCL: supercontinuum laser, SLD: super luminescent diode, LED: light emitting diode.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Type of source</th>
<th>Wavelength range</th>
<th>Lateral resolution</th>
<th>Axial resolution</th>
<th>Depth of imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD-OCT [23]</td>
<td>SLD</td>
<td>1310 nm ± 55 nm</td>
<td>7 µm</td>
<td>7 µm</td>
<td>1 mm</td>
</tr>
<tr>
<td>m-OCT [35]</td>
<td>SCL</td>
<td>550-950 nm</td>
<td>1 µm</td>
<td>1 µm</td>
<td>550 µm</td>
</tr>
<tr>
<td>µOCT [11]</td>
<td>SCL</td>
<td>800 nm ± 150 nm</td>
<td>2 µm</td>
<td>1 µm</td>
<td>300 µm</td>
</tr>
<tr>
<td>PS-SD-OCT [36]</td>
<td>Ti:Sapphire</td>
<td>735-865 nm</td>
<td>12 µm (in air)</td>
<td>2.6 µm (in air)</td>
<td>1.5 mm</td>
</tr>
<tr>
<td>PS-SS-OCT [37,38]</td>
<td>SS laser</td>
<td>1310 nm ± 50 nm</td>
<td>19 µm</td>
<td>14 µm</td>
<td>3.6 mm</td>
</tr>
<tr>
<td>FF-OCT [13]</td>
<td>LED</td>
<td>810 nm ± 25 nm</td>
<td>0.4 µm</td>
<td>0.4 µm</td>
<td>230 µm</td>
</tr>
</tbody>
</table>

#### 2.2. Scanning FD-OCT

Dynamic scanning FD-OCT is indeed a very broad definition that includes different declinations of the technique, starting from how the signal is recorded to how the data is processed. FD-OCT involves the use of a broadband source (wavelength bandwidth on the order of hundreds of nanometers) usually emitting in the infrared range, such as super luminescent diodes \[23\], supercontinuum lasers \[35\] or Ti:Sapphire lasers \[36\], depending on the axial resolution one is hoping to achieve. The light source is imaged on one transverse position in the sample and then scanned by means of two galvanometric mirrors in order to acquire B-scans and C-scans. In swept source OCT setups, instead, a wavelength sweeping laser is used in order to achieve the required broad bandwidth \[37\]. In the case of polarization sensitive devices, polarized light impinges on the sample and on the reference arm and the information related to the polarization of the light scattered back by the sample is conserved in the detection process \[36,37\]. By turning the incident polarization, and by measuring orthogonal scattered polarizations, it is possible to obtain the full Jones’ matrix. Nonetheless, to the best of our knowledge, no dynamic polarization effect was ever reported, although several dynamic birefringent structures may be found in cells \[39\].

#### 2.3. Full field OCT

In addition to the tailor-made systems, a commercial setup has also been used to perform D-OCT: the Light-CT (LLTech), now renamed Celtivity (Aquyre Biosciences) \[40\]. This system is based
on full-field OCT (FF-OCT), whose principle was first presented in [41], and it is able to get en face images in one shot. The temporal interference signal is recorded and a dynamic contrast is generated based on the intensity fluctuation statistics of each pixel, as described in [7]. The technique is hence known as dynamic full-field OCT (D-FF-OCT) but it also can be found under the name of dynamic cell imaging (DCI) [42]. The system has been used by several research groups for applications that range from breast cancer surgery assistance [43] to lung biopsy assessment [44] to pancreatic cancer analysis [45] to HeLa cell-related studies [5,6,46,47].

3. Signal analysis and metrics

Before passing to the signal processing and rendering, we will give a small introduction on the potential origin of the dynamic signal and how it can be linked to cell metabolism.

Mitochondria are some of the main cell organelles which are key for energy production and cell metabolism. Groux et al. demonstrated that mitochondria are involved in D-FF-OCT signal in retinal pigmented epithelial cells [7]. However, this can hardly be generalized since, depending on cell types, and across time and space within the same cell, mitochondria can have many different shapes, sizes and motion [48]. Mitochondria alternate between random Brownian motion, and directed, ATP-driven motion via molecular motors and cytoskeleton interactions [49], while speed in the directed motion phase can equal few hundred nm/s [48,49] and up to 1000 nm/s at 37°C [49].

Even though the basic mechanisms of mitochondrial transport are quite clear, it is still difficult to fully explain how these translate in the actual dynamic signal that is collected. This is mainly related to two different effects that might hint at a better reconstruction of the considered phenomena: there is spatial filtering due to the diffraction limit and D-OCT is mostly sensitive to the axial motion, so there is a modulation of the signal depending on the actual angle of displacement.

In D-FF-OCT, it has been demonstrated that the intensity fluctuations due to the scatterers motion in the voxel of interest are related to ATP-driven transport, opposed to Brownian random motion, since the dynamic signal disappears when the ATP production is blocked [7]. It was also shown that a relaxation of the cytoskeleton promoting cell motility increased the frequency of the D-FF-OCT signal [6].

As shown in section 2.1, there are two main ways to acquire the signal in D-OCT and, consequentially, two different ways to follow the evolution of the sample over time, namely either to look at the (x, z) cross-sectional plane or at the (x, y) transverse plane, before moving to the third spatial dimension. Once the time-varying signal is acquired, the main question is how to treat this type of data. Cell dynamics is a complex problem that involves many bodies’ interactions. The post-processing algorithms developed so far aim to reduce this complexity, defining a standard that allows the discrimination of different types of samples. This is achieved by extracting different metrics from the interferometric signal that will result in different types of images. The sensitivity to nanometric shifts of cells and subcellular organelles is reflected in a time-dependence of the recorded intensity, that itself gives information on the aforementioned shifts, thus on the metabolism of ATP-driven processes. It may seem counterintuitive that D-OCT relies on small movements of scatterers affecting the phase, and all the D-OCT methods use intensity-based measurements. Although appropriate phase measurements might emerge in the future, our understanding is that direct phase measurements are somehow weighted by signal intensity in D-OCT, and that phase signals show large unrealistic temporal fluctuations where the signal is low, as demonstrated in [50]. In contrast, using the signal intensity enables one to couple scattering information to fluctuations, and damp the influence of the noise. To conclude, to our knowledge, no D-OCT method is able to retrieve independently and at the same time phase and intensity of the signal, problem that remains an open question of high interest in the OCT community.
3.1. Motility coefficient

In scanning D-OCT (BM-mode), \( N \) cross-sectional frames will be acquired over a certain time interval \( \Delta t_{TOT} \). The \( i \)-th frame’s intensity will depend on the \( x \) and \( z \) coordinates and on the time point \( i \), thus \( I(x, z, i) \), therefore time dependency and intensity magnitude are two quantities to look at for dynamics inspection. Oldenburg et al. first introduced a metric called motility coefficient \( (M) \) to assess how the intensity of the OCT signal was changing over time [11]. \( M \) was defined as:

\[
M(x, z) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [S_{OCT}(x, z, i) - \langle S_{OCT}(x, z) \rangle]^{2}} \tag{4}
\]

where \( S_{OCT} \) is the signal amplitude and \( \langle \cdot \rangle \) indicates the time average. This parameter is useful to identify regions of the sample that are not active, leading to speckle fluctuations that are shot noise limited. In the cited study, this was used to distinguish regions with live cells from regions where gold nanorods could diffuse [29]. The time scale for these variations was calculated by fitting \( \log \gamma(t) = \frac{-t}{\tau} + c \), where \( \gamma \) is the normalized temporal autocorrelation of \( S_{OCT} \) and \( \tau \) is the decorrelation time. Gold nanorods showed shorter correlation times than mammary epithelial cells (MEC) and extracellular matrix (ECM). In this particular case, a polarization-sensitive setup was employed and cross-polarization \((C)\) was also computed because nanorods are optically anisotropic so it constituted another way to recognize them on top of biological material. The combination of \( M \), \( \tau \) and \( C \) provides a specific signature for each component of the experiment [11]. In a more recent contribution [51], a new motility coefficient \( M' \) was defined as follows:

\[
M'(x, z) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \Gamma_i(x, z, \Delta t)} \tag{5}
\]

where \( \Gamma_i(x, z, \Delta t) \) is the temporal autocorrelation of \( I(x, z, i) \) and \( \Delta t \) is the imaging sampling time. This was done because the new formulation of \( M \) was demonstrated to be insensitive to the intensity itself and to the depth coordinate. The \( M' \) coefficient was also used in a study where mouse cornea were exposed to Benzalkonium Chloride [52].

3.2. Power spectral density

Another approach to the problem of recognizing cell types in a sample takes into account the power spectrum of the intensity. In the same paper [51], the power spectral density (PSD) for each pixel is calculated as the square modulus of the Fourier transform of the intensity, \( PSD(x, z, f) = |F\{I(x, z, i)\}|^{2} \), and then fitted to an inverse-power-law model \( PSD(f) = c_{0}f^{-\alpha} + n \), where \( n \) is additive white noise. After some tests on MEC organoids embedded in ECM both in unfixed and fixed samples, by making the fit on selected regions of interest (ROIs), it was shown that this model was properly describing the cell activity so \( \alpha \) was selected as an additional coefficient for sample characterization. Visualization of this comparison can be found in Fig. 2(i). In a later work on MEC organoids’ response to toxicants, \( \alpha \) and \( M \) were joined in a single metric called intracellular kinetic energy \((E)\) and defined as \( E = \alpha M \), that supposedly describes cellular activity in toto [53].

The idea of obtaining the PSD per pixel was also exploited by Münter et al. in [35]: instead of fitting the curve to a predetermined model, the authors integrate the PSD in three different bandwidths, namely 0-0.5 Hz, 0.5-5 Hz and 5-25 Hz, obtaining three coefficients that will describe the frequency variation of the pixel intensity. A similar procedure was employed in [11] but the frequency ranges were different and non-adjacent.
Fig. 2. Some D-OCT based cell dynamics metrics i) results for motility index (M) and power spectral density (PSD) decay rate ($\alpha$) of mammary epithelial cells (MEC) spheroids before (blue) and after (red) fixation. As it can be seen, those parameters are significantly different depending on the spheroid condition: the motility is higher for the live sample, the exponential is indeed lower, as expected. Figure taken from Ref. [51] ii) temporal trend of the static power ratio (PR) in gelatin phantoms inseminated with (red curve) and without (green curve) yeast taken from Ref. [23]. The higher activity of the yeast makes the PR decrease faster iii) differences in the time evolution of the logarithmic intensity variance (LIV), and the OCT correlation decay speed (OCDS$_e$ and OCDS$_l$), respectively, in areas with high and low dynamics. x-axis shows time, y-axis shows the ratio between the area of low/high dynamics and the total area of the sample. Figure taken from Ref. [55]

3.3. Intensity variance

Moreover, the pixel intensity variance $V(x, z) = \frac{1}{N} \sum_{i=1}^{N} \left[ I(x, z, i) - \frac{1}{N} \sum_{i=1}^{N} I(x, z, i) \right]^2$ was calculated as an additional parameter in a different study in order to better visualize the changes that a spheroid of HeLa cells undergoes during interactions with bio-printed scaffolds [54].

Similarly, Abd El-Sadek et al. defined a new metric called logarithmic intensity variance (LIV) in order to evaluate the pixel-by-pixel evolution of the signal over time [55]. The idea of the metric is based on the possibility for the signal to be divided into static and dynamic contributions, $I(x, z, t) = I_D(x, z, t)I_S(x, z)$. Once this is stated, the temporal variance of the signal is calculated, after converting the intensity into dB, a step that allows the suppression of the static component.

$$I_{DB}(x, z, i) = 10\log_{10}I_D(x, z, i) + 10\log_{10}I_S(x, z)$$

$$LIV(x, z) = \frac{1}{N} \sum_{i=1}^{N} \left[ I_{DB}(x, z, i) - \langle I_{DB}(x, z, i) \rangle \right]^2 =$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left[ 10\log_{10}I_D(x, z, i) - 10\langle \log_{10}I_D(x, z, i) \rangle \right]^2$$

Hence, there will be a coefficient per pixel evaluating the variance: the collected piece of data will be only related to the magnitude of the signal and will not give any information about the speed of the dynamics.
3.4. OCT correlation decay speed (OCDS)

This is the reason why, in the same paper, another coefficient was introduced in order to estimate the speed of variation of this intensity signal: the OCT correlation decay speed (OCDS). This parameter studies the temporal evolution of the autocorrelation of the intensity, defined as follows:

\[
\rho_A(\tau_j, x, z) = \frac{\text{Cov}(I_{dB}(x, z, i), I_{dB}(x, z, i + \tau_j))}{\text{Var}(I_{dB}(x, z, i))\text{Var}(I_{dB}(x, z, i + \tau_j))}
\]

where \(\text{Cov}(\cdot)\) and \(\text{Var}(\cdot)\) indicate the covariance and the variance of the signal and \(\tau_j = j\Delta t\), with \(\Delta t\) being the repeating time of sequential frame acquisition and \(j \in \mathbb{N}\). The curve of \(\rho_A\) is then plotted against \(\tau_j\) and the slope is calculated in different delay time intervals, obtaining the OCDS. The authors identify two zones, the early and late OCDS: if cells have fast dynamics the early OCDS is greater than the late one; they are similar if the observed dynamics are slow. This is related to the fact that the autocorrelation function will quickly decay if the scatterers are oscillating fast: in other words, taking into account the Wiener-Khinchin theorem, this would imply having a PSD oscillating at high frequency. In this particular case, these three parameters synthesize the information about the sample. In addition to the LIV, using a Jones matrix system, Mukherjee et al. were able to extract a birefringence (BR) and an attenuation coefficient (AC) image at the same time. Principal components analysis is then performed to understand how much each parameter (LIV, BR or AC) is contributing to the different sample components [56].

3.5. Eigenvalue decomposition and static power ratio

A new set of metrics was introduced by Wei et al. based on eigenvalue decomposition (ED) of the dynamic signal [23]. First of all, the authors demonstrate the ED-based method to be almost twice as sensitive to Brownian-like motion with respect to intensity-based decorrelation algorithms. This is shown by calculating the contrast to noise ratio (CNR) as:

\[
\text{CNR} = \frac{(\mu_s - \mu_b)}{\sqrt{\sigma_s^2 + \sigma_b^2}}
\]

where \(\mu_s\) and \(\sigma_s\) are the mean and standard deviation over the signal ROI and \(\mu_b\) and \(\sigma_b\) are referring to the background ROI. The parameter is almost twice bigger in the ED-based case.

That being said, subcellular dynamics were evaluated by ED analysis on yeast phantoms, given the high mitochondrial activity of this type of sample. The analysis was performed on \textit{en face} images obtained by means of maximum intensity projection of cross-sectional stacks. Pairs of A-lines acquired at different time points (with \(\tau\) the time interval that separates them) are coupled together to form the so-called cluster and a \(2 \times 2\) covariance matrix is obtained by multiplying the cluster by its transpose complex conjugate as follows:

\[
R_C^\tau = \frac{1}{M} \begin{bmatrix} (A_1^*)^T & A_1 \end{bmatrix} \begin{bmatrix} A_1 & A_\tau \end{bmatrix} = \frac{1}{M} \begin{bmatrix} \text{Var}(A_1) & \text{Cov}(A_1, A_\tau) \\ \text{Cov}(A_1, A_\tau) & \text{Var}(A_\tau) \end{bmatrix}
\]

By assuming that the static and dynamic behavior of the sample are linearly uncorrelated, this matrix is diagonalized in order to find its eigenvalues (EVLs), minimizing the difference between the components: the first EVL will be related to the static composition of the sample while the second EVL is related to the dynamics.

\[
R_C^\tau = E^\tau \begin{bmatrix} \lambda_1^\tau & 0 \\ 0 & \lambda_2^\tau \end{bmatrix} (E^\tau)^T
\]

where \(E^\tau\) is the eigenvectors matrix and \(\lambda_j^\tau\) are the corresponding EVLs for the specific delay \(\tau\). Once this is done, it is possible to follow the dynamics of each A-line over time by calculating
what has been defined as the static power ratio $PR^\tau = \frac{\lambda^1}{\lambda^1 + \lambda^2}$. As for the Brownian motion theory, the behavior of static particles can be described with an exponential function, thus it is reasonable to expect the $PR^\tau$ to decay as $PR^\tau = Ce^{Gr}$, where $G$ is the fitted damping coefficient and $C$ is the peak static power ratio. As expected, phantoms with yeast show faster decay with respect to phantoms without yeast in the short term ($G$ is negative so $G_{YEAST} < G_{NOYEAST}$), due to higher subcellular activity (see Fig. 2(ii)). The same procedure is then applied in vivo while looking at mouse brain, comparing a healthy case to a mouse which suffered an ischemic stroke. After removing the contribution of the blood vascularization from the images, the activity of the tissue was evaluated averaging the $PR^\tau$ over an en face image in both the conditions. It was possible to deduce that the healthy mouse showed higher brain activity than the diseased one. A similar idea of signal decomposition is employed by Scholler in [57] where a singular value decomposition (SVD) is applied in order to remove motion artifact in D-FF-OCT in vivo measurements. Here also, the idea is to decompose the signal fluctuations into some static signal, including vibrations creating spatially correlated noise, and a dynamic part, associated with cell activity. In contrast to [23] and [56], the SVD algorithm was applied to the full dataset at once. Hence, many eigenvalues and associated eigenvectors can be found, and the difficulty is to automatically extract the meaningful ones. Interestingly, Scholler could show that the time evolution of eigenvectors associated with mechanical vibrations exhibit specific features, in particular displaying a higher zero-crossing rate, enabling the extraction and cancellation of these specific eigenvectors, hence enabling one to automatically keep only the eigenvectors associated with cell dynamics.

### 3.6. Analysis for D-FF-OCT

On the other hand in D-FF-OCT, $N$ transverse frames ($I(x, y)$) are acquired over a certain time interval $\Delta t_{TOT}$: the first attempt of dynamic imaging using FF-OCT was made by Apelian et al. in [7] where the standard deviation of the time-varying signal is computed per pixel on subsets of the original time-stack and then averaged to get one single dynamic image, $D(x, y) = \left( \sqrt{\frac{1}{N} \sum_{i=1}^{N} (I(x, y, i) - I(x, y))^2} \right)^2$. The temporal size of the sub-stack is chosen empirically so that it roughly matches typical frequencies of cell fluctuations. Hence, the standard deviation saturates and better contrast can be obtained by averaging several standard deviation images. The result is then compared to a static FF-OCT image of the same ROI to show that different features are revealed. The evolution of the same sample (rat liver) is followed over time in order to demonstrate that the dynamic signal decreases as the sample starts dying, so it must be related to ATP-driven processes as postulated. In a later study [58], and as currently performed in the commercial system from Aqurye to the best of our knowledge, the pixel intensity was Fourier-transformed and integrated in three arbitrary normalized frequency bandwidths in order to get information about the temporal evolution of the signal, as is done in some scanning D-OCT [24]. However, the current state-of-the-art of signal analysis for D-FF-OCT is described in [12,13]: three parameters are retrieved from the recorded intensity and used to describe the observed sample, providing a more complete overview of what happens at a sub-cellular level. These three calculated metrics are recapitulated into one colored image, formed in an HSB color space to allow for fast and physically meaningful visual inspection. First of all, the mean frequency of the pixel intensity is computed as the scalar product of the PSD and the frequency array, $f_{MEAN} = PSD \cdot f$. The standard deviation of the frequencies is calculated as $\sigma = PSD \cdot f^2 - (PSD \cdot f)^2$ as a second parameter to have an idea of the frequency variations of the intensity. Lastly, the running standard deviation average of the signal is computed to complete the picture about the sample, gaining information both on frequency and intensity variance. There has also been an attempt of using the windowed cumulative sum of the signal in [57], as it is proven to be more sensitive to non-Brownian motion. Directed motion becomes more evident on
top of noise with this operator, thus a nearly 2-fold increase in SNR is registered with respect to the running standard deviation average.

3.7. Visual rendering and colored image generation

Once the signal is acquired and the metrics are computed, the last step of the D-OCT imaging process is the visual rendering of what has been observed: this constitutes the link between label-free non-invasive techniques and immunohistochemistry and fluorescence imaging. In the latter techniques, the color is provided by differential absorption of exogenous molecules whereas in D-OCT it has to be sought intrinsically, in order to be able to compare what has been observed with these gold standard techniques. In this paragraph, different ways of producing colored images will be explored.

In the cases where a single parameter is calculated, grayscale images are produced, assigning to each pixel the value of the computed coefficient. This is the case of [23] where the $G$ parameter (See subsection 3.5) is plotted for an en face image and [54] where the image obtained by the pixel intensity variance is shown.

More complete visualizations employ three parameters to generate the pictures that are then compared to validation images: this is the case for Red-Green-Blue (RGB), Hue-Saturation-Brightness (HSB) and glyph representations.

RGB space is employed in the studies where the curve of the PSD is integrated in three bandwidths that are assigned to the three channels of the color space [58,11,59,35,60,43]: the blue channel is used to encode slow dynamics, the red channel for fast dynamics and the green

Fig. 3. i) H&E-stained histology and scanning dmOCT (RGB method) image of a mouse tongue. The numbers (I-V) indicate the different layer of the tissue, identifiable in both the images: (I) cornified layer, (II) granular and spinous layers, (III) basal layer, (IV) lamina propria and (V) muscle. (*) indicates cells nuclei. Images taken from Ref. [35] ii) H&E-stained histology and scanning D-µOCT (RGB method) image of a human esophageal biopsy. Cell cytoplasm (green), nuclei (red dots inside) and perinuclear region (blue) can be observed in the µOCT image and they can be found in the histology as well. Frequency bandwidths and corresponding color channels are shown. Images taken from Ref. [11] iii) H&E and dynamic cell imaging (DCI) images of a benign breast lobule (RGB method), double layered structure is identifiable in both the frames. Images taken from Ref. [43]
channel for medium dynamics. The three channels are then fused together to get a single colored image, as shown in Fig. 3(i)-(iii).

The HSB approach has been used in two different ways. In recent D-FF-OCT work [9,12,13,61], the hue is used to encode the mean frequency of the PSD, the inverse of the frequency bandwidth standard deviation is assigned to the saturation channel while the value is used to depict either the average running standard deviation [12], or cumulative sum [57] on small time periods of the intensity. The hue channel is rescaled in order to prevent a circular lookup table: in this way, the information on the speed of oscillation is still carried by the color (blue areas represent slower dynamics with respect to red areas). The saturation channel is used to discriminate areas with broader bandwidth, that might contain white noise and are artificially rendered as dull, from areas with narrower bandwidth, which appear more vivid. The brightness channel is instead used to display the information coming from the time-dependent intensity signal, making the pixels with highest activity the brightest. An example of this technique is shown in Fig. 4(i). The mean frequency and the normalized integral of the PSD were also used by Park et al. in [5] to generate colored images of HeLa cells.

A different meaning was assigned to two channels in [55,62]: the saturation channel has been set to 1 for all the pixels, leaving just H and V as information carriers. Either the LIV or

---

**Fig. 4.** i) Immunohistochemistry (IHC) and D-FF-OCT (HSB method) images of a retinal organoid. Multipotent retinal progenitor cells are tagged with Visual System Homeobox 2 (VSX2) transcription factor, depicted in green. The formation of a rosette is visible in both the images and it is highlighted by the red dotted line. Images are taken from Ref. [12] ii) Cross-sectional OCT images of a MCF-7 spheroid. “Static” image shown in the top left corner and dynamic (HSV method) images shown elsewhere, comparing what can be obtained when the Hue channel is assigned to LIV, OCDS$_l$ and OCDS$_e$. Images are taken from Ref. [55] iii) Glyph representation of MEC organoids embedded in fibroblasts. The graphical meaning of the three parameters is explained on the bottom. Image taken from Ref. [51] iv) Mosaic of 5 × 5 tiles HSV renders of human fibroblasts, on the right a zoom of the red square area. Actin filaments highlighted with the purple arrow, ring-like subcellular structures pointed at by the orange arrow. Pictures taken from Ref. [61]. Scale bars stand for 40 µm in Fig. 4(iii), 100 µm in Fig. 4(iv) (left panel), 10 µm in Fig. 4(iv) (right panel).
the OCDS (early and late) are encoded in the hue, with a color scale that goes from red (slow dynamics) to green (fast dynamics) whereas the brightness of the pixels is related to the OCT signal intensity itself (See Fig. 4(ii)).

The third way to generate metadata for the same image is the so-called glyph representation: Oldenburg et al. use this in [51] showing three parameters on the same image, $\alpha$, $M$ and the coefficient of determination $R^2$ obtained from the fitting of the previous two parameters. $R^2$ states the reliability of the aforementioned coefficients. The idea behind this type of rendering is to generate a grayscale image using one of the parameters as base for the color-scale ($M$ in this case) and then overlay a matrix of glyphs that take into account small ROIs (9 × 9 pixels in this case): there is no pixel-to-pixel correspondence between the base image and the glyphs matrix. The diameter of the glyph is related to $R^2$ while the color is related to $\alpha$, as shown in Fig. 4(iii).

Once the pseudo-colored images are generated, it is easier to identify the components of the sample, making D-OCT comparable to what can be addressed in immunohistochemistry (IHC) and standard fluorescence (F). Fig. 3(i)-(iii) summarizes and compares what can be obtained with D-OCT image rendering and corresponding IHC and F images.

4. Applications

D-OCT offers an alternative or a complement both to immunohistochemistry and fluorescence imaging, overcoming two main drawbacks of these techniques: invasiveness and need for labeling. The idea of getting a meaningful signal from cellular movement and, more specifically, intracellular motion of organelles, opened up a broad range of applications, especially concerning biology and medicine. As a result, one could think of following cell behavior over time in response to different conditions, thus evaluating the effectiveness of treatments or the evolution of diseases.

4.1. In vitro cells and organoids

Monitoring cell state was one of the first applications of dynamic imaging by means of the OCT signal: Farhat et al. in [63] detect apoptosis in myeloid leukemia cells in vitro by analyzing the time-varying scattering signal and calculating the decorrelation time, that decreases during the apoptotic process due to cell swelling leading to de-condensation of the cell and more space for organelles to freely diffuse.

Since the introduction of spheroids and organoids, there has been a growing need for imaging them at high resolution in order to understand the processes taking place in their development and in the case of drug testing. Scanning D-OCT has played an important role in this framework as it is capable of following single cells over time and allows cell recognition in space. Longitudinal studies were performed on adenocarcinoma spheroids (MCF-7 cell line) to assess their response to anti-cancer drugs, responsible of the depolymerization of microtubules. This happens in vitro for a duration of 28 or 72 hours for the two different studies [55] and [62]. The spheroids were characterized by analyzing the temporal variance and the decorrelation time of the OCT signal, as well as the measurement of the spheroid volume over time but, given the resolution of the system, (14 $\mu$m in the axial direction and 19 $\mu$m in the transverse), it was not possible to identify the intracellular motion, even though the change in signal is justified by the degeneration of microtubules.

There is another approach used to study the response of MCF-7 organoids to microtubule suppression longitudinally, that takes into account the speckle fluctuation variance to quantify the intracellular motion [64]. Also, mammary epithelial organoids embedded in stromal fibroblasts have been studied in order to understand their growth by calculating cell motility [51], intracellular kinetic energy and cross-sectional area of the organoid itself [53]. Transverse resolution is about 12 $\mu$m so, also in this case, sub-cellular organelles were not visible, even though they are responsible for the generation of the bulk dynamic signal. Speckle variance analysis has also
been used in higher resolution systems (microscopic OCT, mOCT, with a resolution of around 1
µm in all directions) in order to assess cell vitality in HeLa cell spheroids [54].

In contrast, thanks to the high spatial resolution of D-FF-OCT, it is also possible to observe
sub-cellular features and their behavior over time, as shown in [12] for retinal organoids. Apelian
et al. demonstrated at the beginning of D-FF-OCT the ability to resolve nuclei and nucleoli inside
single cells, thanks to sub-micron transverse resolution [32]. In addition to this, Monfort et al.
showed recently, for the first time, long-term time lapse imaging that allows the identification of
organoid growth over 17 days while keeping the same type of resolution [13]. Moreover, Groux
et al. were able to follow mitochondrial and lysosomal behavior change under mechanical stress
of the tissue, in parallel with the validation with immunohistochemistry [65,9]. D-FF-OCT was
demonstrated to be effective in the detection of invasive fungal infections, that could potentially
be used in hospitals [71]. As for two-dimensional cell cultures, D-FF-OCT coupled with a
segmentation algorithm was used to extend the imaging range, including the area at the sample-sample holder interface, thus widening the potential applications
to thin 2D cell cultures, as the technique was employed to image human fibroblasts [61].

4.2. Ex vivo tissues and human biopsies

Another category of samples for D-OCT imaging is ex vivo tissues. In this case, the goal of the
researchers is to establish this technique as the standard for optical biopsy, as it acts as a tool for
cell type recognition.

Its use has been effectively demonstrated in several studies on murine liver tissues, as it was
possible to detect the different components of the tissue itself [35,7,60] as well as its microvascular
complex in fibrotic samples [56] and changes of OCT signal variance due to the movements of
lipid droplets in non-alcoholic fatty liver disease [66]. Mouse trachea has been inspected as well,
being able to identify cells that are not visible in “static” OCT [59].

In the context of tissue analysis, there have also been studies on human biopsies. It has been
possible to locate different parts of the tissue, given their intrinsic dynamic signature due to
cell metabolism, similarly to what happens in murine tissues. Scanning D-OCT was used to
inspect esophageal and cervical biopsies [11], in this case achieving micrometer scale resolution
and recognizing nuclei on top of cytoplasm by means of their different frequency response,
while D-FF-OCT served as a tool to recognize breast cancer [42], facilitating the use of machine
learning assisted diagnosis [43].

4.3. In vivo demonstrations

In addition to the previous listed applications, scanning D-OCT was used on yeast, chosen for its
particular sensitivity to mitochondrial activity, and in vivo to detect cerebral activity in mice,
comparing a healthy case to one where the animal had suffered a stroke, as outlined in section 3.1.
As this latter case was carried out in vivo, researchers had to try to filter out the overwhelming
signal coming from blood vessels [23]. A three-dimensional model of tissue invasion by breast
cancer cells grown on a scaffold made of paper was studied as well [67]. Another application
of in vivo scanning D-OCT was performed in [68,69,70], where the goal of the studies was to
calculate the ciliary behavior in the murine oviduct. Time-series of OCT signal were recorded
and spectral analyses were performed in order to detect the beating frequency: this was possible
even though the resolution of the system was larger than the cilia themselves, as the dynamics are
retrieved by extracting the time variation of the signal due to the ciliary motion.

Also, there have been an attempt of in vivo D-FF-OCT imaging in zebrafish larvae, as shown by
Thouvenin et al. in [10], and in murine liver [57]. The first experiment was partially impaired by
the strong unwanted signal coming from the blood vessels, that to our knowledge still constitutes
a limit for further *in vivo* applications whereas the second was realized in the framework of a study about denoising D-FF-OCT from motion artifacts.

This list of applications highlights the importance and the transversality of D-OCT, with all its variants, in the biomedical sector.

5. Discussion and conclusion

Over the past few years, D-OCT has become a technique of reference for high resolution imaging for *in vitro* and *ex vivo* biological samples, both human and animal. Scanning dynamic microscopic OCT and D-FF-OCT show transverse and axial resolution at up to 1 µm resolution, allowing the study of single organelle behavior under different conditions. As for the imaging range, scanning OCT seems to always go deeper than D-FF-OCT in the sample before the SNR lowers too much (see Table 1). Also, most of the studies taken into account already show validation against standard H&E histology, paving the way for tissue and cell recognition using D-OCT contrast alone. That being said, one can imagine an imaging hub in which it is possible to combine both fluorescent microscopy and D-OCT on the same sample, ideally at the same time: this would mean having a complete view of the sample itself in one shot, as already attempted in [12,72]. Also, this would imply the possibility to exploit the best characteristics of both the techniques, without having to compromise on any aspect.

While being similar to each other and leading to a comparable contrast in the resulting images, the various approaches differ quite significantly in the way the signal is analyzed, weighing different quantities in different ways. All in all, information about the frequency of intensity oscillation seems to better reveal fine differences between different cell types, while fluctuation amplitude usually gives higher contrast images of the cells.

Given this ability of D-OCT to recognize cells, the next step for future studies might be the search for an automatic segmentation approach, in order to categorize different cell types based on their intrinsic metabolism without the need of external markers. In addition to this, three-dimensional automatic segmentation would allow one to gain information about the shape and the location of the single cells in space. In longitudinal studies, this would yield single-cell level data at different time points over the long term, opening the way for the evaluation of the effectiveness of drugs with individual cell response, or the tracking of cell development in disease modeling processes.

**Funding.** European Research Council (OPTORETINA (101001841)); Agence Nationale de la Recherche (IHU FOReSIGHT (ANR-18-IAHU-0001), OREO (ANR-19-CE19-0023), VISCO (ANR-21-CE30-0024)).

**Disclosures.** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data availability.** No new data were generated for this manuscript.

**References**


