Sensory Particles with Optical Telemetry

Karthik Ganesan,
Thomas A. Flores,
Binh Q. Le,
Dante G. Muratore,
Neal Patel,
Subhasish Mitra,
Boris Murmann,
and Daniel Palanker

Abstract—Current retinal prostheses provide electrical stimulation without feedback from the stimulated neurons. Incorporation of multichannel recording electronics would typically require trans-scleral cables for power supply and data transmission. In this work, we explore a wireless, optoelectronic, miniature, modular, and distributed electro-neural interface for recording, which we call Sensory Particles with Optical Telemetry (SPOT). It can be used in an advanced, bi-directional retinal prosthesis and other sensory applications. Emphasis is placed on the novel telemetry stage. SPOTs are powered by near-infrared light and transmit information by light. As a proof of concept, we designed and built a low-power, small-footprint linear transconductance circuit utilizing chopper stabilization in 130nm CMOS. Our design achieved 57 mS transconductance within 3.5 kHz bandwidth, and a near-infrared (NIR) power density of 0.5 mW/mm², well within the ocular and thermal safety limits. The telemetry circuit consumes 0.015 mm² area, and each SPOT can be powered by a single photovoltaic (PV) supply of area 0.0056 mm². Electrical spikes transmitted by an 850nm LED were detected with 15 dB SNR, at the output of the optical link.

Keywords—Biomedical Electronics, Biomedical Telemetry, Optical Telemetry, Near-Infrared Light, Neural Recording.

I. INTRODUCTION

Development of electrical neural interfaces has provided a multitude of tools to further our understanding and enable control of the central and peripheral nervous systems. Devices with several electrodes have been used for therapeutics and prosthetics across multiple medical disciplines, including cardiology [1], urology [2], otolaryngology [3], and many others, providing effective treatment of chronic disorders. Progress in miniaturization of the recording and stimulating systems has allowed the development of more sophisticated and less invasive electronics, enabling higher spatial and temporal precision of the electro-neural interfaces. Sensory and motor neural systems, for example, depend upon complex, interconnected networks [4], and therefore the associated neural prostheses require sophisticated recording and stimulation capabilities to approximate the neural code.

Retinal prostheses have seen a surge of development in recent years, including demonstrated feasibility of eliciting visual percepts in patients blinded by degenerative retinal disorders, such as Retinitis Pigmentosa and Macular Degeneration [5]. In this disease, the “image capturing” cells in the retina – photoreceptors – degenerate, leading to loss of sight, while the “image processing” retinal neurons are preserved to a large extent [5]. Photoreceptors relay visual information to the graded-response neurons in the inner nuclear layer of the retina, where horizontal cells, a dozen types of bipolar cells, and amacrine cells [6, 7] process the visual signals. Retinal interneurons pass on visual information to about 20 distinct classes of retinal ganglion cells (RGCs) that generate action potentials, propagating to the brain via their axons, which constitute the optic nerve [8]. Overlapping mosaics of the various types of RGCs encode different aspects of the images, such as increase or decrease in light (ON and OFF RGCs), color, differing receptive field sizes, and others.

There are two distinct approaches to electronic restoration of sight. In the subretinal approach, arrays of electrodes located between the inner nuclear layer and the pigment epithelium replace the degenerated photoreceptors and target primarily the surviving bipolar cells [9, 10, 11]. Since these are graded-response neurons, stronger stimuli are encoded with larger amplitude or longer duration pulses. Signals are then passed via the retinal network to RGCs, which convert them into trains of action potentials. Due to utilization of the remaining retinal network, this approach preserves many features of natural vision [10]. However, due to indiscriminate stimulation of different types of bipolar cells, and some rewiring of the retinal network during retinal degeneration, the resulting retinal output does not perfectly replicate the natural retinal code. In the epiretinal approach, prostheses directly target RGCs, using electrodes placed on top of the retina [12, 13]. Since RGCs are spiking neurons, each stimulation pulse can elicit a single spike, which provides an opportunity to properly encode visual information, if each cell type can be identified and selectively stimulated [14]. Stronger responses are encoded using higher frequency or larger number of spikes. In this approach, stimulation of the retinal network should be avoided. Currently, the only approved epiretinal prosthesis, called ARGUS II (Second Sight Inc.), has very few (60) and rather large (200 μm) electrodes spaced about 500 μm apart. Therefore, it provides acuity no better than 20/1260 [12], utilizing indiscriminate stimulation of multiple RGCs.

Classification of the RGCs is typically based on their responses to light [15], which cannot be used in the degenerate retina. However, recent studies demonstrated that RGCs can also be distinguished based on their spontaneous firing rate, and signatures such as velocity of propagation of the action potential extracted from electrical recordings [16].

An advanced retinal implant that would allow not only stimulation, but also recording, with electrode density approaching the natural density of RGCs, might enable identification of the target RGCs for delivery of the visual information in a manner similar to the natural retinal code. Recording of the retinal response to stimulation allows distinguishing between somatic and axonal responses [14, 17], and fine-tuning of stimuli to avoid electrodes that elicit axonal responses or indiscriminate stimulation of multiple cells [18].

Manuscript received Feb. 7, 2020. This work was supported in part by Wu Tsai Neurosciences Institute at Stanford University under Grant BII-04. K. Ganesan and T. A. Flores contributed equally to this work. Asterisk indicates corresponding author: K. Ganesan, B. Q. Le, D. G. Muratore, N. Patel, S. Mitra, and B. Murmann are with Dept. of Electrical Engineering, Stanford University (e-mail: karthik3@stanford.edu). T. A. Flores is currently with Byers Center for Biodesign, Stanford University. B. Q. Le is also with Dept. of Electrical Engineering, San Jose State University. S. Mitra is also with Dept. of Computer Science, Stanford University. D. Palanker is with Dept. of Ophthalmology and Hansen Experimental Physics Laboratory, Stanford University. D. G. Muratore, S. Mitra, B. Murmann, and D. Palanker are also with Wu Tsai Neurosciences Institute at Stanford University.
Recording electrodes in such a bi-directional retinal prosthesis should be able to detect action potentials of 50 μV peak-to-peak amplitude, with a sampling rate of 20 kHz per electrode [19]. The output telemetry must support hundreds of parallel channels for cell-type classification. The total dissipated power is limited to about 5 mW/mm² due to thermal safety considerations [20]. Designing such a low-power system with RF coupling is exceedingly difficult, as significant data compression and multiplexing must take place on-chip to support the appropriate number of channels within the available bandwidth (at most tens of MHz) [19, 20].

We present the design of a distributed electrical neural interface based on micro-modules. Each micro-module is powered photovoltaically and transmits information in the optical domain. This allows simultaneous recording from thousands of such implants, using a fast camera. Photovoltaic stimulation, currently used in subretinal prosthesis [10], can be used in conjunction with this circuit, which will convert the system into a bi-directional prosthesis. We refer to this as SPOTS: Sensory Particles with Optical Telemetry and Stimulation. In this work, we focus on the sensory and telemetry part of this circuit, and hence it is abbreviated to SPOT. Our system has the following features:

1. The approach is scalable, and can be used to record information from a large population of neurons.
2. The micro-modules emit light as a highly-linear function of the differential voltage sensed between two electrodes. Low noise recordings and on-chip offset cancellation benefit the following signal processing steps outside the implant, especially spike detection. Alternatively, the spikes can be detected on the chip by thresholding, and only the spike trains need to be transmitted, which greatly reduces the power requirements for data telemetry.
3. Information transmission via NIR light allows each micro-module’s position to be localized to within 3 μm on tissue.
4. Using 130nm CMOS technology, large batches of micro-modules can be realized at low cost. The systematic methodology in [21] also allows for fully-automated design for neural signals of varying amplitude and/or bandwidth.

II. METHODS

A. Distributed Neural Recording System Overview

The SPOT concept is illustrated in Fig. 1. Multiple SPOTs interface with RGCs for recording [22, 23] over several millimeters-wide visual field (see Fig. 1a). The front-side of each micro-module consists of one photodiode acting as a photovoltaic power supply by an external NIR laser, and a micro-LED for the output telemetry (see Fig. 1b). The back-side contains CMOS recording circuitry and two electrodes (see Fig. 1b). To maximize the differential input for recording, the electrodes are separated by, at least, a cell size (30 μm).

B. Photovoltaic Considerations

Information and power transmission occur in the NIR range (815 nm – 900 nm) [24]. The silicon photovoltaic (PV) cell provides a potential of ~0.5–0.6 V, with a light-to-current conversion efficiency of 0.3 A/W (see [25; Section IV]). The required current for the neural recording circuit (~8 μA) and the LED (~0.25 μA; see Section III for derivation) is therefore generated by a photodiode of 0.0056 mm² in size, illuminated by an NIR light of about 5 mW/mm², which is within the ocular safety limits [10]. The micro-LED can have dimensions 65μm x 65μm [26], for an area of 0.0042 mm². It can be designed to emit light at wavelength ~840 nm [27].

Fig. 1. SPOT concept (best viewed in color). a) Multiple micro-modules distributed over the retinal surface are powered externally by an NIR laser; b) Each micro-module contains a photovoltaic power supply (PD) and micro-LED for information telemetry on its front side. Different wavelengths of light (depicted in different colors) are used for power and for the information output. Electrodes for signal acquisition and CMOS circuitry for amplification and transmission are located on the back side of each micro-module.

C. Neural Transducer Circuit Architecture

The neural signal telemetry circuit consists of three stages with intermediate AC-coupling (see Fig. 2a). The differential neural signal sensed at the inputs undergoes two amplification stages, followed by a transconductance stage that drives current through the LED. The cascaded circuit therefore provides continuous conversion of the extracellular voltage into NIR-light, with high-linearity. The circuit has a bandpass characteristic between 50 Hz and 3.5 kHz, sufficient for recording action potentials [28]. The circuit footprint is 0.015 mm² and it consumes 8.2 μA current from a 1V supply. The power density does not exceed 520 μW/mm², well within heat dissipation limits for chronically-implanted electronics [20].

The first amplification stage consists of a high-gain, low-noise, fully-differential recycling folded cascode (RFC) operational transconductance amplifier (OTA) [29, 30] that utilizes chopper-stabilization for dynamic offset cancellation [31]. The RFC amplifier is designed with a current mirror factor of K=3 for power-efficiency [30]. The amplifier topology is shown in Fig. 2b. All transistors operate in saturation, with the input pairs biased at 2I_{DSS} of around 100 mV [21]. The capacitors C_m and C_f take values of 3.1 pF and 15 fF, respectively, to provide a closed-loop gain of 207. C_m is implemented using a dual-MiM capacitor, and C_f is implemented using four minimum-sized single-MiM capacitors in series. A low-distortion, common-mode feedback circuit [32] is used in the fully-differential amplifier (Fig. 2c).

Importantly, on-chip offset-cancellation is implemented at first stage in order to allow a simple spike detection algorithm to be implemented on-chip. This enables optimal telemetry of all relevant information, by pulsing only a minimal detectable current through the LED, strictly during spiking events.
Chopping is chosen as the offset-cancellation method due to its suitability for micro-power applications and the continuous nature of the input signal [31]. To minimize input-referred residual offset, and this exponentially (see Fig. 4.28). The output-referred residual offset achieved is 9.8 µV, rms (N=994 Monte Carlo runs). The spike detector can thus be implemented using thresholding, or a simple discrete derivative logic circuit [34].

The second amplification stage is implemented as a differential amplifier with PMOS-input and active load (Fig. 2d). The input capacitor and feedback capacitors are sized at 400 fF and 40 fF, respectively, to provide a closed-loop gain of 10. The capacitors are both implemented using thin-oxide transistor gate capacitances, which achieve higher density than could be achieved with the MIM capacitors.

The final stage is a linear transconductance circuit. It contains a differential amplifier with PMOS-input and active load, connected in feedback with a resistor. A 40 kΩ resistor is implemented with 20 segments of 2 kΩ precision polysilicon resistors, to minimize process variation. The stage has a 25 µS transconductance. The output current flows directly through the LED (with typical forward voltage of ~1.45 V [35]).

Four voltage levels are needed for the telemetry circuit (common-mode, CMOS supply (and ground), and LED anode). The four levels can be generated using a single silicon photovoltaic cell, and two stages of voltage doubling. For a typical silicon photovoltaic cell, these four voltages will be related equivalently as 0.5V, 1V, 0V, and 2V respectively. Doubling can be realized with two stages of the current-starved charge pumps [36]. Alternatively, multiple diodes in series could be connected and other voltages can be generated using a bias circuit. This tradeoff should be analyzed in terms of the area cost, fabrication complexity (if series diodes are required), and total current consumption.

To minimize the residual offset, the first-stage amplifier, common-mode feedback circuit, second-stage amplifier, and the transconductance stage are all laid out using common-centroid templates [37]. The pseudoresistors used to set the high-pass pole for each stage are implemented using 3.3 V thick-oxide I/O transistors. All circuits are biased using a single nano-power PTAT current reference [38] (see Fig. 2e, where the start-up circuit is outlined in dashes). The complementary clocks needed for chopping are generated using a nano-power digital oscillator [39] (see Fig. 2f).

III. Results

The recording circuitry was fabricated in a 130nm CMOS process with 8 metal layers. An annotated die-photo of the test-chip is given in Fig. 3. Each test-chip contains three complete recording circuits. Each configuration includes I/O buffer(s) with additional probing pads for measurement of internal voltages. An additional debugging configuration for the first amplifier stage is also included on the test-chip. The recording circuit can have a compact footprint of 0.015 mm².

The performance was measured to verify the gain and bandwidth (Fig. 4a), input-referred noise (Fig. 4b), strength of the received sequence from LED output (Fig. 4c) and power consumption. A discrete LED with peak NIR wavelength of 850nm was used for system testing [35]. To measure light emission from the LED, a photodetector was fixed at 3 cm away from the test-chip (emulating the distance to the back of the eye). The design provided a 57 ms net transconductance, for spikes of amplitude up to 400 µV, with a bandwidth of 50 Hz to 3.5 kHz, while drawing 8.2 µA from a 1V supply. In the band of interest (300 Hz to 3 kHz [40]), the total input-referred noise is 13.05 µV, rms. Signals recovered at the output of the optical link had 15 dB SNR (in the worst case).

The minimum detectable signal from the LED was also measured within a numerical aperture of 0.1, matching the geometry of the human eye. Pulses of 1 µA amplitude with pulse width of at least 0.5 ms could be detected. RGBs emit at most ~200 spikes/sec [41], with each spike lasting ~1 ms. Thus, 1 nC of charge should be stored within each ~4 ms idle period, for sustained current through the LED during each spike. This corresponds to ~0.25 µA drawn continuously from the PV supply – a small fraction of the 8.2 µA required for signal amplification (see Section II.B).

IV. Conclusions and Future Directions

In this paper, we describe an optoelectronic, miniature, modular, and distributed neural interface for recording and information transmission. This system enables a wireless and high-throughput design, that is applicable to advanced retinal prosthesis. The system may also be used for other applications, such as monitoring of physiological parameters using transdermal communication with a smart watch. Several opportunities for the future work remain:

1. Further optimization of the recording circuitry for lower power and smaller area, by combining amplifier topologies that can achieve extremely high gain at low current (e.g.,...
inverter-based amplifiers [42], with complex capacitor structures that achieve high area-efficiency [43].

2. Implementation of the spike detection logic [34] on each micro-module, so only the minimal detectable current is pulsed through the LED during spiking events.

3. Multiple recording electrodes can be used in each micro-module, if they can be switched by an optical control signal. At least 4 electrodes can be accommodated within the empty area on lower metal layers, by expanding the input chopper to a larger pass-transistor logic function. This allows ~2400 electrodes to be implanted within a 9 mm² area of the retina.

4. Complete heterogenous integration of the NIR photodiode for powering and micro-LED for data telemetry with the CMOS recording circuitry on a single micro-module [44].

5. Validation of the SPOT concept with electrophysiology experiment in-vitro (e.g., using QPI [45] as a control).

**ACKNOWLEDGMENTS**

We thank the MOSIS Educational Program for IC fabrication, and the ARM University Program for I/O cells used in our test-chip. This work was supported in part by the Wu Tsai Neurosciences Institute at Stanford University under Grant BI1-04. We thank Dr. Danny Bankman, Tony F. Wu, and Dr. Sean Fischer, for helpful discussions.

---

**Fig. 3. SPOT test-chip die photo (best viewed in color). a) Four test setups are outlined. Config 4: contains I/O buffer for probing clock (all pads labelled). Config 3: contains I/O buffers for externally generated clocks. Config 2: contains I/O buffer for probing second amplifier. Debug: RFC amplifier with chopping; b) Details of config 4. Layout snapshots of the active sub-blocks are drawn to-scale and super-imposed. A compact 0.015 mm² layout was obtained by removing the buffer, stacking capacitors CL1 on top of the active circuits (cf., [34; Section II.A]) and re-positioning the oscillator capacitor.**

**Fig. 4. SPOT Measurements (best viewed in color). a) Gain/Bandwidth of amplifiers: Measured (red) and post-layout simulation in typical (black), fast (dashed), slow (dotted) corners. Non-idealities of the chopper (~5 kHz) include harmonics, which are further filtered out in the last stage; b) Input-referred noise PSD: Measured (red) and post-layout sim. in typical (black), fast (dashed), slow (dotted) corners; c) Input: neural spikes from a waveform generator. Output: spikes detected at output of a photodetector.**
REFERENCES

[35] Datasheet “0.8mm Height Flat Top Infrared LED HJR19-21C/L11/TR8,” Everlight Electronics Co., LTD.