Silicon photomultipliers for improved detection of low light levels in miniature near-infrared spectroscopy instruments

R. Zimmermann,1,2,* F. Braun,3 T. Acht nich,3 O. Lambercy,2 R. Gassert,2 and M. Wolf1

1Biomedical Optics Research Laboratory, University Hospital Zurich, Switzerland
2Rehabilitation Engineering Laboratory, ETH Zurich, Switzerland
3Institute for Biomedical Engineering, ETH Zurich, Switzerland

*raphaelz@ethz.ch

Abstract: Silicon photomultipliers are novel solid state photodetectors that recently became commercially available. The goal of this paper was to investigate their suitability for low light level detection in miniaturized functional near-infrared spectroscopy instruments. Two measurement modules with a footprint of 26 × 26 mm² were built, and the signal-to-noise ratio was assessed for variable source-detector separations between 25 and 65 mm on phantoms with similar optical properties to those of a human head. These measurements revealed that the signal-to-noise ratio of the raw signal was superior to an empirically derived design requirement for source-detector separations up to 50 mm. An arterial arm occlusion was also performed on one of the authors in vivo, to induce reproducible hemodynamic changes which confirmed the validity of the measured signals. The proposed use of silicon photomultipliers in functional near-infrared spectroscopy bears large potential for future development of precise, yet compact and modular instruments, and affords improvements of the source-detector separation by 67% compared to the commonly used 30 mm.

© 2013 Optical Society of America

OCIS codes: (120.6200) Spectrometers and spectroscopic instrumentation; (040.3780) Low light level; (040.5250) Photomultipliers; (170.2655) Functional monitoring and imaging.

References and links

1. M. Ferrari, L. Mottola, and V. Quaresima, “Principles, techniques, and limitations of near infrared spectroscopy,” Can. J. Appl. Physiol. 29, 463–487 (2004).
2. M. Wolf, M. Ferrari, and V. Quaresima, “Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications,” J. Biomed. Opt. 12, 062104 (2007).
3. M. Ferrari and V. Quaresima, “A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application,” Neuroimage 63, 921–935 (2012).
4. B. N. Pasley and R. D. Freeman, “Neurovascular coupling,” Scholarpedia 3, 5340 (2008).
5. A. Villringer and B. Chance, “Non-invasive optical spectroscopy and imaging of human brain function,” Trends Neurosci. 20, 435–442 (1997).
6. G. Strangman, D. A. Boas, and J. P. Sutton, “Non-invasive neuroimaging using near-infrared light,” Biol. Psychiatry 52, 679–693 (2002).
7. T. J. Germon, P. D. Evans, N. J. Barnett, P. Wall, A. R. Manara, and R. J. Nelson, “Cerebral near infrared spectroscopy: emitter-detector separation must be increased,” Br. J. Anaesth. 82, 831–837 (1999).
8. A. Kienle and M. S. Patterson, “Improved solutions of the steady-state and the time-resolved diffusion equations for reflectance from a semi-infinite turbid medium,” J. Opt. Soc. Am. A 14, 246–254 (1997).
1. Introduction

Near-infrared spectroscopy (NIRS) is an optical and non-invasive method that can be used to determine the concentration of oxy- and deoxyhemoglobin (O₂Hb and HHb, respectively) in human tissue in vivo [1, 2]. In functional NIRS (fNIRS), this principle is used to assess the cortical activation of brain tissue [1, 3]: increased cortical activity alters the oxygenation state of brain tissue through neurovascular coupling [3–5], which causes a focal increase (decrease) of O₂Hb (HHb) [1, 3, 5].

In continuous-wave fNIRS, steady-state NIR light of at least two discrete wavelengths is guided to the scalp, and its diffuse reflectance is measured at some centimeters from the source site [1–3]. The modified Beer-Lambert law (MBLL) is used to relate concentration changes of O₂Hb and HHb to changes in the detected light intensity [1, 5, 6]. A fraction of the detected photons follows a path that passes through cortical tissue, thereby carrying information about the brain. This fraction depends on the source-detector separation (SDS): the larger the SDS, the deeper the detected photons have penetrated the tissue and hence the larger the probed cortical volume, resulting in improved measurement signals [1–3, 7]. However, increasing the SDS also dramatically increases the light attenuation (approx. 20 dB per cm SDS [8]).

One of the major advantages of fNIRS over other neuroimaging modalities is the fact that the instruments can be miniaturized and made wearable, easing the restrictions to head movement and offering increased comfort to the subjects [3]. Such fNIRS devices have to be lightweight and small, which narrows the selection of their optical and electrical components. In particular, special attention is due to the choice of the detector, since it directly influences the measured signal quality and determines the readout- and supply circuitry. Because of the high attenuation of the light at longer SDS, it is beneficial to select very sensitive photodetectors with internal signal amplification [6].

The detection of such low light levels can be done by photomultiplier tubes (PMTs) since they feature high internal gains, but PMTs are bulky, heavy and require high operating voltages,
and are hence not suitable for miniaturized and/or wearable fNIRS instruments. Furthermore, PMTs are sensitive to magnetic fields (i.e., incompatible with magnetic resonance imaging) and vulnerable to overexposure at ambient light [9]. Avalanche photodiodes (APDs), in contrast, are solid state devices that are small in size, insensitive to overexposure and not affected by magnetic fields. But their internal gain is typically lower, and their operating voltage is also rather high [9].

Recently, silicon photomultipliers (SiPMs) were recognized as a solid state alternative to PMTs [9–11]. SiPMs consist of a dense array of parallelly connected tiny APDs, called microcells, which operate independently in Geiger mode [9, 10]. When a photon hits the active area of a microcell, an ionization cascade (“avalanche”) can be triggered that results in a largely amplified current peak. The sum of the currents from all the microcells in the array forms a quasi-analog output signal proportional to the incident light intensity [9–11].

Since SiPMs share the benefits of APDs as they are small, lightweight, robust and relatively low-priced, but operate at lower voltages and exhibit an internal gain that is comparable to PMTs [9–11], they undoubtedly bear advantages that have yet to be exploited for fNIRS instrumentation. The goal of this work hence was to demonstrate the suitability of SiPMs as photodetectors in low level light detection in miniaturized fNIRS instrumentation, and is partially based on a recently completed semester thesis [12], conducted by two of the authors (F.B. and T.A.). To the best of our knowledge, we are the first to report on the design and evaluation of an fNIRS measurement module that features SiPMs as detector units.

2. Methods

2.1. fNIRS sensor module design and implementation

Two fNIRS modules were built, each consisting of two separate printed circuit boards, stacked on top of each other (Fig. 1(A)), and designed to be used as both, source and detector module. A brief overview of the module design is given here, while details can be found in [12].

The used SiPM (PM3360, KETEK GmbH, Munich, Germany) had an active area of $3 \times 3$ mm$^2$ with 2500 microcells. The SiPM current was converted to a voltage signal by a transimpedance amplifier (TIA) as illustrated in Fig. 1(B). The TIA stage had a low-pass characteristic (1st-order, cutoff-frequency: 1 kHz) and a signal amplification of 510 $\Omega$. The resulting voltage was subsequently filtered on-board using a 2nd-order low-pass filter with a cutoff-frequency of 1 kHz and an additional signal amplification stage with a gain of 2. This second amplifier stage allowed to apply an offset voltage in order to keep the range of the preamplified voltage level within the input swing of the subsequent ADC. In order to control the detector gain, the bias voltage was adjustable between $-32$ and $-24$ V. Each module was equipped with two LEDs (Osa Opto Light GmbH, Berlin, Germany), emitting at 680 $\pm$ 8 and 850 $\pm$ 8 nm, respectively. The LED package contained a small lens that allowed for a half viewing angle

![Fig. 1. A: the proposed fNIRS module besides a 2-E coin. B: SiPM readout circuitry (TIA: transimpedance amplifier, LPF: low-pass filter). C: block diagram of the module (GUI: graphical user interface, $\mu$C: microcontroller, ADC: analog-digital converter, DAC: digital-analog converter).](image-url)
of 20 degrees. The emitted light intensities were adjustable by controlling the current through the LEDs, which was fixed at 30 mA for all measurements presented here. The power density impinging on the sample was estimated based on the radiation characteristics and values for the radiant intensity from the datasheet of the selected LEDs. Under the assumption that i) the radiation peaks along the optical axis with a maximum value of $3.5 \,(7.0) \, mW/\text{sr}$ for the LED emitting at 680 (850) nm, respectively; and that ii) in the time multiplexed scheme each LED has a duty cycle of one third (see below), the time-averaged power densities were estimated to be $28 \,(55) \, mW/cm^2$ at 680 (850) nm. This is in the range of the power densities generally used in fNIRS, and only slightly higher than the irradiation from sunlight at these wavelengths [13]. The peak total optical power was calculated to be approximately $3.2 \,(6.3) \, mW$. This is comparable to the values previously reported in the literature for wearable fNIRS instruments [14, 15]. Under normal operating conditions and based on assumptions adapted from [16], the LEDs used in this work can be considered safe for both the eye (including the retina, cornea and the lens) as well as the skin under the guidelines given by the ICNIRP [17].

A cover made of black silicone and held in place by black rubber tape was used to shield the detector from stray light and to minimize lateral radiation from the LEDs. The modules were powered by a lab power supply (HM7042-5, HAMEG Instruments GmbH, Mainhausen, Germany) providing three voltage levels ($−32.0, −3.0$ and $3.5$ V) in combination with on-board low-dropout regulators (Texas Instruments Inc., Dallas, TX, USA). The silicon cover served as a provisional isolator between the module and the subject. Furthermore, the power supply was operated in a current-limited mode in order to eliminate the risk of severe electrical hazards. The limits of the currents were $10$ mA for the $−32.0$ V output, $110$ mA for $3.5$ V and $40$ mA for $−3.0$ V.

Measurements were performed in a time-multiplexed scheme, by using one of the modules as source device, and the other as detector, allowing for easy variation of the SDS. One sample consisted of an intensity measurement for both wavelengths and a background light measurement. To this end, each LED was switched on separately for $3.3$ ms. After waiting $1.8$ ms in order to let the low-pass filtered output voltage reach a steady state, a 16-bit ADC (ADS8330, Texas Instruments Inc., Dallas, TX, USA) was used to sample both the TIA-output and $V_{out}$ (Fig. 1(B)) 50 times at a rate of 50 kHz. These ADC samples were finally averaged, and the procedure was repeated for the other LED and for the background light condition, in which both LEDs were off. In this configuration, the overall sampling rate was 100 Hz. An on-board microcontroller (ATxmega128A4U, Atmel Corp., San Jose, CA, USA) was used to read the ADC, execute the averaging operation and communicate via USB with a PC running a graphical user interface (GUI), which was developed in MATLAB® (The MathWorks Inc., Natick, MA, USA) and used to set the SiPM bias voltage, the LED current and apply an offset correction prior to the A/D conversion as illustrated in Figs. 1(B) and 1(C).

2.2. Signal-to-noise measurements

The signal-to-noise ratio (SNR) was measured using optical phantoms (ISS Inc., Champaign, IL, USA) and was defined as $20 \log_{10} (\bar{s}/s_{\sigma})$, where $\bar{s}$ and $s_{\sigma}$ were the signal mean and the standard deviation across 300 data points, respectively. The modules were placed on the phantom at SDSs ranging from 25 to 65 mm and were firmly pressed to the surface with adhesive tape. To eliminate stray-light effects or light-piping from the LEDs directly to the SiPM, black rubber tape was used to shield the modules, and the measurements were performed in a light-tight box. Adjustments of the SiPM gain (by tuning the bias voltage) and the offset voltage in the second amplifier stage were performed prior to the measurements, to prevent ADC saturation while allowing for maximum signal amplitude. The placement of the modules was carried out once per measured SDS. For each placement (i.e. SDS), the SNR was determined ten times.
consecutively. For each phantom and SDS, these ten SNR values were then used to calculate the mean SNR and the corresponding standard deviation (SD).

In order to relate phantom data to in vivo measurements, the optical properties of the phantom should match those found in humans. Besides the SDS, the measured light intensity – and hence the SNR – essentially depends on the effective transport coefficient \( \mu_t = \sqrt{3} \mu_a (\mu_a + \mu'_s) \), which is a function of the absorption coefficient \( \mu_a \) and the reduced scattering coefficient \( \mu'_s \) [8]. Own frequency-domain measurement data of the optical coefficients of the forehead of 5 healthy subjects (details not reported) are given in Table 1. None of the available phantoms exhibited the exact optical coefficients, hence the SNR measurements were conducted on two different phantoms (A and B, Table 1), which were selected such that the \( \mu_t \) from the human forehead is between that of the two phantoms.

A lower bound for the SNR was defined as follows: (conservatively) assuming a functionally induced change in \( \text{O}_2\text{Hb} \) and \( \text{HHb} \) of 1 and \(-0.25 \, \mu\text{M}\), respectively [3, 6], inversion of the MBLL yielded the corresponding intensity changes \( \Delta I \) with respect to the measured baseline intensity \( I_0 \). In order to obtain adequate SNR requirements for functional measurements, the inverted MBLL was parameterized with values corresponding to the head, i.e. with extinction coefficients and differential pathlength factors for the forehead. The extinction coefficients were taken from [18], and the differential pathlength factors for 680 and 850 nm were linearly extrapolated from the data published in [19], since the published data did not match the wavelengths used in this work. The determination of a minimum SNR, based on the inversion of the MBLL, allows inferring the required SNR that would be needed to detect changes in \( \text{O}_2\text{Hb} \) and \( \text{HHb} \) – due to changes in brain activity – by analyzing the SNR of the detected optical power on a static optical phantom. To adequately resolve these changes, the maximally tolerable noise level was chosen to be one third of \( \Delta I \), yielding a minimally required SNR

\[
\text{SNR}_{min} = 20 \log_{10} \left( \frac{3I_0}{\Delta I} \right).
\]

2.3. In vivo experiments

Easily reproducible hemodynamic changes were induced in one of the authors by occluding the blood flow to and from the forearm with a pneumatic cuff by applying a pressure of more than 280 mmHg during 120 s, followed by 180 s of recovery [12]. The modules measured muscular tissue during three occlusion periods at an SDS of 50 mm. In addition to shielding with black rubber tape, the modules were covered with black cloth to limit the background light. Changes in \( \text{O}_2\text{Hb} \) and \( \text{HHb} \) were calculated according to the MBLL, with differential pathlength factors for the forearm linearly extrapolated from [19] and the extinction coefficients from [18]. The data were low-pass filtered with a digital 3rd-order Butterworth filter at a cutoff-frequency of 1 Hz.

3. Results

Figure 2 (top) shows the obtained SNR values (mean \pm SD across the ten repetitions) for the different SDSs and the two phantoms, together with the estimated minimally required
SNR for both wavelengths (gray lines). Due to saturation of $V_{\text{out}}$, the data of phantom A for $\text{SDS} \leq 30 \text{ mm}$ and the data of phantom B for $\text{SDS} \leq 40 \text{ mm}$ were obtained directly from the TIA output (Fig. 1(B)). For phantom A, the critical $\text{SDS}$, up to which the design criteria are fulfilled for both wavelengths, was roughly 50 mm, while for phantom B, it was 55 mm.

For the \textit{in vivo} experiment, the traces of $O_2\text{Hb}$, $HHb$ and the change in total hemoglobin ($t\text{Hb} = O_2\text{Hb} + HHb$) during the occlusion periods are shown in Fig. 2 (bottom). With the onset of the occlusions, $O_2\text{Hb}$ and $t\text{Hb}$ increased instantly. During occlusion, $HHb$ constantly increased, and $O_2\text{Hb}$ decreased at a similar rate, while changes in $t\text{Hb}$ were much smaller. Releasing the occlusion induced a “rebound”-reaction after which the hemodynamics equalized.

4. Discussion & outlook

This article proposes SiPMs as detectors for fNIRS instrumentation, and introduces a modular prototype measurement system. SNR measurement data obtained from optical phantoms showed that SiPMs enable measurements at large $\text{SDS}$s, while affording small, lightweight and modular probes. The proposed lower bound for the SNR enabled a requirement-driven assessment of the range of $\text{SDS}$s which are expected to yield a sufficiently high resolution that allows to detect changes in brain activity on a single-trial basis.

The $\text{SDS}$ up to which the requirements were met was 50 mm. This is twice as large as the $\text{SDS}$ used in [20] and 67% more than the commonly used 30 mm $\text{SDS}$ [1, 3], and is in line with the work in [7], where a minimum $\text{SDS}$ of 48 mm has been suggested for optimal fNIRS measurements. The backlight-corrected current measured at the SiPM output for the critical $\text{SDS}$ of 50 mm was 1.12 (1.47) mA for phantom A and 2.49 (2.61) mA for phantom B, at the wavelengths 680 (850) nm, respectively. These comparably high currents are indicative for the large internal amplification of SiPMs, and allow for simple and miniaturizable data readout without the need for elaborate and complex ultra-low noise circuitry even at longer $\text{SDS}$s.

Digital low-pass filtering of the data is expected to increase the SNR performance further.
SDSs beyond 50 mm are thus possible. The sampling rate of 100 Hz is – compared to the relatively slow hemodynamics – rather fast, and the fNIRS data are usually low-pass filtered during preprocessing. Assuming white Gaussian noise and a simple 10 point moving average filter, a 10 dB increase in the SNR is expected.

The in vivo measurement data demonstrated the validity of the obtained results at an SDS of 50 mm. The results showed the expected increase in HHb and accompanying decrease in O$_2$Hb during the occlusions, which result from the closed system dynamics due to the blocking of blood in- and out-flow. The steep increase in O$_2$Hb and tHb at the occlusion onset can be explained by the non-zero time that is required to apply the pressure required for arterial occlusion. The signal traces, as well as the magnitudes of the induced hemodynamic changes, were comparable to those reported in [1, 20].

4.1. Limitations

The phantoms used for this study did not perfectly match the optical properties of the human head. As the effective transport coefficient of phantom A was larger than that measured on the human forehead, it yields a too low estimate of the SNR. Conversely, $\mu_t$ was too small in phantom B, leading to an optimistic SNR measurement. We expect the SNR that corresponds to the human head to lie somewhere in between the results from phantom A and B.

The developed modules suffered from a certain drift in the intensity signal, which was larger for short SDSs and vanished for longer SDSs. The origin is assumed to be internal heating of the SiPM [12]. Thermally stabilized SiPM modules are commercially available, and might be alternatively considered in a future redesign. However, the background light measurement in the time multiplexed scheme allowed for a partial correction of the drift, which becomes evident in the occlusion data, where no apparent signal drift is visible.

4.2. Future work

The modules are first prototypes developed to test the potential of using SiPMs in fNIRS applications. The electrical isolation of these prototypes was thus provisional and protective means were implemented directly at the level of the power supply. A redesign including safety circuitry with current limiters and a proper casing will be required prior to extensive testing of the device in humans, especially when fully wireless and battery powered operation with medical certification is pursued.

No substantial difference in the SNR was observed when comparing the data obtained from the TIA output and $V_{out}$ (Fig. 1(B)). The second stage of the amplifier served two purposes: it featured a 2$^{nd}$-order low-pass filter as well as the possibility to shift the amplified voltage level into the input range of the ADC. Since the TIA already had a 1$^{st}$-order low-pass characteristic, the subsequent 2$^{nd}$-order filter was most effective in the frequency band close to 1 kHz (where the noise rejection from the 1$^{st}$-order filter was not yet very potent). As inferred from the marginal improvements when comparing with the TIA output data, however, this noise rejection was already achieved through the onboard digital filtering (i.e. a data point was based on the average of 50 ADC samples which were obtained at 50 kHz). Hence, in a future redesign of the modules, the analog filtering circuitry could be simplified accordingly, allowing for reduced system complexity and size.

Each module communicated via an on-board microcontroller, and hence would be individually addressable in a bus, making parallel operation in a sensor network possible. Such a setup would benefit from a flexible probe placement with multiple SDSs on the one hand, and from a dense grid of light paths for optical topography on the other hand.
5. Conclusion

Due to their excellent SNR performance, small size and weight and low operating voltage, SiPMs offer new opportunities for the design of small, modular, flexible and highly precise fNIRS instruments.

Acknowledgments

This work was supported by the CHIRP1 ETH Research Grant CH1-02 09-3, and the National Center of Competence in Research on Neural Plasticity and Repair of the Swiss National Science Foundation. The authors would like to thank T. Ganka from KETEK GmbH and Prof. Dr. S. Kozerke for technical support, and F. Scholkmann, A. Metz and J. Mata Pavia for useful discussions.