Towards Next-Generation Time-Domain Diffuse Optics for Extreme Depth Penetration and Sensitivity

Alberto Dalla Mora$^{1,†}$, Davide Contini$^{1,†}$, Simon Arridge$^2$, Fabrizio Martelli$^3$, Alberto Tosi$^5$, Gianluca Bosio$^4$, Andrea Farina$^{5,7}$, Turgut Durduran$^6$, Edoardo Martinenghi$^1$, Alessandro Torricelli$^{1,8}$, Antonio Pifferi$^{1,5}$

$^1$Dipartimento di Fisica, Politecnico di Milano, Italy
$^2$Department of Computer Science, University College London, United Kingdom
$^3$Dipartimento di Fisica e Astronomia, Università degli Studi di Firenze, Italy
$^4$Dipartimento di Elettronica Informazione e Bioingegneria, Politecnico di Milano, Italy
$^5$Istituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche, Italy
$^6$ICFO-Institut de Cències Fotòniques, Spain
$^†$These authors contributed equally to this work

Abstract: Light is a powerful tool to non-invasively probe highly scattering media for clinical applications ranging from oncology to neurology, but also for molecular imaging, and quality assessment of food, wood and pharmaceuticals. Here we show that, for a paradigmatic case of diffuse optical imaging, ideal yet realistic time-domain systems yield more than 2-fold higher depth penetration and many decades higher contrast as compared to ideal continuous-wave systems, by adopting a dense source-detector distribution with picosecond time-gating. Towards this aim, we demonstrate the first building block made of a source-detector pair directly embedded into the probe based on a pulsed Vertical-Cavity Surface-Emitting Laser (VCSEL) to allow parallelization for dense coverage, a Silicon Photomultiplier (SiPM) to maximize light harvesting, and a Single-Photon Avalanche Diode (SPAD) to demonstrate the time-gating capability on the basic SiPM element. This paves the way to a dramatic advancement in terms of increased performances, new high impact applications, and availability of devices with orders of magnitude reduction in size and cost for widespread use, including quantitative wearable imaging.

© 2015 Optical Society of America

OCIS codes: (170.5280) Photon Migration; (290.7050) Turbid media; (230.6046) Smart pixel systems; (170.3660) Light propagation in tissues; (170.6960) Tomography.

References and links


1. Introduction

A widely-used method to non-invasively investigate scattering media, such as biological tissues [1, 2], is to use a continuous-wave (CW) source within the 600-1100 nm range, and detect light re-emitted at some distance $\rho$ from the injection point. Upon increasing $\rho$, deeper structures are probed. Such approach is severely hampered by the interplay between light absorption and scattering causing strong attenuation and impaired spatial resolution. The adoption of a time-domain (TD) approach introduces one more dimension -- the photon time-of-flight $t$ -- that provides added benefits, such as uncoupling absorption from scattering [3], and probing a medium in depth by exploiting long-lived, deep-travelling photons [4]. Further, we have shown that a null-$\rho$ approach optimizes light confinement, resulting in optimal signal level, contrast and spatial resolution [5], provided that an efficient gating mechanism extracts the few late photons out of an overwhelming burst of early photons [6].

The real limitations of TD as compared to CW systems are the cost, complexity and poor signal-to-noise ratio. Originally, large table-top lasers combined with streak-cameras were employed [7, 8]. More recently, pulsed diode lasers or fibre-based sources together with compact time-correlated single-photon counting (TCSPC) led to rack instruments, compatible with clinical use [9–13]. Still, the number of optodes is quite limited and cumbersome fiber bundles are needed for light collection. Gated intensified cameras are suitable for parallel detection [2, 7, 8] yet poor gating suppression and dynamic range are insufficient for the null-distance approach [2]. Alongside, new detection strategies were proposed for e.g. use of a simple gated counter [2] instead of the whole TCSPC electronics or of compressive sensing approaches for optimal information gathering [2]. For what concerns the sources, external laser modules were always used, most often with fiber delivery, apart from few non-contact implementations [2, 7, 8]. So far, we found no attempt to embed the pulsed source directly onto the probe.

In this paper we lay the foundations of next generation TD diffuse optics. In particular, we identify on simulations the exciting breakthrough in depth penetration and sensitivity reachable when pushing the technology to the limit. Then, we demonstrate the first basic steps needed to go in that direction, which are 1) use of a pulsed source embedded onto the probe (VCSEL); 2) use of a large area detector embedded onto the probe (SiPM); 3) demonstration of the fast-
gating of the basic element of a SiPM (gated SPAD) so as to permit null source-detector distance measurements. Compared to our previous achievements with fast-gated SPADs [2, 21], here the novelty is on the demonstration of the null-distance approach with a new design of the device, suitable with standard CMOS SiPM construction, and direct integration onto the probe, as key requirements towards the long-term scenario prefigured by the simulations.

In the following, we will first study with simulations the extreme limits accessible with an ideal, yet realistic TD approach, and identify the key conditions required to reach this goal. Then, we present the new components designed to address those specified conditions, and show a first validation on phantoms using the NEUROPT protocol for performance assessment of time-domain diffuse optical imagers [14]. Further, we propose a first example of in-vivo use on brain functional imaging. Finally, we draw a possible path for further advancements and present the potential high impact in the medium-long run.

2. Materials and Methods

2.1. Simulations

Simulations were generated with a forward solver accounting for the effects of an absorption perturbation placed inside a diffusive slab. The Born approximation [15] was applied to the photon diffusion equation solved with the extrapolated boundary condition [16]. As background optical properties we assumed a reduced scattering coefficient $\mu'_s = 10 \text{ cm}^{-1}$, an absorption coefficient $\mu_a = 0.1 \text{ cm}^{-1}$, and a refractive index $n = 1.33$.

Diffuse Optical Tomography was simulated on a cubic portion of a semi-infinite homogeneous diffusive medium (60x60x60 mm$^3$, $\mu_a = 0.1 \text{ cm}^{-1}$, $\mu'_s = 10 \text{ cm}^{-1}$, n = 1.33). Sources and detectors, both having an area of 10x10 mm$^2$, were placed on the surface and completely covered the portion surface. A cubic absorptive perturbation (10x10x10 mm$^3$) was placed at 40 mm depth with an absorption coefficient twice the background value.

Concerning TD data, a practical maximum limit for single-photon counting was fixed at 1 MHz/mm$^2$ assuming each detector made up of 10x10 single pixels of 1 mm$^2$ each of them respecting the maximum count rate of 1 MHz typical of time-correlated single-photon counting systems. The limit was distributed over the number of temporal windows considered (100 windows of 100 ps), thus implying a maximum count rate of 10 kHz/mm$^2$ for each time window. CW data were not limited in terms of maximum count rate since different detection strategies can be used. All data were subsequently perturbed by additive Gaussian noise ($\sigma = 0.1\%$) representing fluctuations of light sources, detector, electronics and other systematic errors. Finally, Poisson noise was further added to take into account shot noise.

Reconstructions were performed using the Generalized Minimal Residual Method [17] applied, with a maximum of 150 iterations and $10^{-6}$ of tolerance, to the regularized normal equation

$$\left(J^T J + \lambda^2 I\right) \Delta \mu_a = J^T y$$

where $J$ is the Jacobian matrix, $\Delta \mu_a$ is the variation in absorption for each voxel and $\lambda$ is the regularization parameter set to 1 [18]. The volume was divided into cubic voxels of 2.5x2.5x2.5 mm$^3$.

2.2. Experimental Setup

The schematics for the two devices (PROBE1 and PROBE2) used in the experiments are shown in Fig. 1. In both cases, the source and the detector are hosted onto the probe. The source is a pulsed VCSEL for both devices, while for detector either a free-running SiPM [19] or a time-gated SPAD are mounted, at a relative distance of 30 or 5 mm, respectively.
The source is based on a VCSEL die (VCSEL Part number 680C, Vixar Inc, USA) with a dedicated pulsing electronic circuit developed at Politecnico di Milano. The pulsed VCSEL provides 250 ps full-width at half maximum pulses at 690 nm, with a repetition rate of tens of MHz and up to 900 µW average power.

For PROBE2 we adopted a fast-gated SPAD developed at Politecnico di Milano [20], together with the required wide-band front-end circuitry for detector bias, signal read-out and avalanche quenching. The SPAD is fabricated with a high voltage CMOS process. It features an active area diameter of 200 µm and a photon detection efficiency of 16% at 690 nm and still higher than 5% up to 800 nm. Moreover, it can be gated ON and OFF in less than 200 ps. The PROBE2 size is 20 mm x 25 mm with a source detector distance of 5 mm. It still requires some external electronics (see Figure 1a) that are: (i) a fast pulse generator (220 ps, 7 V pulses, 40 MHz repetition rate) operating the VCSEL in gain-switching mode and providing the global synchronization signal; (ii) a second pulse generator (6 ns, 6.5 V) to gate the SPAD; (iii) a passive delay timer to shift in time the gate window; (iv) a comparator to pick up the SPAD avalanche signal [21] and restore the detector operating condition after each photon detection; (v) a TCSPC board [22] (SPC-130, Becker & Hickl GmbH, Germany).

For PROBE1 we implemented a SiPM (C30742-11) by Excelitas Technologies driven by a dedicated electronics so as to obtain optimal time performances. It features a large active area of 1 mm² with a detection efficiency >10% at 690 nm. The experimental setup is a modified ver-

![Diagram](image-url)

Fig. 1. (a) Scheme of the experimental setup employed to operate the TD PROBE2 based on small (5 mm) source-detector separation. (b) Scheme of TD system (using PROBE1) for large (30 mm) source-detector separations.
sion of the previous one (see Figure 1b): (i) the probe is split in two parts for large separations (e.g. 30 mm) between laser and detector; (ii) the delayer and the second fast pulse generator are removed; (iii) the comparator is replaced by a signal amplifier (VT120, Ortec, USA) due to the lower amplitude of the SiPM avalanche signal with respect to a SPAD; (iv) the reset signal is no longer needed since each triggered microcell of the SiPM is passively recharged back to its quiescent level via an integrated resistor.

2.3. Phantom Test

A water solution of Intralipid and India ink with \( \mu_a = 0.1 \text{ cm}^{-1} \) and \( \mu'_s = 10 \text{ cm}^{-1} \) was employed as bulk medium [23] in which a black plastic (PVC) cylinder (5 mm diameter, 5 mm height) worked as absorption perturbation. It was demonstrated that a totally absorbing object of the proper volume can mimic the effect of any absorption perturbation of a given value and volume within a fairly large range of applicability [24]. The probe, opportunistically waterproofed, was placed in contact with the free surface of the liquid phantom, with the PVC cylinder moved at different depths (from 2.5 mm down to 40 mm). For the fast-gated SPAD (PROBE2) the detector was gated-ON with a 5 ns window, starting few ns (3.5 ns for PROBE2 and 4.8 ns for the 100x power case using PROBE2) after the peak of the Instrument Response Function (IRF). For PROBE1 the IRF was taken by directly facing source and detector with a thin Teflon layer and some gelatin filters in-between in order to both provide an isotropic illumination of the SiPM and to properly decrease the optical power to reach the single-photon regime. Vice versa, using PROBE 2 source and detector cannot be faced being hosted into the same printed circuit board at a relative distance of just 5 mm, hence the IRF was acquired by placing the board onto a thin Teflon layer so as to provide a weakly-diffused optical pulse reflection to the SPAD, while gelatin filters were used to cover the VCSEL to properly adjust the optical power. All measurements were taken at 690 nm, apart from the 100X power case using PROBE2, where the source was a 80 mW, 745 nm laser (the phantom recipe was modified accordingly in order to maintain the same optical properties), with an acquisition time of 1 s (20 repetitions to compute standard deviation).

For data analysis, the background noise was computed at the last delay on each time-resolved curve and then subtracted, thus removing the average value of both dark counts and memory effects [25]. The relative contrast with respect to the unperturbed situation (inhomogeneity at 50 mm of depth) was calculated as \( C = (N_0 - N)/N_0 \), where \( N_0 \) are the total counts within a given temporal window during the unperturbed situation, while \( N \) are the total counts in the perturbed case. A temporal window of 500 ps was used to compare with results reported for a state-of-the-art TD system [29]. For each system, the time-window for contrast calculation was placed at the time position (delay) which maximizes the contrast for the deepest regions of the phantom in order to investigate the system penetration capability.

2.4. In-vivo Test

A motor task experiment was performed to explore the use of the proposed approach for in vivo applications. PROBE2 was positioned over the sensorimotor area of the left hemisphere, centred on C3 according to the 10/20 International System. The correct position was previously checked with a frameless Neuronavigation system (Softaxic, EMS, Bologna, Italy). PROBE2 was fixed to the scalp by Velcro stripes and medical grade adhesive tape. The subject (male, 45 y, right handed) was comfortably sitting in a dimmed room. The motor task consisted in right hand finger tapping at 2 Hz rate. Ten consecutive trials were executed by alternating 20 s baseline (hand still), 20 s task (finger tapping) and 20 s recovery (hand still). The total duration of the protocol was 10 minutes. The protocol on volunteers was approved by the “Ethical committee for clinical investigations of Hospital Clinic de Barcelona”. The acquisition time
Fig. 2. (a) Simulations of the contrast $C$ produced by a 1 cm$^3$ cubical perturbation set at increasing depths $z$ in a homogeneous medium ($\mu_a=0.1$ cm$^{-1}$, $\mu'_s=10$ cm$^{-1}$) and related to an ideal TD (blue continuous lines) or CW (red dashed lines) system. The detectability region ($C>1\%$ and $\sigma<1\%$) is represented in white. TD data are obtained at $\rho=0$ cm at increasing delays from 1 to 9 ns. CW data are obtained at $\rho=1,2,3,4,5$ cm. (b, c) Reconstruction results of the tomography performed using CW (b) and TD (c) simulated data. The represented surfaces correspond to the half of the maximum of the absorption change $\Delta \mu_a$ obtained in all the volume. The simulated perturbation is also depicted in red as reference.

was 1 s and the photon count rate was set initially at about 1.35 Mcounts/s. Data analysis was performed with a Matlab (The Mathworks Inc. Natick, US) script to calculate the time course of the contrast at 690 nm, calculated as for the phantom measurements in each trial. The unperturbed situation is obtained by averaging the photon counts in the baseline period, while the perturbed intensity is calculated at each time instant. Average with standard deviation over the 10 trials of the experiment are then calculated to enhance the signal-to-noise ratio (folding average).

3. Results

To understand the ultimate limits of TD approaches, here we study the problem – problematic of many practical cases, such as the assessment of tumour masses or brain functional activations – of the detection of a localised absorption inhomogeneity set at a depth $z$ within a semi-infinite diffusive medium. We consider an optimal system under realistic conditions, where ideal sources and detectors are densely packed on a 1 cm$^2$ optode, with 100% fill factor, unitary quantum efficiency, perfect gating capability, and a power density of 4 mW/mm$^2$. As a measure of sensitivity, we simulate the relative contrast $C$ produced by a 100% absorption increase in a 1 cm$^3$ cube embedded in a homogeneous turbid medium (Fig 2a). We compare ideal CW and TD systems for an acquisition time of 1 s, assuming a single source-detector pair, set above the inclusion, using – for CW – different $\rho$ values, and – for TD – $\rho=0$, different temporal delays $t$, and a temporal window $\Delta t=1$ ns. Assuming as criteria for detectability $C>1\%$ and a Poisson noise $\sigma<1\%$ (white region in Fig 2a), TD extends well beyond CW reaching a maximum depth $z_{\text{max}}=6$ cm for $t=9.6$ ns, with at least 10,000 photons ($\sigma<1\%$). In this extreme case, the TD contrast is $10^4$ higher than the CW contrast, but also for $z=2$ cm – typical of brain cortex imaging in adults – the TD contrast is 8-fold higher than the CW contrast at $\rho=5$ cm.
The ideal TD scheme is optimal also for 3D tomography in reflectance geometry (Fig 2b-c). To emphasize the depth sensitivity of the TD approach, we assumed an absorption perturbation at a depth $z = 40$ mm. The perturbation can be easily reconstructed with TD (Fig 2b), whereas the ideal CW case (Fig 2c) fails in retrieving the true object (red cube in the figure). The reason for this opposite behavior is clearly seen in Fig 2a. Whereas TD provides a contrast $C > 10\%$, conversely CW yields $C < 0.03\%$ (well below the boundaries of the figure), thus cannot retrieve the object out of the $0.1\%$ added Gaussian noise.

A dense, distributed source-detector arrangement, approaching the ideal case, has already been constructed for the CW approach, demonstrating exciting results for instance on the study of brain functional connectivity [26,27]. For the TD case, to exploit the full potential expressed in Fig 2, three conditions are needed, that are: a dense coverage of pulsed sources (condition 1) and detectors (condition 2) onto the probe so as to inject and collect enough photons, and the gating capability (condition 3) for extracting late photons even at null-$\rho$.

Here we demonstrate the basic building block – a TD source-detector pair directly integrated onto the probe – that sets the basis of next generation TD systems. For the source, we provide the first demonstration of a pulsed VCSEL [28] in TD diffuse optics, as the basic element suitable for dense pulsed light injection (condition 1). For detection, we exploit both a free-running SiPM [19], and a time-gated CMOS-SPAD. The former attains already a large $1 \text{ mm}^2$ active area, and is compact, low cost, and densely packable (condition 2) in matrices of square dies, being developed on large scales for positron emission tomography (PET) systems. The latter demonstrates the gating capability (condition 3) on the basic constituent – a 0.2 mm diameter SPAD in standard CMOS technology [20] – of the SiPM, which is composed of hundreds of small area SPADs with high fill factor. The source-detector pair is integrated onto a small probe to be placed in contact with the medium. Since the SiPM (PROBE1) is not yet fast-gated, we were forced to implement a standard large source-detector distance $\rho = 3$ cm. Conversely, for the fast-gated SPAD (PROBE2) we could use a short $\rho = 0.5$ cm. All the additional electronics – power supply, VCSEL and SPAD gating, amplifiers, TCSPC board – are still external.

To validate the devices – following the NEUROPT Protocol for performance assessment of diffuse optical imagers [14] – we measured the relative contrast $C$ produced by a 100 mm$^3$ black cylinder (equivalent to a 160% absorption increase over a 1 cm$^3$ sphere [24]) set at a depth $z$ within a tissue-like homogeneous phantom (Fig. 3a). PROBE1 can detect the object ($C > 1\%$) down to $z_{\text{max}} = 3.0$ cm, with performances already comparable with a bulky, and costly state-of-the-art TD system [29]. PROBE2 demonstrates the effective rejection of early photons by the gating mechanism permitting to reach $z_{\text{max}} = 2.5$ cm with $\rho = 0.5$ cm, thus already at the level of adult brain cortex. Yet, its performance is worse as compared to PROBE1 because of the tiny area – 0.03 mm$^2$ – of the single SPAD element, which limit the maximum reachable time window. To prefigure what could be the gain in performances applying the gating to each SPAD element of a SiPM, we increased the power delivered in PROBE2 by a factor of 100 using an external laser source, reaching $z_{\text{max}} = 3.5$ cm (PROBE2, 100x power). Plotting the contrast-to-noise ratio $\text{CNR} = (N_0 - N)/\sigma(N_0)$, both PROBE1 and PROBE2 show striking better performances as compared to the state-of-the-art system (Fig. 3b).

As a first exploratory use of the proposed approach for in vivo measurements, we performed a standard brain activation exercise – finger tapping – positioning PROBE2 directly on the head over the motor cortex area (Fig. 4). The decrease of $C$ at 690 nm followed by a slow recovery is the standard signature of deoxy-haemoglobin decrease in the activated area due to neurovascular coupling [30]. Thus, the compact pixel can already detect standard motor-cortex activations even in this first single-small area SPAD and at short $\rho$. A complete in-vivo validation of the device will require a more mature setup with full imaging capabilities resulting
from the realization of multiple optodes.

4. Discussion

The first demonstration of this technology to TD diffuse optics will be disruptive in two ways. First, it will permit construction of small, wearable, low cost TD devices, of great interest in personalized health monitoring, or even deployment of smartphone consumer applications. Although the present paper is focused on the problem of imaging through inhomogeneous media, nonetheless the proposed probe could be considered also for spectroscopy of homogeneous media, in applications like breast cancer risk assessment, on-line quality assessment of fruit or pharmaceuticals [31,32], monitoring of porous materials [33,34]. In terms of performances, even its first implementation using off-the-shelf $100$ VCSELs and $100$ SiPMs is already comparable to state-of-the-art bulky TD systems (Fig. 3), with large room for improvements. For what concerns the additional electronics, the most crucial constituent is the TCSPC board, which can be replaced by CMOS time-to-digital converters (TDC) already demonstrated in SiPMs or SPADs technology [35], while the other blocks are already suitable for smooth integration in a CMOS chip.

On a parallel path, combining a large number of small (e.g. 1 cm$^2$) single optode blocks into a flexible Pad will push TD Diffuse Optics towards a new general purpose imaging technique with a depth of view of 4-6 cm, comparable to other modalities such as ultrasonography, with worse spatial resolution but with chemical and functional specificity. In addition to the enormous benefit for the imaging of brain functions and pathologies, new organs will be addressable such as the lungs or heart (e.g. with the chance to detect the early onset of myocardial infarction). To reach this goal, we can envision to implement the gating mechanism at the level of the single SPAD component to the whole SiPM matrix, following a similar approach that led to the creation of digital SiPM [36], with $>50\%$ fill factor.

Without pushing the technology to the ultimate TD limits, which could be a goal for the next decades, we can already foresee in next few years the integration on 1 cm$^2$ optode of
Fig. 4. (a) Time course of the contrast $C$ at 690 nm during the motor task experiment. The shaded vertical area marks the task period. (b) Scheme of the experimental set-up. PROBE2 (shown in photo) was positioned over the sensorimotor area of the left hemisphere centered on C3 according to the 10/20 International System. Source detector distance is 5 mm. (c) Raw photon counts acquired within a time-gate during the right hand finger tapping motor task experiment. The shaded vertical areas mark the task period.

$\approx 25$ gated SiPMs with 1 mm$^2$ area, 3 VCSELs firing 1 mW each, 3 single chips for VCSEL pulsing, SiPM gating, and TDC. Such device would provide just a fraction $f=10^{-4}$ of photons calculated for the ideal TD case in Fig.2, still it could reach $z_{\text{max}} \approx 5$ cm. For comparison, PROBE2 stands at $f=10^{-8}$, $z_{\text{max}} \approx 2.5$ cm, while the boosted SPAD at $f=10^{-6}$, $z_{\text{max}} \approx 3.7$ cm.

5. Conclusions

In conclusion, we have shown on simulations that TD approaches could greatly enhance depth penetration and sensitivity as compared to CW systems given that a dense distribution of sources and time-gated detectors is provided. Towards this aim, we demonstrate 1) pulsed VCSEL as a possible component for dense integration of sources on the probe; 2) use of SIPM towards a large area dense coverage of detection; 3) fast gating on the basic cell of a SIPM using a CMOS SPAD. Phantom measurements applying the NEUROPT Protocol [14] showed that the VCSEL-SiPM probe yields already comparable contrast and better contrast-to-noise ratio as compared to a bulky state-of-the-art TD instrument. Conversely, the VCSEL-SPAD probe demonstrated operation of the basic SiPM cell in fast gated mode for null source-detector dis-
tance measurements. At present it is limited by the small active area, but we showed that with a 100x increase in power – simulating the gating of the whole SiPM matrix – striking better performances as compared to existing systems are obtained. Further, even in its present form, the VCSEL-SPAD probe was capable to monitor in vivo a standard brain motor cortex activation.

Such technology could lead on one hand to smart, low cost, compact devices for widespread use in personalized health-care and non-destructive testing of diffusive media, and on the other hand to a dramatic increase in contrast and depth sensitivity, yielding a more powerful non-invasive diagnostic modality. The path is opened, with rapid progress feasible exploiting achievements in parallel fields such as PET systems. We are just at the dawn of a new bright era for Diffuse Optics.

Acknowledgments

Funding for this work was provided by the European Community’s Seventh Framework Programme under grant agreements NEUOPT n°201076 and LASERLAB EUROPE n°284464.