Review Article

Prospects on Time-Domain Diffuse Optical Tomography Based on Time-Correlated Single Photon Counting for Small Animal Imaging

Yves Bérubé-Lauzière, 1,2,3 Matteo Crotti, 4 Simon Boucher, 1 Seyedrohollah Ettehadi, 1 Julien Pichette, 1 and Ivan Rech 4

1Department of Electrical and Computer Engineering, Université de Sherbrooke, 2500 Boulevard de l’Université, Sherbrooke, QC, Canada J1K2R1
2Centre d’Imagerie Moléculaire de Sherbrooke (CIMS), Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke (CR-CHUS), 3001 12e Avenue Nord, Sherbrooke, QC, Canada J1H5N4
3Institut Interdisciplinaire d’Innovation Technologique (3IT), Parc Innovation, Pavillon P2, 3000 Boulevard de l’Université, Sherbrooke, QC, Canada J1K0A5
4Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Correspondence should be addressed to Yves Bérubé-Lauzière; yves.berube-lauziere@usherbrooke.ca

Received 2 October 2015; Revised 5 December 2015; Accepted 20 December 2015

1. Introduction

Similarly to widespread medical imaging modalities, such as X-ray computed tomography (CT), magnetic resonance imaging (MRI), nuclear imaging (positron emission tomography (PET) and single photon emission computed tomography (SPECT)), and ultrasound (US), diffuse optical tomography (DOT) seeks to obtain deep (≥1cm) interior images of living organisms with noninvasive exterior-only measurements. Rather than using X-rays, radio waves and magnetic fields, gamma rays, or mechanical waves, DOT uses light and optical techniques (laser or incandescent sources, light detectors, and optical components) to achieve this. More precisely, DOT aims at reconstructing 3D spatial maps of the optical properties of biological tissues (absorption and scattering coefficients) in the case of intrinsic imaging or of the concentration of a light-emitting compound in the cases of fluorescence or bioluminescence imaging [1, 2]. This is achieved by measuring light emerging from a tissue at several locations on its boundary and feeding the measured data to an image reconstruction algorithm that will recover an image therefrom.
Optical methods present several interesting features for the purpose of medical imaging. First, light in the visible and near-infrared (NIR) ranges is nonionizing, and, thus, it does not harm biological tissues provided some maximum exposure limits are not exceeded [3, 4]. Second, optical methods are sensitive to blood oxygenation owing to the fact that blood absorption, which is a major component of the optical absorption of tissues, is directly linked to the oxy- and deoxyhemoglobin concentrations in blood [1, 5] through the relationship

$$\mu_a,\text{blood} (\lambda) = \ln(10) \left[ \epsilon_{\text{HbO}_2}(\lambda) C_{\text{HbO}_2} + \epsilon_{\text{Hb}}(\lambda) C_{\text{Hb}} \right]. \quad (1)$$

Here $\mu_{a,\text{blood}}$ is the optical absorption coefficient of blood, which is a function of wavelength $\lambda$, $\epsilon_{\text{HbO}_2}$ and $\epsilon_{\text{Hb}}$ are, respectively, the wavelength dependent molar extinction coefficients of oxyhemoglobin ($\text{HbO}_2$) and deoxyhemoglobin ($\text{Hb}$), and $C_{\text{HbO}_2}$ and $C_{\text{Hb}}$ are their respective concentrations. Figure 1 depicts absorption spectra of Hb and HbO2 at typical blood concentrations (150 mg/mL) along with those of other important chromophores present in biological tissues (water, fat, and melanin). For water and fat, the absorption spectra are for pure water and pure fat. For melanin, the spectrum displayed is that typical of skin obtained with the formula

$$\mu_a (\text{cm}^{-1}) = 1.70 \times 10^{12} \times \lambda^{-3.48},$$

with $\lambda$ given in nm. Blood oxygenation is not accessible with other medical imaging techniques, except for functional MRI (fMRI) through blood oxygen level dependent (BOLD) signals, which, however, represent a complex mixture of blood flow, blood volume, and oxygen metabolism [6]. The $e$ is no simple relationship for BOLD signals such as (1) to directly access oxy- and deoxyhemoglobin concentrations. This, optical imaging proves to be of high interest for in vivo molecular imaging due to the wide availability and relative ease of synthesizing molecular fluorescent and bioluminescent probes that possess high specific ty. In contrast, probes for PET require handling radioactivity and synthesizing molecules to which a large radioactive atom (e.g., copper) needs to be attached, which is chemically challenging [7]. Similarly, probes for MRI also require binding large and heavy atoms (typically gadolinium). Furthermore, owing to MRI’s very low sensitivity, large probe concentrations are oft n necessary to obtain sufficient signal. This can pose problems related to toxicity or to the possibility of altering the biological system under study via pharmacological effects [8]. Finally, optical technology is relatively cheap compared to other medical imaging technologies.

DOT finds applications in diagnostic imaging of the breast [9–12], joints [13], and prostate [14, 15], as well as in brain imaging for diagnostic and fundamental research purposes [16–18], whereas fluorescence DOT (FDOT) and bioluminescence DOT (BDOT) find most of their applications in small animal preclinical imaging [2, 8, 19–24], whereby one seeks to study biomolecular processes in small animal models of human diseases using light-emitting probes. This is of interest in a wide range of fields (i) biology and medicine to develop an understanding of the fundamental mechanisms of disease development; (ii) pharmacology to label a drug or medication and follow its progression to see if it reaches targeted tissues (this can lead to faster drug development, which is of great interest to the pharmaceutical industry), or to follow the evolution of a pathology under treatment on the same subject (longitudinal studies) to monitor treatment efficacy, which also allows reducing the number of animals used/sacrificed, a subject of great ethical interest; (iii) oncology for cancer cell tracking by labeling specific membrane proteins and monitor medication and radiotherapy effects [25]; (iv) toxicology to study chemicals and drugs toxicity and their effects on the health of living organisms and in several other areas (pH imaging, antibody labeling, enzymatic reactions, peptides imaging, etc.).

DOT involves 3 major aspects: (1) how light interacts with biological tissues, which determines the imaged parameters and leads to light propagation modeling; (2) instrumentation to carry out multiview measurements and acquire the data necessary for tomographic imaging; and (3) tomographic image reconstruction algorithms that allow reconstructing interior images from the acquired data.

DOT is a relatively immature imaging modality compared to CT, MRI, PET, SPECT, and US, primarily because the image reconstruction problem is highly ill-posed and ill-conditioned owing to the high degree of light scattering in tissues. One way of reducing this ill-posedness in both DOT and FDOT is to resort to time-domain (TD) optical measurements (note: this excludes BDOT since bioluminescence emission is intrinsically a continuous-wave signal). This will be discussed further later on.

The TomOptUS group at Université de Sherbrooke has undertaken the development of a TD-DOT scanner for small animal (mouse) molecular imaging. The first generation of this scanner has been discussed at length in [26]. It is not the purpose to go over this again here in detail but rather to overview the architecture and hardware, summarize the highlights and shortcomings of that first generation along with a possible upgrade.
with some illustrative results, and more importantly discuss what will come next and the work undertaken in that direction and associated early results.

Although image reconstruction will be alluded to, the present paper focusses on hardware aspects of TD-DOT and on some future prospects on it, notably as regards enabling technologies for the success of TD-DOT, specifically single photon avalanche diodes (SPADs) and parallel multichannel time-correlated single photon counting (TCSPC), two areas in which the SPADlab at Politecnico di Milano (PolyMi) has made significant contributions in recent years.

The paper is subdivided as follows. Section 2 reviews the basics of how light interacts with biological tissues. Hardware requirements for TD-DOT and associated challenges are discussed in Section 3. This is followed by Section 4 that overviews the first-generation TomOptUS scanner along with imaging results obtained with it. Section 5 discusses key advances related to TCSPC technology that will allow improving on the performances of the first-generation scanner; a second-generation preliminary prototype is presented along with associated imaging results which are compared with those of the first generation. Finally, Section 6 draws some conclusions, gives a summary, and presents an overview of current work.

2. Review on the Interaction of Light with Biological Tissues

When a photon propagates through a biological tissue, it can undergo two types of interactions: absorption and scattering. Figure 2 illustrates simplistic microscopic views of these two processes, whereby a photon absorbed disappears as its energy is converted into some other form (typically heat) by an absorber (typically some complex molecule in biological tissues, such as hemoglobin), whereas scattering causes the photon to be deflected in another direction without loss of energy (elastic scattering is considered here; scatterers in biological tissues are generally structures of different sizes such as cells, cell nuclei, and mitochondria down to cell membranes [1]).

The probability for a photon coming from a direction specified by the unit vector \( \hat{s}' \) to be scattered in another direction \( \hat{s} \) is given by the so-called phase function \( p(\hat{s}', \hat{s}) \) which can be obtained from electromagnetic theory such as Mie’s theory for scattering by spheres or ultimately from quantum mechanics by considering the fundamental processes underlying scattering (in tissue optics, however, one never goes that far as it is not necessary). In biomedical optics, the Henyey-Greenstein phase function, which has been determined empirically to represent well the scattering directional properties of biological tissues [27–29], is often resorted to [1, 30, 31].

Absorption and scattering provide the two possible contrast mechanisms for optically imaging a tissue in its native state. On a macroscopic scale, absorption, characterized by the so-called absorption coefficient denoted \( \mu_a \), is responsible for the attenuation of light energy in a medium (Figure 3(a)). The macroscopic intensity of a light beam is exponentially attenuated following Beer’s law \( I = I_0 e^{-\mu_a l} \), where \( I_0 \) is the injected intensity and \( I \) is the intensity after having traversed
the medium with thickness $l$. The absorption coefficient $\mu_a$ is wavelength dependent as indicated by the dependence on $\lambda$. Absorption in biological tissues is weakest in the near infrared (NIR) between 650 and 1000 nm owing to the absorption spectra of HbO$_2$ and Hb (Figure 1), which are the main absorbing chromophores. Hence, light in that range (the so-called therapeutic window) is preferably used to image biological tissues as it can travel centimeters before being absorbed, whereas visible light is practically completely absorbed after a few millimeters. Scattering, which is strong in biological tissues, is responsible for the macroscopic diffusive nature of light propagation therein (much like heat flow in a metal), a fact that degrades the spatial resolution achievable by optical imaging. Macroscopically, scattering is characterized by the so-called scattering coefficient $\mu_s$ denoted by $\mu_s$ along with the phase function. For scattering, the macroscopic intensity also obeys Beer’s law, with the scattering coefficient $\mu_s$ of the medium characterizing the strength at which scattering occurs within the medium (Figure 3(b)). Under some simplifications (diffusion theory), scattering is also customarily characterized by the reduced scattering coefficient $\mu' = (1 - g)\mu_s$, where $g$ is the anisotropy factor representing the average cosine of the angle between the incoming and scattered directions [1].

Photons in diffusive media can be thought of as propagating along 3 different regimes. These regimes can be discriminated by way of TD measurements made at the boundary of a medium of the photon time-of-flight curve [32] (the so-called time point-spread function (TPSF)—Figure 4). First are ballistic photons, which by definition propagate in straight lines, only suffering absorption without being scattered (their contribution is grossly exaggerated in Figure 4). Such photons are, however, too few to be used for any practical purposes in tissue imaging as soon as the propagation distance exceeds a few millimeters ($\approx$2-3 mm), beyond which practically all photons have suffered a scattering event [32]. Second are the so-called snake or early photons which suffer few scattering events and can be considered to propagate along nearly straight paths [33, 34]. The are the photons appearing early in a measurement and contribute to the rising edge of a TPSF (Figure 4). Finally are diffused photons following highly meandering paths. These are of lesser interest for imaging as their paths are completely random, making the regions of the medium which they probe unknown. They are ultimately responsible for the blurry nature of DOT images.

Since TD measurements directly allow discriminating the different regimes of photon propagation, they are generally accepted as providing for the richest data [1, 2]. Indeed, the shape of a TPSF directly depends on the absorption and scattering coefficients $\mu_a$ and $\mu_s$, whereas continuous-wave (CW) measurements provide only a number (light intensity) with no signature about how much a photon may have suffered scattering events while propagating. TD-DOT aims at exploiting this data richness to improve the quality of the reconstructed images in terms of obtaining higher spatial resolution and more quantitative imaging. This has been shown to be the case in both intrinsic and fluorescence imaging [35–37], and indications of this can also be found in our work on the development of image reconstruction algorithms [38]. Further, TD measurements also allow probing parameters not accessible with CW measurements, such as the fluorescence lifetime, which is the mean time for an excited fluorescent molecule to relax to its ground state. In simplest form, this relaxation is described by a single decaying exponential. In tissue, this decay is convolved with TPSFs to account for the diffusive nature of light propagation resulting in a compound curve being measured and to be called here a fluorescence TPSF (FTPSF).

Lifetime is important for imaging. It is widely used in microscopy (fluorescence lifetime imaging (FLIM)) [39], because the lifetime of a fluorescent molecule is sensitive to the biochemical environment in tissues (e.g., the pH) and to chemical binding. It thus offers a high potential for in vivo molecular imaging. Of particular interest is Förster resonance energy transfer (FRET) [40–42] whereby two fluorescent molecules, a donor and an acceptor interact, with the donor’s emission spectrum overlap with the acceptor’s absorption spectrum. As a donor gets in proximity to an acceptor, its fluorescence will get absorbed (“quenched”) by the acceptor, with quenching getting more efficient the shorter the distance between the two. This leads to a modulation of the donors fluorescence lifetime that depends on the donor-acceptor distance. More precisely, the transfer rate of energy between
the donor and the acceptor varies as the inverse of the 6th power of their distance. The effect of the increase of the transfer rate from the donor to the acceptor is to decrease the donors fluorescence lifetime as the donor has access to a faster channel to return back to its ground state. This effect can notably be used to study the conformation of molecules [43]. FRET has of course an influence on intensity-based measurements because if quenching is stronger, the light signal will be weaker. However, intensity-based measurements are also dependent on the fluorescent probes concentration, thus confounding quenching with concentration, which is not the case with lifetime measurements. Another application of FRET is in the study of interactions between biomolecules (e.g., protein-protein interactions or antibody labeling). With one type of biomolecule labeled with a donor and the other type with an acceptor, the donor’s lifetime will be different depending on whether the molecules interact or not (when interacting, they will be a shorter distance apart). Of course not all molecules will interact, and thus the overall signal will be formed by two decaying exponentials. Fitting measured signals with a double-exponential model allows determining the lifetimes associated with the interacting and noninteracting molecules along with the fraction of interacting molecules, which is not possible with intensity-based measurements. As a note, FRET allows the design of smart probes that are normally in the OFF state (no light emission) but start emitting light (ON state) when cleaved by a specific enzyme [22]. In this case intensity-based measurements suffice.

3. Challenges and Requirements in TD-DOT

Typically, to carry out TD measurements, one shines an ultrashort excitation light pulse (ideally approximated by a Dirac delta function in time and typically output by an ultrastate laser) on the meadium of interest and measures the pulse’s temporal shape after its propagation through the medium. In effect, one measures the transfer function of the medium of which Figure 4 exemplifies a case. Such measurements can be performed with a certain number of means depending on the characteristics of the medium, which determines the amplitude and the time scale of the duration of the pulse to be measured. In biological tissues, an ultrashort pulse will typically temporally broaden by about 1ns FWHM per centimeter of tissue traversed. For instance, laboratory mice, of interest in our work, typically have a diameter of about 2.5 cm and thus typical measured pulse durations are on the order of 2.5 ns. To resolve such signals, which can be modeled as the convolution of the true signal from the process to be measured with the system’s IRF, the system must have a very high temporal resolution on the order of 100 ps or less in order not to distort (smear out) the signal too severely and lose information. Mathematically, the true signal can be recovered by deconvolving the IRF out of the measured signal, but since deconvolution is an ill-posed problem, it is advised to keep distortions to a minimum with deconvolution, then giving better results [44, 45]. Furthermore, an excitation pulse is attenuated by several orders of magnitude after its propagation through a biological tissue due to absorption and owing to its spreading in all directions as a consequence of scattering. Thus, imaging instrumentation shall be able to measure very faint optical signals (a few picowatts are typical). As regards fluorescence measurements, fluorophores used in optical imaging typically have fluorescence lifetimes ranging from 500 ps to 2-3 ns. Again, resolving exponential decay curves with such lifetimes requires temporal resolutions down to 100 ps or less. In view of FRET measurements, discriminating nearby fluorescence lifetimes may call for even finer temporal resolutions. Fluorescence signals emanating from biological tissues also tend to be extremely faint. All these requirements (very short and weak signals) are a natural setting for the TCSPC technique.

As mentioned in the introduction, a further requirement of DOT is that measurements need to be made at a plurality of positions around a medium on its boundary to obtain an interior image thereof. Several configurations depicted in Figure 5, are possible for achieving such multiview detection. This, however, poses significant challenges both to the optomechanical design of a DOT scanner and to TCSPC because one needs several detection channels working in parallel to carry out the measurements in reasonable time. Recent advances in parallel multichannel TCSPC make it possible nowadays to carry out such multiview detection at high temporal resolution and exquisite sensitivity. This will be discussed further later, and it will be seen that multichannel TCSPC is a critical technology for the success of TD-DOT.

An important feature of TCSPC is that it allows measuring full TPSFs, which we believe is crucial in exploiting the full potential of TD measurements in image reconstruction algorithms to improve the spatial resolution and quantification achieved by DOT [38]. Moreover, measuring full FTPSFs is a necessity for extracting lifetime.


The TomOptUS group has developed a first-generation prototype of a TD-DOT scanner for small animal imaging. The purpose here is not to go through all the details of that scanner as this has already been done elsewhere [26], but rather to provide an overview of its architecture and hardware and review some recent results obtained therewith.
This will lead us in the next section to the needs for the realization of second-generation scanners and the work we have undertaken towards that end.

To put this into a better perspective, we first now briefly review other approaches and systems that have been developed for small animal time-resolved (i.e., time-domain and frequency-domain) DOT imaging (the discussion on systems will be restricted to those for tomographic 3D imaging and excludes systems solely for bioluminescence tomography, as bioluminescence provides by nature a CW signal). The advantages of TD measurements over CW measurements have been previously discussed as regards information richness. By performing frequency-domain (FD) measurements, it is in principle mathematically possible to obtain the same information as with TD measurements since both are related by a Fourier transform. FD systems resort to a modulated light source to measure the modulation index and phase shift between the light injected into a tissue and the light emerging therefrom after its propagation therethrough. Such systems are limited in small animal imaging owing to the small tissue volumes involved requiring source modulation at frequencies above 1 GHz to obtain sufficient phase shift contrasts [47]. Modulating light sources at such high frequencies is technically difficult and is limited to around 2 GHz (this is what is usually found in the literature). It is thus in practice impossible to measure the full Fourier spectrum and hence get FD data that is equivalent to TD data; some information is lost. This is a reason why TD measurements better allow reaching the full information that can be obtained from measurements on a tissue with light. This is in fact an aspect that motivates our work.

Notwithstanding the previous discussion on TD versus FD and CW measurements, owing to lower costs, most small animal imaging systems resort to CW measurements. Noteworthy is that commercial systems with tomographic capabilities exclusively resort to CW measurements [48–51]. An in-depth broad review of commercial scanners has been provided by Leblond et al. [23]. Regarding research prototypes, most acquire CW data [52–60] and a few are FD systems [56, 61–64]. Interestingly, a study comparing CW and FD systems and image reconstruction suggested that FD measurements may result in superior images but at much higher costs and efforts as regards instrumentation and computational resources [65].

Regarding systems exploiting TD signals, which is the focus here, the most important characteristics are the temporal resolution and the acquisition time. Acquiring TD data requires more time than CW data and that is a major issue. Another issue is the cost of TD instrumentation as will be discussed later. Temporal resolution can be characterized by the full-width at half maximum (FWHM) of the instrument response function (IRF) which is measured by directly injecting light from the system’s pulsed laser source into a detection channel.

Our systems described herein, through the free-space optics design for noncontact measurements, the short pulse laser along with the detectors, and TCSPC electronics used show unprecedented temporal resolution for multiview dual-wavelength TD measurements: IRF ≈ 200 ps FWHM on average for the first-generation scanner (Section 4.5 provides more details on this) and IRF ≈ 55 ps FWHM on average for the second-generation prototype (Section 5.4). For comparison, an 8-channel system with an IRF of 260 ps using TCSPC, MCP-PMTs, and optical fibers has been reported [66]. In other work, system of 5 dual-wavelength detection channels has been developed to be used in conjunction with a commercial X-ray CT scanner [67]. Th detection channels use optical fi ers and red-enhanced PMTs, achieving a temporal resolution of ≈465 ps. To our knowledge, the only TD system that has surpassed the temporal resolution of our first-generation system was a single view single detector transillumination noncontact free-space optics system [36] (IRF 163 ps FWHM) resorting to a supercontinuum pulsed fi er laser (pulse width ≈ 30 ps) and TCSPC with preselected PMT modules for temporal resolution (150 ps). This system, which measures a single wavelength at a time, is an adaptation of an earlier system that used a time-gated intensifie CCD (ICCD) camera with 200 ps gate widths and 25 ps temporal step sizes [68, 69]. Similar transillumination ICCD systems have been developed for fluorescence lifetime tomography with minimum 300 ps gate widths [70] or a reported IRF of 300 ps FWHM [71]. The systems cited thus far resort to sequential pointwise laser illumination scanned over the surface of the animal to inject light at different locations as needed for tomography. Such scanning is generally time-consuming.

An alternative is to use wide fi ld illumination with different spatially modulated patterns of light sequentially projected onto the surface of the object to be imaged and to measure the transmitted spatial light pattern [72]. This has been adapted [73, 74] for TD small animal imaging using a short pulse laser, an array of micromirrors to generate the illumination patterns, and a time-gated ICCD camera for measuring the temporal signals (200–300 ps gate widths and 25–40 ps temporal step sizes).

To further highlight the unique features of our systems compared to others, ours (1) possess the most dual-wavelength channels that can acquire signal simultaneously, this with the best temporal resolution leading to less signal distortion and (2) are able to perform both backscatter and transillumination measurements.

### 4.1. Architecture Overview

The first-generation scanner, depicted in Figure 6(a), provides at-a-distance measurements without contact with the animal. It comprises 7 dual-wavelength detection channels separated by 40° for both intrinsic and fluorescence multiview parallel detections. Figure 6(c) shows one such channel (optical design was carried out with the Zemax™ optical design software, whereas mechanical design of the scanner was carried out with the SolidWorks™ CAD software). Th channels are mounted on a turntable allowing detection at all angles around the subject, except for a small blind range of ±15 around the laser beam (Figure 7). Th configuration allows both epi-illumination and transillumination, along with intermediate cases (refer back to Figure 5).
Figure 6: Drawing and photographs of the first-generation TomOptUS scanner (a, b); drawing of one dual-wavelength detection channel (c).

Figure 7: First-generation scanner. Drawings of top and side views.
In this generation, the subject is imaged at the vertical. It can be translated vertically using a motorized translation stage and can also be rotated independently from the detectors around the vertical axis so that the light beam from a laser can be injected anywhere around it. This configuration reduces hardware complexity, which was important to develop a first prototype, compared to a configuration where the subject would be imaged lying horizontally as the latter would require the laser beam to be rotated around the subject. Vertical rotation avoids physiological changes and anatomical displacements within the subject due to gravity.

The scanner is endowed with two cameras for 3D stereo computer vision to measure the subject’s outer surface shape. This is necessary in a noncontact architecture since the surface determines where light propagation changes from diffuse (in biological tissues) to straight line (in air). This allows fixing the boundary conditions to be obeyed by the forward model used to compute the light distribution inside the subject, an essential part in any DOT image reconstruction algorithm. In distinction to other scanners, the TomOptUS scanner uses the same laser beam for both surface and DOT measurements. This allows simultaneous surface and DOT measurements, cutting down on the acquisition time, while reducing system complexity. Additionally, this allows measuring the exact positions where laser light is injected into the subject. These are also required by any forward model used in DOT image reconstruction [75].

4.2. Electrooptics Hardware for TCSPC. To perform TCSPC measurements, the excitation light source used is a mode-locked Ti:Sapphire laser (Tsunami, Spectra-Physics, USA) emitting ultrashort light pulses (pulse width 4 ps FWHM) at a repetition rate of 80 MHz. This is an ideal light source for a prototype scanner since its emission wavelength can be tuned in the range from 700 to 1000 nm which is best suited for imaging biological tissues as mentioned earlier and also since it provides ample power at any wavelength in a high quality TEM00 mode circular beam. The laser’s tunability also allows high flexibility to image different NIR fluorescent probes. In our system, laser beam steering is carried out as much as possible with reflections of optical components to minimize temporal distortion.

On the detection side, as mentioned previously, high sensitivity and high temporal resolution are needed for measuring the short and faint TPSFs typical of small animal imaging. This calls for TCSPC, which is a technique exactly suited to such conditions [76]. Specifically, we use an SPC-134 stack of 4 PCI computer plug-in cards (Becker&Hickl GmbH (bh), Germany), with each card achieving an electrical temporal resolution of 6.6 ps FWHM. The 14 detectors used are PMC-100-20 PMT modules (bh) with a spectral response curve in the wavelength range from 300 to 900 nm, having a nominal transit time spread (TTS) of $\approx$ 180 ps FWHM (thermoelectric coolers (TECs) are incorporated to reduce dark counts), and with a photocathode diameter of 8 mm.

The reference trigger electrical pulses for the TCSPC cards are derived from the laser pulses with a fast PIN photodiode (PHD-400 N, bh) connected to the STOP input of the TCSPC cards operated in reversed START-STOP mode. Four routers (HRT-41, bh), one for each card, allow connecting up to 4 PMT modules to each card for a maximum of 16 detection channels. An electronics box developed in-house allows operating up to 16 PMT modules: it supplies and controls power to the PMT modules and their TECs. The box also provides hardware shutdown of the modules to prevent overload light fluxes from damaging the PMTs. Careful optomechanical design of the detection channels prevents optical reflections within them, thus avoiding the TPSFs from being contaminated by such reflections.

4.3. Software. The custom software for operating the scanner, initially developed in LabVIEW® [26], has completely been redesigned and rewritten in the C++ language for better structure and greater flexibility. Open-source tools have been used throughout: NetBeans and the gcc compiler for Windows for software development and Qt for the graphical user interface (GUI) and its libraries for communications (USB, serial) with the hardware. The GUI allows configuring the scanner’s internal hardware settings (TCSPC parameters, number of active detectors, etc.), along with parameters related to an imaging session (ranges and step sizes of subject motion, TCSPC collection time, and laser beam attenuation). The software allows completely automated operation of the scanner during experiments along with complete manual control, which is useful for development and testing purposes and for preliminary data acquisition for setting up an experiment. The GUI also displays the following information in real time as measurements are in progress: actual position of subject and detectors, elapsed time of experiment, photon counting rates, and TCSPC histograms (TPSFs).

4.4. Results Obtained with the First-Generation Scanner. The first-generation scanner has recently allowed developing entirely new techniques for imaging the scattering properties of turbid media mimicking biological tissues [33] and for localizing fluorescent point-like inclusions therein [34]. These techniques critically rely on the high temporal resolution and exquisite timing accuracy and stability of TCSPC which is at the core of our scanner. They will now be briefly reviewed and illustrative results will be presented.

4.4.1. Intrinsic Imaging with Diffuse Photons Density Wavefront Speed. As mentioned earlier, snake photons suffer few scattering events. Because of this, they can to a good approximation be considered to propagate along straight paths (called rays) diverging in all directions from where they originated, that is, where light was injected in the medium (Figure 8). The collective propagation of snake photons originating from a given point inside a medium or at the boundary thereof defines an expanding diffuse wavefront (a so-called diffuse photon density wavefront (DPDWF)) therein [33]. When the medium is homogeneous, this wavefront is spherical [34] but gets distorted in a nonhomogeneous medium [77].

The time of arrival of a DPDWF at a given point on the boundary, which corresponds to the arrival time of early
Figure 8: Planar section of a cylindrical medium partitioned with triangular elements. Light is injected at the red point and detection positions are shown in blue. For each measurement, a ray (dotted black) is traced between the corresponding detection position and the injection point.

 photons (so-called early photons arrival time (EPAT)), can be obtained from a measurement of the TPSF thereat. In fact, as snake photons arrive early, they contribute to the rising edge of a TPSF, and the time of arrival of that edge is taken as the time of arrival of early photons (i.e., the EPAT), which can be stably and reliably obtained from a measured TPSF using numerical constant fraction discrimination (NCFD) previously introduced by our group [78].

The EPAT obviously relates to the speed of propagation of the wavefront inside the medium. As we have shown with our scanner, the EPAT depends on the optical properties of a medium [26], and thus the speed of DPDWFs also depends on those properties since the speed is the distance traveled over the time of travel, with the latter being the EPAT.

The considerations lead to the possibility of reconstructing DPDWF speed maps in order to image the intrinsic properties of media. Such maps have indeed been obtained, showing sensitivity to variations in the optical properties of a medium [33]. Thus far, this has been achieved for variations in the scattering coefficient, and we are currently developing a similar approach for obtaining contrast to variations in the absorption coefficient as well.

Briefly, to obtain a speed map (full details can be found in a previous paper [33]), the medium is partitioned into small cells as depicted in Figure 8 (triangles are chosen for simplicity, but it could also be other shapes). Choosing the cells sufficiently small, the speed within each cell can be assumed constant. Consider a ray, such as one shown in Figure 8, defined by a laser injection position and a detection position (a so-called source-detector (SD) pair). The total time of travel of early photons along that ray (i.e., the measured EPAT) is given by the sum of the individual times spent in each cell intersected by the ray, with the time within each cell being simply the length of the ray intersecting the cell divided by the speed in that cell. Mathematically, this can be expressed as a sum extending over all cells of the partition as

\[ t_{EPAT}^i = \sum_{k=1}^{N_c} \frac{d_{ik}}{v_k}, \]

where \( N_c \) is the total number of cells in the partition, \( d_{ik} \) is the distance crossed by ray \( i \) within cell \( k \) of the partition, and \( v_k \) is the DPDWF speed within cell \( k \) (\( d_{ik} \) is set to 0 for cells not crossed by the ray). Measuring EPATs for a large number of different rays (i.e., SD pairs) leads to a system of linear equations such as (2) that can be solved in the unknowns \( 1/v_k \) which are then mapped back onto the partition to obtain a speed map. Figure 9 displays reconstructions of such maps from experimental data showing speed contrast where inclusions with scattering coefficient differing from that of the otherwise homogeneous background were inserted (see figure caption for details). The big advantage of this imaging approach exploiting EPATs is that since straight line propagation is considered, image reconstruction of a medium can be carried out slice by slice as in X-ray CT, in contrast to other approaches developed in DOT which necessitate considering the medium as a 3D entity. This makes reconstruction with this approach much faster and also relatively easy to implement as only linear equations need to be solved. Perhaps a disadvantage of this approach is that it relies on detecting snake photons and this may involve low signal-to-noise ratios. This is not a problem with our first-generation scanner owing to single photon detection capabilities and the high light collection efficiency of the detection channels thanks to the large photosensitive area of the PMTs used (8 mm diameter). As will be seen later, this may, however, degrade results with smaller diameter single photon detectors. This calls for optical design optimization as will be discussed.

4.4.2. Fluorescence Localization by Time-of-Flight Measurements. We have developed an algorithm for localizing in 3D a plurality of fluorescent inclusions located therein by using distance ranging. This is based on the measurement of fluorescence EPATs at different positions around a turbid medium (e.g., a biological tissue) following ultrashort laser pulse illumination at a given position, with this being repeated for several laser injection positions. Such an algorithm is relevant, for instance, to locate small tumor masses or metastatic cells targeted by a fluorescent labeling agent [25].

Referring to Figure 10(a), the time to obtain a fluorescence signal at a given detection position \( D \) following the injection of a laser pulse at another position \( S \) is given by (i) the time \( T_2 \) taken for the laser pulse (which becomes a DPDWF inside the medium) to travel from the injection point to the fluorescent point-like inclusion at \( P \) plus (ii) the time \( T_3 \) for the fluorescence wavefront pulse ensuing the excitation of the inclusion by the laser DPDWF to propagate from the inclusion to the detection position \( D \). For a given SD pair of laser injection and detection positions, the time \( t_{\text{meas}} = T_2 + T_3 \) is what can be measured by a time-domain system. Not considered are the other propagation times indicated by
Fig. 9: (a) Top view of phantom used in experiments for obtaining speed maps. Inclusions labeled A, B, D, and E of different sizes and at different locations can be inserted in the medium for obtaining regions with scattering coefficient differing from that of the otherwise homogeneous background. In the results given in (b) and (c), the background’s reduced scattering coefficient has value $\mu_s' = 5 \text{ cm}^{-1}$. (b) Speed map obtained with inclusion A2 having $\mu_s' = 2.5 \text{ cm}^{-1}$ and A3 with $\mu_s' = 10 \text{ cm}^{-1}$. (c) Speed map with inclusions B1, B2, and E6 having, respectively, $\mu_s' = 20 \text{ cm}^{-1}$, 2.5 cm$^{-1}$, and 2.5 cm$^{-1}$.

$T_0$, $T_1$, $T_4$, and $T_5$ in Figure 10(a) (the time for an electrical signal to propagate in the system’s cables, the propagation of light in air and in the optical channels, etc.), since these can be calibrated out as explained in [26, 34]. Measuring the time for several SD pairs allows localizing the inclusion as depicted in Figure 10(b). This relies on the fact that different detectors will not measure the same time, with the time for a given detector being dependent on the depth of the inclusion relative to that detector. More precisely, as Figure 10(c) shows, the time measured for a given SD pair defines an oval with

$$t_{\text{meas}} = t_{\text{LP}} + t_{\text{PD}} = \frac{d_{\text{LP}}}{v_{\text{ex}}} + \frac{d_{\text{PD}}}{v_{\text{flu}}},$$

where $v_{\text{ex}}$ and $v_{\text{flu}}$ are, respectively, the speeds of DPDWFs at the laser excitation and fluorescent wavelengths (should the speeds $v_{\text{ex}}$ and $v_{\text{flu}}$ be equal, then the oval would be an ellipse as the sum of the distances $d_{\text{LP}}$ and $d_{\text{PD}}$ would then be a constant). This oval define the locus on which the inclusion is to be found, since all points on such an oval will give the same measured time and this is the sole information available from a measurement associated with an SD pair. The arguments thus far have been expounded in 2D for simplicity, but they are equally valid in 3D whereby ovals become ovoids. In 2D, two ovals are in principle sufficient to find the position of an inclusion, with that position being the intersection of the ovals. In 3D, one needs to find the intersection of 3 ovals. In the case where multiple inclusions are present in the medium, then more ovals (ovoids) must be obtained, which amounts to have more SD pairs of EPAT measurements. In practice, we do not find intersections of ovals (ovoids) but rather trace all ovals associated with SD pair measurements and the regions of highest density in the maps thus obtained correspond to inclusions positions. This is shown in Figures II(a) and II(b) in the case of a simulation, where (a) displays a density map and (b) displays that same
Figure 10: (a) Times and delays involved in an EPAT measurement for a given source-detector (SD) pair. (b) Several SD pair measurements for localizing an inclusion. (c) The result of a given SD pair measurement constrains the inclusion to lie on an oval.

map as a height function from which maxima are extracted. Figure 11(c) shows a localization result for a 3-inclusion case (all inclusions in a plane, experimental data), and Figure 11(d) shows the localization of 4 inclusions at various heights in a 3D cylindrical medium, with a worst case error of 1.7 mm on the localization of one of the inclusions, which for diffuse optical imaging is excellent.

Further details and results on the localization technique summarized here are given in [34]. Considering that typical speeds for diffusive wavefront pulses are on the order of 1/10th the speed of light in vacuum and that this technique allows localizing inclusions with mm accuracy, it is seen that, for such a localization approach to succeed, the system providing the necessary measurements must have a very high temporal resolution and an extremely stable timing accuracy allowing extracting reliably the arrival time of such pulses. Our system has enabled such developments. Further features of the localization approach reviewed here are that (1) it is extremely fast since the operations to obtain an image are very simple and (2) it can localize a fair number of inclusions (we have achieved up to 5). Some shortcomings are that it is limited to finding the positions of point-like inclusions; it cannot image smooth distributions of fluorescence and it cannot localize a large number of inclusions.

4.5. Specifications. As mentioned earlier, the temporal resolution is a critical performance index for a TD measurement system. In the case of our scanner with the dual-wavelength channels at the two wavelengths at which it operates, we measured IRFs FWHM at 780 nm of 195 ps on average with variations from 116 to 285 ps, and at 830 nm this was 169 ps on average with variations from 114 to 256 ps.

To give an idea of typical acquisition times with our scanner, it takes ≈4.5 minutes per slice (corresponding to a given value of the height z of the subject) to acquire a set of measurements comprising a total of 144 measurements (36 laser injection points with an angular separation of 10°,
detector stepping of $10^\circ$ from 0 to $30^\circ$, and photon counting collection time of 2 s). Such typical acquisition times are very long as this is for only one slice. The bottleneck is the necessity to rotate the detectors. Furthermore, the current scanner is somewhat bulky and resorts to high cost TCSPC electronics, which makes scaling up the number of detectors difficult.

5. Towards a Second-Generation DOT Scanner

To acquire data in reasonable time for a live subject, measurements will need to be performed faster than what is currently possible with the first-generation scanner. Achieving this requires addressing two major aspects. First, it is necessary to increase the density of detection channels around the subject (i.e., more channels) to avoid rotating the detectors. This necessitates miniaturizing the detection channels by resorting to smaller optics and to replace the bulky PMT modules currently in use with smaller footprint detectors. Second, dealing with more channels requires replacing the expensive TCSPC electronics cards with few channels by more compact denser fully parallel TCSPC electronics at a reasonable cost, which is not currently commercially available.
Two enabling technologies, to which the SPADlab at PoliMi and its spin-off Micro Photon Devices S.R.L. (MPD) have been leaders in recent years, allow addressing these two major aspects. These technologies, discussed below, are (i) the development of single photon avalanche diodes (SPADs) based on planar silicon technology thin layer junctions [79] and associated front-end electronics circuits [80, 81] and (ii) the development of highly integrated electronics for implementing compact parallel TCSPC hardware [82, 83].

An alternative to higher detection density and parallel TCSPC could be to use wide field spatially modulated illumination as described in Section 4. Parallelism then comes from the camera and its array of detectors. Such approaches allow injecting more light into a tissue since it is spread over its surface, rather than being concentrated on a small spot (which then requires some upper power limit not to be exceeded so as not to damage the tissue). Wide field illumination thus appears at first sight to offer the possibility of performing measurements at higher signal-to-noise ratio (SNR) than with TCSPC, hence possibly allowing increasing the acquisition speed. However, CCD cameras are less sensitive than TCSPC which has single photon sensitivity. Also, with time-gated ICCD cameras, the gate must be scanned to different delays for resolving a temporal signal, which requires time. TCSPC in turn acquires full temporal signals by counting and binning the counts in a time histogram (no photons are lost, which makes it more efficient). It is thus not clear which approach in the end will be faster and can achieve higher SNRs. An important question in wide field illumination imaging is to determine the best patterns of light to use for a given imaging task and to reduce their number as much as possible. This can be related to compressive sensing and work on this is underway [84]. Finally, with current technology, the smallest gates are 200 ps wide, whereas SPADs can reach temporal resolutions down to 50 ps and even lower. It is thus seen that two competing technologies can be pursued for TD-DOT imaging, each having its advantages and disadvantages. We favor TCSPC with its higher temporal resolution which we believe will be an advantage in small animal imaging applications.

5.1 Advances in TCSPC Technologies

5.1.1 SPADs and Associated Front-End Electronics Circuits. TCSPC requires detectors that can deliver a measurable fast signal in response to the detection of a single photon. Photomultiplier tubes (PMTs) with their high gain have long been the workhorse in this area, but single photon avalanche diodes (SPADs) have become a serious alternative owing to their ruggedness, compact size, higher photodetection efficiency (PDE), and robustness to overloading light signals (whereas a PMT is severely damaged in such cases, a SPAD only saturates).

A SPAD is a pn junction reverse-biased at a voltage $V_A$ exceeding the breakdown voltage $V_{BD}$. It operates in Geiger-mode: due to the strong electric field within the junction, a single charge carrier injected into the depletion layer can trigger a self-sustaining avalanche multiplication process causing the current to rise abruptly (nanoseconds or subnanosecond rise time) to a milliamp range level. When the primary carrier (in fact an electron-hole pair) is generated by a photon, the front edge of the avalanche current pulse tags the arrival time of the detected photon. Since the current continues flowing into the junction until the bias voltage $V_A$ is lowered below $V_{BD}$, a suitable circuit, the so-called quenching circuit, is required. The latter also restores the bias voltage and the device is ready to trigger again.

SPADs. The milestones that led to the development of modern SPAD detectors as pertinent to the present work will now briefly presented (further details and references can be found in [80]). The first planar silicon technology SPAD structure suitable for monolithic integration was based on an early semiconductor diode structure devised at the beginning of the 1960s [85–87]. Further developments [88–92] aimed at improving device performance (reducing the duration and intensity of the diffusion tail which complicated analysis of fast fluorescence decay measurements, reducing the dependence on wavelength of the shape of the temporal response function of the detector, improving uniformity of the PDE over the active area, reducing the dark counting rate (DCR), and afterpulsing probability and their temperature dependence) led to the actual SPAD structure, consisting of a planar double epitaxial structure implemented in a custom technology [79]. Devices are now fabricated starting from an n-type substrate on top of which a p-type epitaxial layer, constituting the SPAD anode, is grown. Finally, an n-type, thin diffusion region that constitutes the cathode of the device is implanted. The p-type region has two purposes: to reduce the breakdown voltage in the central region of the device, avoiding premature edge breakdown and its doping profile is designed to suitably shape the electric field in the multiplication region, with the aim of optimizing device timing response [93]. Eventually, an isolation region can be incorporated, which in conjunction with the n-type substrate forms a well surrounding the device anode. The efore, by reverse-biasing the isolation region-anode junction, it is possible to electrically isolate the SPAD from adjacent devices and to develop monolithic SPAD arrays [94].

The double epitaxial technology, thanks to its compatibility with standard CMOS fabrication techniques, paved the way to further developments allowing directly integrating circuitry onto the detector chip. In the last years, several device prototypes, implemented in standard CMOS technologies, have been reported. However, these devices present a lower PDE [95, 96] or a much worse timing resolution [97], a significant optical higher DCR per unit area [98, 99], and a higher afterpulsing probability with respect to custom SPADs, forcing device designers to use smaller active areas and low excess bias voltages to limit detector noise. On the contrary, optimization of the electric field profile allowed by custom technologies leads to better performance in terms of timing resolution, DCR, afterpulsing (allowing the development of large area SPADs with diameters from several tens of microns up to few hundreds of microns [100]), and higher PDE (further improvements have been recently demonstrated by using a thicker absorption layer [101]).
Quenching and Resetting. Another important aspect when using SPADs is quenching the avalanche buildup after the detection of a photon [102, 103]. The main issues here are the speed at which quenching occurs, the hold-off time, and the reset transition.

Quenching should enter in action as fast as possible after an avalanche is triggered to reduce the avalanche charge. The latter is responsible for energy dissipation within the detector, which should be kept low to prevent excessive power dissipation and damage. It also leads to secondary light emission, giving rise to unwanted optical cross talk between detectors. Finally, as the avalanche current proceeds through the detector, some charge carriers are trapped and get released after a random time. Such carriers can trigger a secondary unwanted avalanche (afterpulsing).

To speed up quenching, both the delay $T_d$, with which the quenching transition starts after the avalanche and the transition time $T_f$, must be minimized. With passive quenching circuits (PQCs), the onset of the transition is immediate after the avalanche, but the transition time can be fairly long. Active quenching circuits (AQCs) can speed up this transition, but their feedback loop may introduce a significant delay.

The hold-off time, which is the time the SPAD spends in the quenched state, must be kept short to achieve high counting rates but should not be too short to leave enough time for trapped carriers to be released. Controlling the hold-off time is a feature that can be obtained with AQCs, but not with PQCs.

Following quenching, the SPAD needs to be reset. This requires the bias voltage to be brought back to its quiescent value above the breakdown voltage with a duration of at most a few nanoseconds. Indeed, during the reset phase, a degradation of the photon timing resolution and of the PDE, due to the variation of the SPAD overvoltage, can occur. This calls for minimizing the reset transition time so that the probability of detecting photons during this transition is kept to a minimum. PQCs inherently show exponential recovery, with typical reset time constants of a few hundred nanoseconds. AQCs on the other hand allow faster reset transitions. Owing to the respective advantages of PQCs (faster response to an avalanche) and AQCs (faster transitions and adjustable hold-off time), mixed active-passive quenching is nowadays the preferred approach [103].

Current Pick-Up. SPADs with timing resolutions in the few tens of picoseconds initially had very small photosensitive areas (typ. 10–20 μm diameter). The difficulty in developing optical systems with such small area detectors in terms of optical alignment and light collection led to efforts to develop larger area SPADs [79]. The timing resolution of planar SPADs, however, strongly depends on the diameter: in the very early part of the avalanche, carrier multiplication is confined within a small area and the current rises with small statistical fluctuations. The later part of the current rise corresponds to the progressive spatial spreading of the multiplication process and is subject to higher statistical fluctuations; hence, the larger the active area, the worse the timing performance. The trade-off between active area and timing resolution may be overcome by detecting the avalanche current during the initial part of its rise [104]. With an appropriate current pick-up circuit, a timing resolution down to 35 ps FWHM was demonstrated at room temperature with thin depletion layer SPADs having an active area diameter up to 100 μm [105]. Further work on current pick-up circuits for their integration near to the detector has made pixel architectures for SPAD arrays and multichannel TCSPC possible [81, 106].

5.1.2. Multichannel Parallel TCSPC Electronics. TCSPC relies on periodic excitation of a process (an example is fluorescence emission) with ultrashort pulses of light and measuring the time interval between the arrival of each single detected photon in response to the process (in our example this would be a fluoresced photon) and the excitation light pulse that gave rise to that photon in the process. Hence, each detected photon is time-tagged according to its measured time interval. At high pulse repetition rates in the range of tens of MHz such as what can be obtained with some solid-state and diode lasers, because of the typical dead time in the processing electronics and possibly in the detectors such as for SPADs, it is impossible to work at photon counting rates where several photons would be detected for each cycle of the periodic excitation. Even, the probability of detecting a photon per excitation cycle must be kept much smaller than 1 (a probability of ≈1% is typically used in practice). Such a low probability is necessary to make the detection of 2 or more photons in the same excitation cycle negligible since all photons arriving after the first one cannot be processed as the electronics are busy with the first.

The electronics are thus blind to photons arriving after the first within an excitation cycle, and so if multiple photons occur too frequently, the waveform to be measured gets distorted towards short times. Given that the condition of low probability is satisfied, accumulating a large number of time-tagged photons allows measuring a temporal waveform (such as a fluorescence decay curve) by forming a histogram of the number of detected photons in terms of their time tags. In effect, TCSPC is a means to statistically reconstruct a waveform that is then thought of as a probability distribution (the probability of occurrence of photons as a function of time for a given process).

Hence, apart from detectors and associated front-end electronics for detecting single photons, TCSPC requires processing electronics blocks to build the histogram. In one implementation, this is done with an analog time-to-amplitude converter (TAC) that converts the time interval between a START electrical pulse (corresponding to a laser pulse) and a STOP electrical pulse (obtained from the detection of a photon) into an analog voltage proportional to that interval [41, 107, 108]. This voltage is then converted to a digital number with a fast analog-to-digital converter (ADC), and this number is used to increment a counter in a bank of counters at the address corresponding to that number. Each counter thus implements a time bin whose address is proportional to time. After time-tagging and counting a large number of photons, the bank holds the
Fig. 12: Block diagram of the 8-channel PoliMi TCSPC module (DAC: digital-to-analog converter, TAC: time-to-amplitude converter, ADC: analog-to-digital converter, FPGA: field-programmable gate array, USB: universal serial bus, and PC: personal computer).

histogram corresponding to the signal to be measured. In another implementation of TCSPC, the time interval between a START pulse and a STOP pulse is directly converted to digital form with a device called a time-to-digital converter (TDC). There are several ways to implement a TDC, which exploit the transit time of the timing signal within a chain of logic gates and are typically based on Vernier delay lines [109] or on ring oscillators [110].

Applications of TCSPC require more and more multichannel detection capabilities to increase measurement throughput, and TD-DOT is one example. A relatively high number of detection channels (typically up to 32) can be handled with modules in the form of cards that can be inserted in mother board slots of computers, as the ones used in our 1st-generation DOT scanner. Such modules exhibit high performances in terms of timing resolution (a few picoseconds), linearity, and dynamic range (hundreds of nanoseconds). However, their limited parallel signal processing capability, large size, high cost, and excessive power consumption prevent further scaling-up of the number of detection channels. Thus, integrated circuit solutions [82] able to reach all the required timing characteristics and ensuring cost and size reduction [111] are mandatory for multidimensional imaging systems.

5.2. Eight-Channel TCSPC Module. To fulfill the requirements of modern multichannel TCSPC applications, an 8-channel TCSPC module prototype, which can eventually be easily scaled up, has been developed by the PoliMi group. Th module is based on an intrinsically multichannel structure, whose detailed description can be found in [83].

Th block diagram of the 8-channel module is depicted in Figure 12. An interface board is used to adapt the single-ended external timing signals in NIM standard analog format (eight START pulses and a common STOP pulse) to the differential TAC input signals, provide the external power supply to the TCSPC board (the module operates with a single DC power supply delivering from 8 V to 16 V), and perform temperature control (regulating the rotation speed of two fans). Th TCSPC electronics themselves are implemented on a 95 × 40 mm eight-layer printed circuit board (PCB) to provide several power planes and avoid electrical cross talk between the analog signal conditioning stages and the digital processing blocks.

Th TCSPC board is based on two 4-channel integrated TACs directly driving a commercial 8-channel ADC. An on-board FPGA processes the ADC data outputs and the computed values are used to build the TCSPC histograms into the FPGA internal memory. Finally, data are transferred to a PC through a USB link, used also as an interface for remote control. The two integrated TAC arrays, each including four independent converters, are built in 0.35 μm Si-Ge technology. Th 4-channel TAC is an improved version of a converter previously developed at PoliMi [82] which was itself based on a single-channel prototype [112]. Each channel inside the TAC array features a variable full-scale range (FSR) that can be chosen among four values: 11 ns, 22 ns, 45 ns, and 88 ns. To maximize system flexibility, the FPGA can digitally set this parameter, even when a measurement is running. To reduce the board dimensions, the TAC outputs were designed to match the ADC input dynamic range, so that no external operational amplifiers are required.

To keep dimensions compact, the AD9252 A/D converter (from Analog Devices) was chosen, as it features eight independent channels, a 50 MSPS conversion rate, and a 14-bit data output. Although it satisfies most of the TCSPC requirements, its ±40% LSB differential nonlinearity (DNL) is too high. To obtain an appropriate DNL, we decided
to implement the dithering technique. A random dithering signal is added to the TAC output voltage through a D/A converter and subsequently subtracted from the digital value resulting from the ADC conversion. The dithering technique implemented here is known as sliding scale [113], and we have previously demonstrated its effectiveness in reducing the overall DNL [114]. To minimize dimensions, the adder stages and the integrated D/A converter (based on a 10-bit resolution current-steering segmented architecture [115,116]) are included in the TAC chip too.

As already described, the ADC outputs are acquired, processed, and stored into the FPGA (Spartan-6 XC6SLX150T from Xilinx) internal memory. At power-on, the FPGA reads the contents of an on-board FLASH memory programming itself automatically, thus making the module ready to use. The FPGA samples and deserializes the eight ADC serial outputs; each deserialized result is stored in a dedicated internal FIFO memory to decouple the sampled data stream from that for data processing. The FIFO output, after dithering compensation, corresponds to the histogram memory cell whose value can then be updated.

The eight histograms are stored in eight dual-port RAMs divided into 2^{14} cells, each having a 32-bit depth. Measurement data are updated by writing into one port and can be simultaneously sent to the PC by reading the memory from the second port (also used to erase the RAM content). Histogram data, along with additional information useful to characterize the measurement, are exported to an external PC using the FT2232H Hi-Speed 2.0 USB transceiver from FTDI.

Photographs of the complete 8-channel TCSPC module are shown in Figure 13. The interface board has been tailored to be enclosed in a small aluminum case (160 × 125 × 30 mm³). The TCSPC board itself is shaped to be connected into an 80-pin connector. This way the board can be included in the stand-alone 8-channel module as presented or in a system featuring more channels by simply parallelizing more boards [111].

The complete module has a power consumption of about 6 W and a time bin width ranging from 0.8 to 6.4 ps depending on the selected TAC FSR. Several experiments were conducted on the TCSPC system to evaluate three critical parameters: the time resolution, the DNL, and the cross talk between channels. Besides these, it is important to evaluate the conversion rate of the channels, which depends on the maximum TAC-ADC dead time. A value of 1 D was measured for the latter and, considering a maximum ADC conversion delay of 88 ns, a total time of approximately 200 ns for a single conversion is obtained. Hence, the system reaches a useful maximum conversion rate of 5 MHz per channel.

To thoroughly characterize the module, all channels and all FSRs were independently tested, but, for brevity, only the time resolution and DNL values for one channel for the 45 ns full-scale range are reported in Figure 14 more details on the performance and on the set-ups used to carry out the measurements can be found in [83, 111]. The timing resolution that can be obtained depends on the selected FSR. As Figure 14 shows, it is 51 ps FWHM (considering the value at mid-FSR) for the 45 ns FSR and, in general, it improves as the FSR is reduced: switching from the 88 ns to the 45 ns FSR, an increase by a factor of two can be observed, with a measured FWHM resolution scaling from 110 ps down to 55 ps. As the FSR is further reduced, the resolution improvement factor is lower: the timing resolution for the 22 ns FSR is 30 ps and, for the shortest range, it decreases to 18 ps. On the contrary, the peak-to-peak DNL value is constant and extremely good for all FSRs, being lower than 4% of the time bin width for almost all time delays. Finally, the cross talk, measured in the worst possible operating conditions [83], resulted in a peak-to-peak disturbance lower than 6% of the time bin width, extremely weak, and completely negligible in an actual TCSPC acquisition.

5.3. A Second-Generation DOT Scanner Test Prototype. In 2007, the TomOptUS group demonstrated the feasibility, for a single detection channel, of resorting to SPAD detectors to replace more bulky PMTs for DOT measurements [117]. This showed that it was possible to obtain sufficient signal with such detectors despite their much smaller photosensitive area (50 μm diameter compared to typ. ≈8 mm for PMTs). Since then, other groups have also investigated the possible use of SPADs for diffuse optical imaging measurements [118–121]. We have now gone one step further by modifying our 1st-generation scanner described above to integrate SPAD detectors and the compact 8-fully-parallel-channel PoliMi TCSPC module described above. This is to demonstrate to a larger scale the feasibility of miniaturizing the detection
channels and establish design criteria in view of developing the second-generation scanner with such technologies. This is a first step towards the realization of a high detection density virtually parallel TD-DOT scanner, whereby each detection channel will have its own electronics rather than resorting to routers which introduce interchannel cross talk and which also further limit the possible photon counting rate.

More specifically, in the fluorescence channels of the first-generation scanner, we replaced the PMTs by SPAD detectors (PDM module, 50 μm active diameter, <50 dark counts/s with thermoelectric cooling, and 50 ps nominal timing resolution but our modules had better performances with timing resolutions averaging at 42 ps) from Micro Photon Devices (MPD, Bolzano, Italy) and adapted the optomechanics (lenses and lens tubes) accordingly due to the smaller photosensitive area of SPADs compared to that of PMTs (50 μm versus 8 mm) (Figure 15). The PoliMi 8-channel TCSPC module replaces the Becker&Hickl TCSPC cards for the fluorescence channels. The scanner’s software has also been modified to acquire the data from the PoliMi TCSPC module. We chose to establish feasibility with the fluorescence channels, since fluorescence signals are much weaker than intrinsic ones, and thus feasibility with fluorescence detection warrants feasibility for the intrinsic case. Furthermore, for small animal imaging, fluorescence imaging is of greater interest than intrinsic imaging.

5.4. Results: A Comparison with the First-Generation Scanner. We have first characterized the temporal resolution of the fluorescence channels with the SPADs and the 8-channel PoliMi TCSPC module. The temporal resolution directly depends on the timing resolution of the TCSPC electronics and SPADs. We operated the TCSPC module with the FSR set at 22 ns for which the timing resolution is 30 ps as previously
Figure 16: Fluorescence localization results for experimental data obtained with SPADs. (a) One inclusion. (b) Three inclusions in a plane (2D case). (c) Three inclusions in 3D.

mentioned. We obtained an average IRF FWHM at 830 nm over all 7 channels of 54 ps with variations from 49 to 62 ps. This is a threefold decrease compared to the PMT channels of the first-generation scanner.

We then performed measurements of fluorescence EPATs with the SPADs and the 8-channel PoliMi module to reconstruct fluorescence inclusions localization maps following the approach presented in Section 4.4.2. Figure 16 shows representative results for the localization of 1 and 3 inclusions in a plane (2D case) and 3 inclusions in 3D. It is not the purpose here to go into a detailed analysis of these results but, generally, localization errors are larger when using SPADs compared with those obtained with PMTs (for which Figure 11 gives a representative example). This can be traced back to the fact that signals with SPADs are weaker with the current optomechanical channel configuration compared to those obtained with PMTs under the same conditions (same photon counting collection time). This calls for an optomechanical rework of the detection channels to obtain a design optimized for SPADs so that light collection can be increased; presently the current versions of the channels for SPADs are mere adaptations of the channels designed for PMTs. Such optomechanical rework is a major challenge as smaller optics will need to be resorted to in order to achieve denser channel packing, while increasing the light gathering power. Furthermore, working with small optical
elements poses more stringent requirements on alignment and thus on tolerancing. These are major practical issues that will need to be addressed. Overcoming such difficulties will be a very significant step forward in TD-DOT technology towards reducing data acquisition time, which is a major bottleneck as discussed earlier.

Figure 17 shows typical EPAT projections (also called “sinograms” [34]) acquired with PMTs and SPADs. Such projections serve as data for obtaining localization maps. A given EPAT projection corresponds to the set of EPATs measured for a given point where laser light is injected into the medium; that is, the location of the source (S) is held fixed. As can be seen, EPAT projections obtained with the SPADs are noisier than those obtained with PMTs. This explains why localization errors are larger for the data acquired with SPADs compared to PMTs. Again, this is not a problem with the SPAD detectors but rather with the nonoptimized current optomechanics.

The results presented in Figure 16 clearly indicate that imaging with SPADs to replace PMTs is possible, and we may hypothesize that, with appropriate optomechanical design, the performances of SPAD channels should approach or equal those of PMT channels. The same conclusion is also motivated by past investigations carried out in our laboratory [17]. This is the reason why we are pursuing the development of our scanner with SPADs.

6. Summary, Conclusions, and Current Work

In this review, we provided background on the interaction of light with biological tissues to motivate the advantages for resorting to TD measurements in terms of the higher information content in the signals thus acquired. The characteristics of such signals represent challenges related to their measurement. This leads to requirements on the instrumentation necessary to acquire them without disturbing their integrity as much as possible. We then went on to present instrumentation that we developed for small animal DOT imaging (first-generation TD-DOT scanner) based on PMT detectors and TCSPC cards which fulfills these requirements; that is, it is able to acquire the signals and preserve their integrity, but it does not do so in sufficiently short time for practical purposes in small animal imaging. This led us to investigate TCSPC technologies that would allow overcoming this difficulty, which is SPAD detectors and fully parallel multichannel highly integrated electronics for TCSPC. We developed a prototype scanner based on these technologies and demonstrated that they will allow developing next-generation TD-DOT scanners, as we were able to obtain results similar to those with PMTs and TCSPC cards. There is, however, a challenge in this endeavor, which is to optimize detection channels optomechanics for SPADs with their “microscopic” photosensitive areas in order to bring the amplitude of the acquired signals to levels comparable to those obtained with PMTs and their “macroscopic” photosensitive areas. This is the real challenge before us along with further electronics integration to be able to deal with at least on the order of 64 or more detection channels in parallel. In this way, it will become possible to carry out TD tomographic data acquisition in reasonable time without mechanical movement of detectors, which is currently the limiting factor to increase acquisition.
speed. This is the purpose of the work undertaken here and currently pursued in our laboratories.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

Yves Bérubé-Lauzière acknowledges financial support from FRQ-NT (Projet de Recherche en Equipe), NSERC (Discovery Grant), and CFI (New Opportunities Fund) for this work. Yves Bérubé-Lauzière is member of the FRQ-S-funded Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke (CR-CHUS).

**References**


