HHS Public Access

Author manuscript

IEEE J Solid-State Circuits. Author manuscript; available in PMC 2020 November 01.

Published in final edited form as:

IEEE J Solid-State Circuits. 2019 November ; 54(11): 2957–2968. doi:10.1109/JSSC.2019.2941529.

A 512-Pixel, 51-kHz-Frame-Rate, Dual-Shank, Lens-less, Filter-less Single Photon Avalanche Diode CMOS Neural Imaging Probe

Jaebin Choi [Student Member, IEEE],
Electrical Engineering Department, Columbia University, New York, NY, USA

Adriaan J. Taal [Student Member, IEEE],
Electrical Engineering Department, Columbia University, New York, NY, USA

Eric H. Pollmann [Student Member, IEEE],
Electrical Engineering Department, Columbia University, New York, NY, USA

Changhyuk Lee [Member, IEEE],
Brain Science Institute, Korea Institute of Science and Technology, Seoul, South Korea

Kukjoo Kim,
Electronics and Telecommunications Research Institute, Daejeon, South Korea

Laurent C. Moreaux,
California Institute of Technology, Pasadena, CA, USA

Michael L. Roukes,
California Institute of Technology, Pasadena, CA, USA

Kenneth L. Shepard [Fellow, IEEE]
Bioelectronic Systems Laboratories, Columbia University, New York, NY, USA

Abstract

We present an implantable single photon shank-based imager, monolithically integrated onto a single CMOS IC. The imager comprises of 512 single photon avalanche diodes distributed along two shanks, with a 6-bit depth in-pixel memory and an on-chip digital-to-time converter. To scale down the system to a minimally invasive form factor, we substitute optical filtering and focusing elements with a time-gated, angle-sensitive detection system. The imager computationally reconstructs the position of fluorescent sources within a three-dimensional volume of 3.4 mm × 600 μm × 400 μm.

Keywords

Time-gated fluorescence imaging; Single photon avalanche diode; CMOS; Compressed sensing; Neural probe

I. Introduction

OPTICAL functional neural imaging has revolutionized neuroscience with optical reporters that enable single-cell-resolved monitoring of neuronal activity *in-vivo*. State-of-the-art microscopy methods, however, are fundamentally limited in imaging depth by optical scattering in tissue even with the use of the most advanced multiphoton microscopy techniques (Fig. 1a) [1, 2].

Instead, if a camera itself can be inserted into the brain, then imaging at arbitrary depth is possible. One method to enhance depth imaging from a multi-photon microscope is to implant a relay graded-index (GRIN) lens [3]. While these lenses allow one to couple laser power into deep nuclei at the expense of the cerebral tissue above the interrogation volume, and enable functional Calcium imaging, their displaced tissue volume is very large (> 0.1 mm^3) compared to the imaging volume (> 0.01 mm^3). These methods achieve deep brain imaging at the cost of significant neural network trauma and still require free-space microscopes for read-out.

In addition to solving the imaging depth problem, there is interest in scaling down the size of the epifluorescence microscopes themselves (Fig. 1b) [4, 5]. These efforts have primarily focused on small-form-factor implementations of full microscopes, including lenses and optical filters. This continued reliance on lenses and filters challenges miniaturization.

More recently, there has been significant efforts directed toward lens-less imaging. These approaches have been based on far-field masking (either phase or amplitude) to produce a spatially diverse illumination pattern on the imager which can be used to computationally reconstruct an image [6]. The requirement for far-field positioning of the mask, however, precludes use of these approaches in shanks, where shank thicknesses want to be maintained to a minimum to prevent tissue displacement. The requirement for optical spectral filters in the scheme of fluorescence microscopy adds to the probe thickness.

In this work, we take an entirely different approach in which the “camera” itself can be inserted into the brain, allowing imaging at arbitrary depth by collecting fluorescence signal very close to the neuron of interest, without its signal having to scatter and attenuate along its insertion depth. In particular, we present a monolithically integrated single-photon imager in the form of a CMOS, shank-based optical image sensor array that can be inserted into the brain (Fig. 1c) [7]. Single-photon avalanche diodes (SPADs) fabricated in CMOS are used as proximal light detectors. Lenses are replaced with near-field diffraction gratings placed on each SPAD, giving each of them a unique angularly modulated field of view (FoV) [8]. The mapping of SPAD counts into a volumetric scene allows the imager to determine the location of light sources. This additional information provides a multi-source localization capability similar to a far-field lens-less imaging approach. An external 480-nm pulsed pico-second laser delivers fluorescence excitation at the insertion point of the shanks by means of a fiber-coupled collimator. Spectral filters are replaced with time-gated operation of the SPADs in which detectors are turned on immediately after the pulsed excitation light has been shut off.

This paper is structured as follows. In Section II, we present the system design, and explain the post-processing into the shank form factor. The SPAD device and quenching circuit that allows for time-gated, filter-less fluorescence photometry is described in Section III. Section IV describes the implementation of lens-less imaging through angle-sensitive detection, and Section V demonstrates the computational multi-source localization of phantom targets.

II. System Overview

Fig. 2a shows the die photo of the implantable imaging probe with two shanks, each 4.1 mm-long and 120 μm-wide, connected to a 2-mm-by-1.2-mm head. Geometrically, this differs from our earlier design [7], extending the shanks by approximately 650 μm in length and reducing the shank separation from 250 μm to 175 μm to increase overlap of individual shank FoVs, aiding reconstruction performance. As a result, it is used to accommodate buffers and decoupling capacitors for delivering signal and power along the shank length. Each shank employs two rows of 128 SPADs operating in Geiger mode at 25.3 μm pitch and 7.7 μm active diameter, delivering a 6.3% fill factor. Each pixel contains a SPAD detector with a quenching circuit, 6-bit memory and addressing logic. Because the shanks have partially overlapping FoVs, a total imaging volume of 3.4 mm × 600 μm × 400 μm is achieved. The probe head contains a digital-to-time converter (DTC) for on-chip control of pixel gating and metal-oxide-metal (MOM) capacitors for decoupling the SPAD bias voltage. In subsequent discussions, we use the x direction to denote distance along the length of the shank, the z direction to denote vertical distance away from the shanks, and the y direction to denote distance along the width of the shank.

Fabricated in a 130-nm high-voltage process, the system is subdivided into four voltage domains including a 1.5-V digital core domain, a 3.3-V digital input and output (IO) domain, a 5-V domain for the SPAD quenching circuits, and a >16-V domain for biasing the SPAD cathodes.

A. System Architecture

A system level block diagram is shown in Fig. 2b. An FPGA is used to synchronize time-gated photon counting with an external pulsed laser. A phase-locked loop (PLL) implemented within the FPGA uses a photodetector-produced laser pulse waveform to deliver a synchronization signal to the DTC, which then produces a global time-gate signal with programmable delay and pulse width. Photon counts are accumulated with this global shutter for a desired number of laser pulses and read out by addressing each pixel's 6-bit memory serially with a 31.25 MHz read-out clock.

This read-out clock frequency is ultimately limited by the rate at which data can be read from the pixel-level memory. A per-pixel sequential read-out rate of 31.25 MHz, combined with a minimum image acquisition time of 3.2 μs when used in synchronization with a 20 MHz laser repetition rate, leads to a maximum possible frame rate of 51 kilo-frames per second (kfps). The 6-bit dynamic range at this frame rate can be maximally utilized at high signal conditions, which corresponds to a 520-nm wavelength irradiance of 2.2×10^6 photons/s/um^2 at the pixel surface. The data is streamed out with a USB 3.0 link, and image reconstruction is performed on an external computer.

B. Post-processing

In order to further miniaturize the imaging probe's volume displacement, excess silicon surrounding the shank is removed to yield a 40 μm-thin structure ready for implantation. Fig. 3a shows the fabrication sequence starting with (i) photoresist (AZ P4620, MicroChemicals) patterning, followed by an O_2-CF_4 reactive ion etch through the (ii) passivation and (iii) CMOS stack. (iv) A Bosch process etches through 100 μm of silicon substrate, after which (v) photoresist is stripped, and (vi) the die is mechanically milled (X-PREP, Allied High Tech Products) to remove 250 μm of substrate silicon, fully releasing the shanks (Fig. 3b). The imaging probe is finally wire-bonded and packaged to a printed circuit board with a flat flexible cable connection and passivated with an epoxy (Fig. 3b inset).

III. Time-gated Fluorescence Imaging

Optical multi-dielectric spectral filters are an important part of any fluorescence microscope. They are required to reject the excitation intensity and allow only the fluorescence response to reach the detector. It is typically accomplished in the spectral domain by three components, an emission filter, an excitation filter, and a dichroic mirror, combined usually in a "cube" to provide an optical density (OD) > 6 rejection of the excitation wavelength. Achieving the same levels of rejection with an integrated filter alone on our shank-based imagers remains difficult and is compounded by the dependency of filter properties on incident angle in the case of thin-film interference filters. Pigment-based absorption filters, as commonly used in color displays and CMOS imagers, do not have adequate rejection ratios [9]. Interference filters are difficult to integrate on CMOS and have problems in the rejection of non-orthogonal incident light [10, 11].

As an alternative to spectral filters, time-domain rejection is a powerful alternative by time-stamping the excitation source (pulse) and time-gating the detectors such that the fluorescence response is detected after the excitation has been turned off. This relies on the fact that fluorescent reporters have a time decay of fluorescence intensity characterized by their lifetimes. For commonly used fluorescent dyes, these lifetimes are in the 1–10 ns range (4.1 ns for EGFP, 4.0 ns for fluorescein, and 1.68 ns for Rhodamine B). The signal-to-background ratio (SBR) of a time-gated fluorescence image is determined by the ratio of the integrated photon count coming from the fluorescence emission to that remnant from the excitation source. The exact position of the time-gate relative to the laser turn-off time is chosen to maximize this SBR for the given laser turn-off and lifetime characteristics of the fluorescence.

A. Pixel Quenching and Reset Circuit

The in-pixel quench-and-reset circuit and pixel layout are shown in Figs. 4a,c. Each of the 512 SPADs is reverse biased between a global cathode and individual anode (*AN*). The cathode is held at a constant voltage above breakdown (V_{BD}), while the anode (*AN*) moves between zero and the reset voltage (V_{RST}) to put the SPAD into and out of Geiger mode. The rising edge of the *ON* signal provided by the DTC triggers edge detection. The large reset NMOS transistor (M1) begins discharging *AN* until it drops below the input threshold (~750 mV) of the inverter-based comparator. When the inverter output goes high, the

internal feedback clears the flip-flop-based edge detector, turning M1 off and setting *AN* to a high-impedance state. The SPAD is then biased in Geiger mode and ready for avalanche breakdown at an incoming photon. The self-timed reset, a design change from the earlier version [7], optimizes the pulse width of the *RST* signal and internally minimizes the reset time for an arbitrary value of between 1.5 V and 5 V without the need of an externally programmed pulse.

The long-channel half-latch transistor (M2) provides high-impedance at the *AN* terminal of the SPAD. During avalanche breakdown, a large current flow causes a build-up of voltage at the input of the comparator. The output of the comparator is flipped and triggers the event detection flip-flop, incrementing the counter by one. Only events within the *ON* signal are detected and at most one photon can be detected per cycle. When the counter reaches 63 counts, the *Full* signal is asserted, and the counter stops incrementing. The timing diagram in Fig. 4b demonstrates a cycle with zero counts and a cycle with one count. During the zero-count cycle, the internal delay of the self-timing reset mechanism and the falling edge of *ON* show the “effective” time-gate. This can be compensated with minor adjustments to the duty-cycle of the *ON* signal. During the one-count cycle, the avalanche is quenched and the counter is incremented by one.

B. Pixel Performance

The P-type (N-implant) shallow-junction SPAD was manufactured in a 130-nm process with a V_{BD} of 15.5 V and its key performance metrics were measured in an integrating sphere [7]. In Fig. 5a, the cumulative percentage of pixels under a given dark count rate (DCR) is plotted for increasing overvoltage ($V_{OV} = V_{CATH} - V_{BD}$), showing a “hot-pixel” percentage of1 2%; hot pixels are defined as those with more than five times the median DCR. The median of the array-wise photon detection probability (PDP) is plotted as a function of wavelength in Fig. 5b. The SPADs show a peak sensitivity at 575 nm.

Figs. 5a,b show an increase in both DCR and PDP, respectively, with increasing V_{OV}, as measured with a monochromatic light source centered at 520 nm which matches the emission peak of EGFP. We use a figure-of-merit (FoM: Eqn. 2 of [12]) based on photon counting statistics for fluorescence imaging to evaluate SPAD performance. This FoM selects the highest probability of detecting photon events while avoiding false positives and noise events. In Fig. 5c, this FoM is plotted for fluences over a range of 10^3-10^5 photons per second, equivalent to a photon flux of 0.4–40 fW/μm^2 at 520-nm wavelength. Due to the relatively low DCR compared to fluence, the optimal operating value of V_{OV} is found to be 1 V. Fig. 5d demonstrates the linearity of photon counts in response to increasing fluence at 520 nm. The linearity is limited by the dark count at low fluence and by photon pile-up at high fluence.

V_{OV} at 1 V also has the benefit of reduced afterpulsing events [13]. Afterpulsing is estimated by measuring DCR while sweeping the active-reset SPAD dead time from 300 ns to 10 ns. The median 40-Hz dark count holds constant across this range, indicating that afterpulsing is not a significant contributor to DCR.

Table I reports comparison with prior time-gated SPAD imagers. While much of the novelty of this work lies in the form factor and application of this imager, this design compares favorably with other time-gated SPAD imagers in conventional planar array formats [14–17], while displacing a volume of only 0.03 mm^3 with the implanted shanks.

C. Time-gating Circuitry

The system employs a DTC to synthesize the global shutter *ON* signal with a tunable delay with respect to the laser pulse. Fig. 6 shows its architecture and timing diagram. *TRIGGER* determines the *ON* repetition rate. The Johnson counter creates an 8-bit sequence signal, *CNT*, where each bit corresponds to an 8×-down-sampled version of *CLKIN*, each 45 degrees delayed and fed into both phase interpolators (PIs). A user-defined 8-bit programmed phase is decoded to select any combination of three phases of *CNT*. These are equally summed into Σ (see Fig. 2b), low-pass filtered and buffered [18]. Whereas *CLK* is generated in the first PI, *REF* is generated from a second PI to flexibly compensate for internal delays, and its rising edge asserts *READY* to allow for the next synthesized *CLK* to propagate to the output *ON*. *READY* is pulled low when *ON* is activated and asserted again on the rising edge of *REF* when *TRIGGER* is high. The dynamic range of pulse width and delay is extended through the final tunable counter stage. A serial peripheral interface (SPI) slave programs phase, width and delay from the control FPGA.

A variable output duty cycle of 20–50% for *ON* allows the SPAD to be moved out of Geiger mode after sufficient decay of the fluorescence signal, avoiding dark count and unnecessary power dissipation. As *CLK* and *REF* are 8×-down-sampled, 50% duty cycle, phase-delayed versions of *CLKIN*, the *CLKIN* frequency determines the duty cycle (D) resolution of *ON*. The *CLKIN* frequency must always be greater than or equal to $8 \times 8 \times \frac{F_{TRIGGER}}{D}$, where $F_{TRIGGER}$ is equal to the laser repetition rate, in order to ensure correct phase and duty cycle for *ON*, with higher frequencies allowing higher resolution phase control.

All tests were performed at a 10 MHz *TRIGGER* repetition rate. If one keeps the *REF* edge constant, the DTC allows *CLK* timing to be tuned relative to *REF* with an 8-bit resolution in a range of the *REF* and *CLK* period.

In the case of fluorescence decay lifetimes on the order of 2–4 nanoseconds, *REF*-to-*CLK* delays up to 12.5 ns are required, equivalent to 50% of the available range (seven bits of the DTC). *CLKIN* is set to a 320 MHz frequency to synthesize a 40 MHz *CLK*, allowing for a 25% duty cycle for *ON*. Fig. 6c shows how 128 LSB (seven bits) for *CLK* results in 97-ps resolution. Fig. 6d shows the measured linearity performance of the converter, with DNL staying within [−1, +2.5] LSB and INL staying within [−4, +2] LSB, with standard deviations over the nonlinearity curve of 0.6 and 1.3 LSB, respectively. From the measurement data in Fig. 6c, we calculate the signal-to-noise-and-distortion (SINAD) of the DTC as [19]:

$$SINAD = 20\log_{10}\left(\frac{T_{fs}}{\sigma_\epsilon}\right) = 40.95\text{dB} \quad (1)$$

where T_{fs} is the full-scale conversion time (12.5 ns), and σ_ϵ = 0.112 ns represents the standard deviation of the output error. This leads to an effective number of bits (ENOB) of 6.5 bit, or an effective resolution of 137 ps.

D. Time-gating Performance

To demonstrate the effectiveness of time-gated fluorescence imaging, the imaging probe was placed under media with properties similar to a fluorescence-expressing mouse brain (Fig. 7a). We used a 0.6% agarose gel (A7777, Teknova) to imitate optical scattering of neural tissue [20–22], and a range of fluorescein concentrations were added to it to simulate bulk fluorescence expression (Fig. 7b). A fluorescein concentration of 10 μM is equivalent in brightness to cytoplasmic expression of GCaMP in a neuron [23]. A picosecond-pulsed laser (Fianium SC450-pp, NKT Photonics) delivers excitation light through a collimated beam 500 μM above the imager with 670 μW average power (1.6×10^6 photon/s), and the time-gate *ON* signal with a 10-ns on-time is moved in steps of 280 ps over the range of 12 ns after the laser pulse.

In the scattering medium, the excitation laser light reaches the pixels after traveling along a Mie scattering path, which we find representative of neural tissue without the presence of fluorescence. The resulting photon count (Fig. 7c, *Agar*) is the convolution of the SPADs' 10 ns time-gate with the 50 ps-wide laser pulse, which is denoted as the instrument response function (IRF) of the time-gated quenching circuit. Its sub-ns exponential decay characteristic is the result of minority carriers, created by the high excitation photon flux, diffusing into the multiplication region of the SPAD and creating an avalanche breakdown in the absence of an incoming photon [14]. The time-gated rejection ratio 2 ns after the pulse is 98%, corresponding to an excitation OD of 1.7.

In the case of a fluorescent medium representative of an EGFP-expressing mouse cortex (Fig. 7c, *FL*), direct laser intensity is not observed on the imager because most of the laser light is converted to fluorescence emission before it reaches the imager. Instead, we see a slow rise and 4.7 ns lifetime fall produced by a convolution of the excitation light time-gate and the 4.1 ns fluorescence lifetime of fluorescein.

Time-gated fluorescence imaging of fluorescent markers in a scattering environment requires a time-window in which the fluorescence response exceeds that of the scattered excitation light. In the case of 10 μM fluorescein in agarose, the SBR reaches 31.6 dB two nanoseconds after the laser pulse. SBR and excitation optical density (OD) continue to increase past 2 ns, reaching 105 dB and 3.2 OD, respectively, at 12 ns, however, this is achieved at the cost of reduced signal yield (5% at 12 ns). While time-gated filtering alone does not match the OD of epifluorescent microscopes, relatively low-OD spectral filters can be added on top of the imager to supplement the rejection from time gating.

When the imager is illuminated with the output-saturating photon flux of 2.2×10^6 photons/s/um^2 and reads out at the maximum frame rate of 51 kfps, operating at maximum power consumption, it consumes a total of 6.24 mW. The charging of the SPAD diode capacitance at quenching events consumes 2.94 mW on the V_{RST} and SPAD Cathode nodes together. The 1.5 V digital core consumes 3.3 mW. The 3.3 V IO power was supplied

through the FPGA. With the majority of the digital logic located on the base of the shank, we estimate a power consumption of ~3 mW for the implantable regions of the imager, comparable to levels reported by other implantable CMOS neural recording shanks [24].

E. Photometry in Scattering Neural Tissue

To test the imaging performance on a fluorescent target with known size and properties, a Monte Carlo photon trajectory simulation [25–27] is performed in a tissue model characteristic of EGFP expression in gray matter. The tissue is modeled with anisotropy factor g=0.88, refractive index n=1.37, scattering coefficient μ_s=21 mm^{-1}, and absorption coefficient μ_a =0.06 mm^{-1} [28], while EGFP is modeled with a quantum yield of 0.6, extinction coefficient of 55000 $M^{-1}cm^{-1}$ [29], and lifetime of 4.1 ns. Using a simulation setup similar to Fig. 7a, we model an imager implanted along the x axis with pixels located between x=0 and 3.2 mm, in parallel with a collimated pulsed laser light source illuminating the tissue boundary of x=0 μm, at a distance z=100 μm away from the imager with 700 μW average power. A group of fluorescently labeled somata contained in a 50 μm-radius located at height z=100 μm, expressing EGFP with an equivalent concentration of 10 μM [23], is swept along the x axis. The model assumes a duty-cycle of 50% and a 12-ns time-gate delay after the excitation pulse, which is the shortest delay at which SBR exceeds one at all depths. At this delay the filter achieves an OD of 3.2 and the fluorescence emission falls to 5% of its peak intensity.

We perform these simulations at a frame rate of 1 fps. For a laser repetition rate of 20 MHz, this means that photon counts for 2×10^7 laser pulses are integrated. Fig. 8a shows resulting counts for the time-gated excitation source, the fluorescent emission, and the dark count. The fluorescent emission assumes that the group of somata is positioned directly above the sensing SPAD at each depth. Fluorescence emission is found to overcome scattered excitation at all simulated depths, although by a small margin. Fig. 8b plots the signal-to-noise-and-background-ratio (SNBR), defined as the ratio of fluorescent signal counts to the background, photon-shot-noise and dark-count-shot-noise. For time gate delays longer than 12 ns, the drop in signal yield for a given frame rate leads to an increase in shot noise relative to signal. A frame rate of 1 fps is necessary to reduce the shot noise to yield SNBR larger than 3 dB. SNBR is largest at tissue depths below 500 μm because scattering has not yet become significant. When the imager is deeper in tissue, excitation collimation is lost and more blue excitation light scatters into the detector. However, due to the limited acceptance angle of 60° (see Section IV.A), the majority is still rejected, and the excitation background and fluorescent signal reduce proportionally between 500 μm and 2.5 mm.

Beyond 2.5 mm, exponential extinction of the excitation light places the detector into dark-count-limited operation. In the absence of better excitation rejection, the SNBR of our imager is sufficient up to a depth limit, and deep brain imaging becomes a problem of light delivery. Potential solutions have been demonstrated including implantable waveguides [30–32], densely integrated GaN μLEDs [33, 34], or multimode fiber endoscopes [35].

Another solution is to incorporate spectral filters that work together with time gating. In Fig. 8b, we repeat the same analysis with the addition of a filter which is able to block the excitation wavelength at an OD of only 0.6 [36], but which acts to further reduce the

background and photon shot noise associated with this background. The SNBR in this case is sufficient to allow imaging at full shank depths. Imaging at the hardware limit of 51 kfps would be possible given high enough illumination. The frame rate is in direct trade-off with shot-noise-limited SNBR.

IV. Lens-less Imaging Using Near-field Angle-sensitive Detection

The lack of refractive focusing optics poses two challenges in image formation. The first is low photon yield, which is addressed in part through the use of sensitive SPAD detectors. The second is the lack of spatial resolution. Simply relying on the limited numerical aperture of the detectors in the array itself (which is an acceptance cone of approximately 60° single-sided) would allow an x- or y- spatial resolution of only ~450 μm to be achieved 200 μm from the detector. Instead, we rely on near-field diffraction gratings on each pixel to provide an angle sensitivity diversity that can be used for computational image reconstruction [16]. Use of these gratings comes at a cost in transmission efficiency, which averages only 3% of incident photons. This can be improved to 29% using non-blocking dielectric phase gratings to achieve the same diffraction gratings instead of metal wires [37].

The inverse imaging problem in this case is formulated as a linear system from the scene x to the array photon count y:

$$y = Ax + \epsilon \tag{2}$$

The sensing matrix A compresses the scene into a much lower dimensionality determined by the pixel count. Dark count and background ϵ are assumed to be spatially and temporally uncorrelated to the sources, and afterpulsing was estimated to be negligible. Image quality is improved by constructing a sensing matrix A that maximizes information extraction from the scene x. Such an optimized A allows for the best reconstruction of the inverted image $\hat{x}$. In particular, we seek an A that has maximally incoherent columns [38], indicating that each location in space is compressed onto the imager's response in a maximally distinct way.

A. Angle-sensitive Pixel Design

A is determined by a diversity of angle-sensitivity introduced into each pixel with near-field diffraction gratings as shown in Fig. 9. These gratings, formed in two layers of back-end metal separated by 1.14 μm (Fig. 9a), constitute sixteen angular-sensitivity variations (which take up a total distance along the length of the shank of 200 μm) consisting of orthogonal combinations of two angular modulation frequencies, two rotations, and four quadrature phases [8]. The gratings are employed across both shanks in a repeating pattern of two by eight pixels to ensure that a source located 100 μm above the center axis of the shank and positioned (in x) at the midpoint of the 16-pixel group, is situated in the 60° FoV of each pixel variation.

The angular response for each grating combination was measured by illuminating the imager with a collimated monochromatic light source at 520 nm and varying the angles of incidence along θ and φ, two azimuthal angles pivoting around the x and y axes respectively, using a rotational and tilting stage under the imager (Fig. 9b). Two examples of grating structures,

each with a low and high angular frequency, are shown in Fig. 9c, along with the parametric fit.

The higher angular modulation amplitudes at extreme angles in Fig. 9b reflect the fact that the Talbot self-image is not generated at the exact height of the secondary grating. At even higher incident angles, the CMOS metal interconnect starts occluding photons for incident angles larger than 45°, and rejects completely at 60°. Additionally, the PDP at green wavelengths is lower for obliquely incident photons due to a deeper effective multiplication layer, shifting peak PDP away to a higher wavelength.

B. Point Spread Function

The point spread function (PSF) describes the system response to a point source, and its width determines how close two point sources can be placed while still being individually reconstructed. The PSF of this imaging system can be computed as a pseudoinverse backprojection of a single voxel:

$$x_{PSF} = \left(A^{+}A\right)x_{source} \tag{3}$$

where A^{+} is the left pseudoinverse of x_{source} and n_{voxel} = is an n_{voxel}-by-1 array which is zero except at a single voxel. n_{voxel} is ~2 million in the case of using a voxel grid of 5 μm. The PSF of a single voxel located at Cartesian coordinates of [2000 μm, 0 m, 200 μm] is plotted in Figs. 9a–c, showing the correlated nature of voxels in this highly underdetermined system. Fig. 10d displays the resolution in the form of full-width at half-maximum (FWHM) of the one-dimensional PSF profile. For continuity of the plot, data points in which the local minima closest to the peak did not fall below the half-maximum were omitted.

The broadening of the FWHM with increasing source distance results from an approximately constant angular resolution translating to a larger spatial spread at further distances. The minimum resolution is measured to be 64 μm, 26 μm, and 65 μm in the x, y, z directions respectively. In comparison to a GRIN lens based miniature microscope [3], while the imager resolution is 1–2 orders of magnitude worse, it images a brain volume larger by the same factor. The imaging volume can be sculpted further by restricting the illumination profile to a volume of interest, drastically reducing the image background.

V. Imaging Performance

A. Single Point Source Imaging

To mimic a cluster of neuron somata, a diffuser tip of 100 μm radius (Thorlabs CFDSB20), coupled to a monochromatic green LED (300 μW, 7.9×10^{14} photons/s), was placed 200 μm above the imager and translated along its length at three different positions (Fig. 11, *p1, p2, p3*). The image was acquired with a 0.8 fps frame rate and photon accumulation time of 200 ms, yielding a maximum of 41k counts at the pixel directly under the light source. The repeating pattern of sixteen gratings is responsible for the periodicity of raw counts in Fig. 11a and contribute to localization of the point source. Pseudoinverse backprojections ($x_{proj} = A^{+}y$) sectioned at z=200 μm show the imager's ability to find centers of brightness within a 30 μm standard deviation (Fig. 11b).

B. Multisource Phantom Imaging

A multisource phantom fluorescent target was constructed by depositing fluorescent microspheres (F8836 10 μm, Thermofisher) on a cover glass, which was then placed 400 μm away from the imager. A picosecond-pulsed excitation (Fianium SC450-pp, NKT Photonics) centered at 480 nm and carrying 700 μW of average power (1.7×10^{15} photons/s) with a 20-MHz repetition rate was delivered in a direction parallel to the shank with a 500 μm-diameter collimator. This parallel incident angle for the excitation power minimizes the image background produced by incomplete time-gate rejection. Figs. 11a,b show two distinct scenes with a difference in the size and arrangement of the imaging target. The fluorescence microscope image is displayed alongside the reconstructed pseudoinverse backprojections of SPAD counts. Photon counts were accumulated for 200 ms at 1 fps with a time-gate delay of 550 ps after the laser pulse, resulting in a maximum SBR of 15 dB. Voxel sizes are 20 μm in x, y, and z dimensions to optimize for computation speed.

To show the efficacy of time-gated fluorescence imaging, a third scene (Fig. 12c) is constructed with two microsphere clusters, one composed of fluorescent spheres and another with scattering latex spheres. Despite the visibility of both clusters under brightfield imaging, the imager selectively detects only the fluorescent microspheres.

C. Source Localization by L1-norm Minimization

Due to the highly compressed nature of the imaging system which maps ~10^5 voxels to 512 pixels when constructing a three-dimensional (3D) image with 20 μm voxels, a direct backprojection image will inevitably have limited spatial resolution and appear blurred. Furthermore, high frequency noise is amplified into the image through high spatial sampling frequencies, associated with low magnitude singular values of the sensing matrix A. By incorporating a sparsity constraint, the brightest few source locations can be solved for in a least-mean-squares manner. Performing this basis pursuit denoising optimization allows an estimation of best-fit source locations while penalizing the number of total sources with a weighting parameter λ [39]. This optimization problem employs a cost function that includes the L1-norm of sources (f_{L1}) and the L2-norm of the residual error (f_{L2}):

$$\hat{x} = \underset{x}{\operatorname{argmin}}\left(\lambda\left\|x\right\|_1 + \frac{1}{2}\|Ax - y\|_2^2\right) = \underset{x}{\operatorname{argmin}}(f_{L1} + f_{L2}) \tag{4}$$

To solve this optimization problem, we use the in-crowd algorithm [40] to solve for sparse locations within the imager's FoV in real time. Fig. 13a shows the 3D reconstructed image of the microsphere arrangement in Fig. 12b solved with 20 μm voxel sizes. Raw imager data is drawn as a heat map at the plane of the imager (z=0 μm). The largest magnitude contributors to the solution ($\hat{x}$) are displayed in circles, while circle diameters indicate the magnitude estimate. The two voxels with largest magnitude, colored red, show the imager's ability to locate the two microsphere clusters at their correct volumetric locations. The next highest 40, colored black, are also found to gravitate around the two clusters with decreasing amplitude estimates. As more voxels are added to the solution $\hat{x}$, the L2-norm of the residual error (f_{L2}) declines as shown in Fig. 13b, but each with a diminishing contribution to

residual error minimization. By combining this information from residual error with the appearance of clusters of estimated sources, we can heuristically infer sparsity and location without relying on an external verification of ground truth.

D. Limit of Reconstruction Accuracy

We perform a Monte Carlo simulation to determine the maximum number of sources the imaging probe can resolve simultaneously. Sparse reconstruction theory states that the maximum number of detectable sources is the number of pixels M. However, for L1-minimization solvers that assume no constraints on the sensing matrix A, phase transition occurs around $\sqrt{M}$ [41], meaning that a sharp decline in reconstruction reliability occurs above that density. In the case of the imaging probe, the columns of $\sqrt{M}$ have a repetitive structure determined by the angular sensitive gratings. As a result, we expect the maximum allowable sparsity to be below $\sqrt{M}$.

We place K sources randomly in the imaging volume of 3.2 mm × 200 μm × 150 μm above a single shank at 250 μm depth. Photon shot noise is applied to the simulated fluorescent data, and additional excitation background is added with the expected SBR of 3.5 dB. Source placement and subsequent localization is repeated N=1000 for each value of K. To correctly compare the localization result against the ground truth, an exhaustive search method links found sources with their closest true locations, where computational tractability limits the analysis to K=10. We analyze the localization error in terms of mean square error (MSE), allowing a decomposition into the average error squared and variance of the error:

$$MSE(x,\hat{x}) = \frac{1}{N}\|x-\hat{x}\|_2^2 = \left(\frac{1}{N}\sum(x-\hat{x})\right)^2 + \sigma^2_{x-\hat{x}} \tag{5}$$

Fig. 14 shows how localization performance is affected by the source cardinality, K. Fig. 14a shows the distribution of the error over all iterations and source cardinalities in each dimension. Large outliers are essentially ground truth locations that are not found, and the larger imaging volume in the x direction permits larger outliers along that axis. The error is distributed with 98% of sources found within 130 μm, or around two times the largest FWHM at 100 μm distance, of the true location. The compressive sensing method has a 2% likelihood for introducing false positives and negatives, which need to be accounted for in subsequent imaging target analyses.

The total ensemble standard deviation of all errors in all directions reduces to 17 μm (Fig. 14b) when corrected for these outliers. The error increases with more sources due to increased noise sensitivity at higher source densities. Both the standard deviation and bias of the error (Fig. 14c) are smaller than the pixel pitch in all directions x, y, and z. The sign of the data point indicates an estimation bias in the respective direction.

Even at the reported large variance, a low error bias indicates the imager's strength at localizing volumetric centers of brightness with micrometer accuracy, where precision improves with repeated measurements and averaging over multiple reconstructions. We, therefore, expect that a single shank with 256 pixels can localize up to ten point sources

simultaneously with 98% certainty at SBR levels comparable with those estimated for an *in-vivo* environment.

VI. Conclusion

In this work, we present an implantable neural shank-based imager with minimal tissue displacement of only 0.03 mm^3, while monitoring a volume of 0.4 mm^3. The minimally invasive imager is able to localize fluorescent objects at depths beyond those of conventional imaging systems. It is capable of distinguishing multiple fluorescent objects with a resolution of 64 μm, 26 μm, and 65 μm in the x, y, and z directions, respectively, without the use of spectral filters and refractive focusing optics. Time-gating adds two orders of magnitude of excitation light rejection, adding to what can be provided by thin-film absorptive optical filters. The imager's modest pixel count of 512 trades off spatial resolution with fast image acquisition; its high maximum frame rate (51 kfps), synchronized with an external pulsed excitation source, allows the investigation of sparsely expressed biomarkers for in vivo neural imaging.

Acknowledgment

This work was supported by the National Institutes of Health under Grant U01NS090596, by the Defense Advanced Research Projects Agency (DARPA) under Contract N66001–17-C-4012, and by the U. S. Army Research Laboratory and the U. S. Army Research Office under Contract W911NF-12–1–0594. We would also like to thank TSMC Foundry for their full support in test chip fabrication, and Byeongjoo Ahn for helpful discussions on image reconstruction.

Biography

Jaebin Choi (S'14) received the B.S. degree in engineering physics from Cornell University, Ithaca, NY, in 2011 and the M.S. degree in electrical engineering from Columbia University, New York, NY, in 2012. He worked as a Research Scientist in the Korea Institute of Science and Technology, Seoul, Korea from 2014 to 2017, and is currently pursuing a Ph.D. degree in electrical engineering at Columbia University, New York, NY. His current research interest is in optoelectronic neural interfaces.

Adriaan J. Taal received the B.S. and M.S. degrees in electrical engineering from Delft University of Technology, Delft, the Netherlands, in 2014 and in 2017, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at Columbia University, New

York, NY. His research interests are in optoelectronic circuit design, computer vision, and developing imaging applications for neuroscience.

Eric H. Pollmann (S'19) received the B.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, in 2017 and the M.S. degree in electrical engineering from Columbia University, New York, NY, in 2018. He is currently pursuing the Ph.D. degree in electrical engineering at Columbia University, New York, NY. His research interests are in implantable CMOS fluorescence imagers for applications in biology and neuroscience.

Changhyuk Lee (S'08–M'14) is currently a Senior Research Scientist at Brain Science Institute of Korea Institute of Science and Technology (KIST). He received the B.S. degree in electrical and computer engineering from Korea University, Seoul, Korea, in 2005, and the M.S. and Ph.D. degrees from Cornell University, Ithaca, NY, USA. Before starting his research at Cornell, he worked at Samsung, Chang-won from 2004 to 2006, and he also worked for Com2us, Seoul, Korea, from 2007 to 2008. His Ph.D. work focuses on the light field (Angle Sensitive) photon counting CMOS image sensors, neural interface, and RFIC. He is a recipient of the Qualcomm Innovation Fellowship 2013–2014. His current research interests include optoelectronic neural interfaces, implantable biomedical devices, computational neuroscience and behavioral studies.

Kukjoo Kim received the B.S. and Ph.D. degrees in 2008 and 2014, respectively, in electrical engineering at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. His thesis was on solution-processed graphene and its electronic applications. He continued as a postdoctoral scholar in the lab of Prof. Kenneth Shepard in Columbia University, New York, NY, in 2016, and is now working as a Senior Researcher at the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests include flexible/stretchable electronics and neural interfaces.

Laurent C. Moreaux is a permanent Research Scientist for the French National Center for Scientific Research (CNRS), and also a Senior Research Scientist at the California Institute of Technology. He earned an optoelectronics degree from Polytech Paris-Sud in Orsay in 1997, an M.S. degree in optics and photonics from the Institut d'Optique in Orsay in 1998, and his Ph.D. degree in physics at the ESPCI Paris with Drs. Jerome Mertz and Serge Charpak focusing on the development of novel, non-linear optical microscopy tools and methods for biological imaging. He continued as a postdoctoral scholar in the lab of Dr. Gilles Laurent in the Caltech Division of Biology in 2002, and was awarded a permanent Research Scientist position with the CNRS in 2004, where he continued working on the development of optical techniques to probe neuronal activity in several model systems. He returned to Caltech in 2009, where he has since held positions in both the Division of Biology and Bioengineering as well as in the Division of Physics, Mathematics and Astronomy. He has contributed his technical expertise to many projects in neuroscience at Caltech, including developing methods for whole-cell recordings combined with extracellular arrays and two-photon microscopy in the hippocampus of awake and behaving mice. With the current depth limitation of free space optical microscopy, he believes that the use of integrated photonics will allow us to directly bring the microscope "into" the brain to probe the activities of an unprecedented number of neurons. With this vision in mind, his most recent efforts have focused upon fundamental conceptualization, device design, and initial characterization of novel integrated neurophotonic probes.

Michael L. Roukes received B.A. degrees in both physics and chemistry in 1978 from the University of California, Santa Cruz, CA; and the M.S. and Ph.D. degrees in physics in 1985 from Cornell University, Ithaca, NY. Thereafter, he was a Member of Technical Staff/ Principal Investigator at Bell Communications Research, Red Bank, NJ. He is currently Frank J. Roshek Professor of Physics, Applied Physics, and Bioengineering at the California Institute of Technology (Caltech), Pasadena, CA, where he has been since Fall 1992. His research has spanned mesoscopic electron and phonon transport at ultralow temperatures, physics of nanoelectromechanical systems (NEMS), biomolecular detection and imaging, ultrasensitive NEMS-based mass spectrometry, and ultrasensitive calorimetry. Dr. Roukes is a Fellow of the American Physical Society, was an NIH Director's Pioneer Awardee, and was named Chevalier (Knight) of the Ordre des Palmes Académiques by the Republic of France.

Kenneth L. Shepard (M'91–SM'03–F'08) received the B.S.E. degree from Princeton University, Princeton, NJ, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1988 and 1992, respectively. From 1992 to 1997, he was a Research Staff Member and the Manager with the VLSI Design Department, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, where he was responsible for the design methodology for IBM's G4S/390 microprocessors. He was the Chief Technology Officer with CadMOS Design Technology, San Jose, CA, until its acquisition by Cadence Design Systems in 2001. Since 1997, he has been with Columbia University, New York, NY, where he is currently the Lau Family Professor of Electrical Engineering and Biomedical Engineering and the Co-Founder and the Chairman of the Board of Ferric, Inc., New York, NY, which is commercializing technology for integrated voltage regulators. His current research interests include power electronics, carbon-based devices and circuits, and CMOS bioelectronics. He has been an Associate Editor of the IEEE TRANSACTIONS ON VERY LARGE-SCALE INTEGRATION SYSTEMS, the IEEE JOURNAL OF SOLID-STATE CIRCUITS, and the IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.

References

[1]. Weisenburger S. et al., "Volumetric Ca(2+) Imaging in the Mouse Brain Using Hybrid Multiplexed Sculpted Light Microscopy," Cell, vol. 177, no. 4, pp. 1050–1066 e14, 5 2 2019. [PubMed: 30982596]

[2]. Wang T. et al., "Three-photon imaging of mouse brain structure and function through the intact skull," (in English), Nat Methods, vol. 15, no. 10, pp. 789–792, 10 2018. [PubMed: 30202059]

[3]. Levene MJ, Dombeck DA, Kasischke KA, Molloy RP, and Webb WW, "In vivo multiphoton microscopy of deep brain tissue," (in English), J Neurophysiol, vol. 91, no. 4, pp. 1908–12, 4 2004. [PubMed: 14668300]

[4]. Skocek O. et al., "Author Correction: High-speed volumetric imaging of neuronal activity in freely moving rodents," (in English), Nat Methods, vol. 15, no. 6, p. 469, 6 2018. [PubMed: 29786093]

[5]. Corder G, Ahanonu B, Grewe BF, Wang D, Schnitzer MJ, and Scherrer G, "An amygdalar neural ensemble that encodes the unpleasantness of pain," (in English), Science, vol. 363, no. 6424, pp. 276–281, 1 18 2019. [PubMed: 30655440]

[6]. Adams JK. et al., "Single-frame 3D fluorescence microscopy with ultraminiature lensless FlatScope," (in English), Sci Adv, vol. 3, no. 12, p. e1701548, 12 2017. [PubMed: 29226243]

[7]. Lee C. et al., "11.5 A 512-Pixel 3kHz-Frame-Rate Dual-Shank Lens-less Filterless Single-Photon-Avalanche-Diode CMOS Neural Imaging Probe," in 2019 IEEE International Solid-State Circuits Conference-(ISSCC), 2019: IEEE, pp. 198–200.

[8]. Wang A, Gill P, and Molnar A, "Light field image sensors based on the Talbot effect," (in English), Appl Opt, vol. 48, no. 31, pp. 5897–905, 11 1 2009. [PubMed: 19881658]

[9]. Koo HS, Chen M, and Pan PC, "LCD-based color filter films fabricated by a pigment-based colorant photo resist inks and printing technology," (in English), Thin Solid Films, vol. 515, no. 3, pp. 896–901, 11 25 2006.

[10]. Yu Y, Wen L, Song SC, and Chen Q, "Transmissive/Reflective Structural Color Filters: Theory and Applications," (in English), J Nanomater,2014.

[11]. Gu Y, Zhang L, Yang JK, Yeo SP, and Qiu CW, "Color generation via subwavelength plasmonic nanostructures," Nanoscale, vol. 7, no. 15, pp. 6409–19, 4 21 2015. [PubMed: 25800353]

[12]. Field RM, Lary J, Cohn J, Paninski L, and Shepard KL, "A low-noise, single-photon avalanche diode in standard 0.13-μm complementary metal-oxide-semiconductor process," Applied Physics Letters, vol. 97, no. 21, 2010.

[13]. Bronzi D, Tisa S, Villa F, Bellisai S, Tosi A, and Zappa F, "Fast Sensing and Quenching of CMOS SPADs for Minimal Afterpulsing Effects," (in English), Ieee Photonics Technology Letters, vol. 25, no. 8, pp. 776–779, 4 15 2013.

[14]. Schwartz DE, Charbon E, and Shepard KL, "A Single-Photon Avalanche Diode Array for Fluorescence Lifetime Imaging Microscopy," (in English), IEEE J Solid-State Circuits, vol. 43, no. 11, pp. 2546–2557, 11 21 2008. [PubMed: 23976789]

[15]. Perenzoni M, Massari N, Perenzoni D, Gasparini L, and Stoppa D, "A 160 × 120 Pixel Analog-Counting Single-Photon Imager With Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging," (in English), Ieee J Solid-St Circ, vol. 51, no. 1, pp. 155–167, 1 2016.

[16]. Lee C, Johnson B, Jung T, and Molnar A, "A 72 × 60 Angle-Sensitive SPAD Imaging Array for Lens-less FLIM," Sensors (Basel), vol. 16, no. 9, 9 2 2016.

[17]. Field RM, Realov S, and Shepard KL, "A 100 fps, Time-Correlated Single-Photon-Counting-Based Fluorescence-Lifetime Imager in 130 nm CMOS," (in English), Ieee J Solid-St Circ, vol. 49, no. 4, pp. 867–880, 4 2014.

[18]. Tsimpos A, Souliotis G, Demartinos A, and Vlassis S, "All Digital Phase Interpolator," (in English), 2015 10th Ieee International Conference on Design & Technology of Integrated Systems in Nanoscale Era (Dtis),2015.

[19]. Cao Y, De Cock W, Steyaert M, and Leroux P, "1–1-1 MASH Delta Sigma Time-to-Digital Converters With 6 ps Resolution and Third-Order Noise-Shaping," (in English), Ieee J Solid-St Circ, vol. 47, no. 9, pp. 2093–2106, 9 2012.

[20]. Pogue BW. and Patterson MS, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," J Biomed Opt, vol. 11, no. 4, p. 041102, Jul-Aug 2006. [PubMed: 16965130]

[21]. Pomfret R, Miranpuri G, and Sillay K, "The substitute brain and the potential of the gel model," Ann Neurosci, vol. 20, no. 3, pp. 118–22, 7 2013. [PubMed: 25206029]

[22]. Sindhwani N, Ivanchenko O, Lueshen E, Prem K, and Linninger AA, "Methods for determining agent concentration profiles in agarose gel during convection-enhanced delivery," IEEE Trans Biomed Eng, vol. 58, no. 3, pp. 626–32, 3 2011. [PubMed: 21342811]

[23]. Akerboom J. et al., "Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design," J Biol Chem, vol. 284, no. 10, pp. 6455–64, 3 6 2009. [PubMed: 19098007]

[24]. Jun JJ. et al., "Fully integrated silicon probes for high-density recording of neural activity," Nature, vol. 551, no. 7679, pp. 232–236, 11 8 2017. [PubMed: 29120427]

[25]. Wang L, Jacques SL, and Zheng L, "MCML–Monte Carlo modeling of light transport in multi-layered tissues," Comput Methods Programs Biomed, vol. 47, no. 2, pp. 131–46, 7 1995. [PubMed: 7587160]

[26]. Pediredla AK. et al., "Deep imaging in scattering media with selective plane illumination microscopy," (in English), J Biomed Opt, vol. 21, no. 12, p. 126009, 12 1 2016. [PubMed: 27997019]

[27]. Fang Q, "Mesh-based Monte Carlo method using fast ray-tracing in Plucker coordinates," Biomed Opt Express, vol. 1, no. 1, pp. 165–75, 8 2 2010. [PubMed: 21170299]

[28]. Yona G, Meitav N, Kahn I, and Shoham S, "Realistic Numerical and Analytical Modeling of Light Scattering in Brain Tissue for Optogenetic Applications(1,2,3)," eNeuro, vol. 3, no. 1, Jan-Feb 2016.

[29]. Cranfill PJ. et al., "Quantitative assessment of fluorescent proteins," (in English), Nat Methods, vol. 13, no. 7, pp. 557–62, 7 2016. [PubMed: 27240257]

[30]. Zorzos AN, Boyden ES, and Fonstad CG, “Multiwaveguide implantable probe for light delivery to sets of distributed brain targets,” Opt Lett, vol. 35, no. 24, pp. 4133–5, 12 15 2010. [PubMed: 21165114]

[31]. Shim E, Chen Y, Masmanidis S, and Li M, “Multisite silicon neural probes with integrated silicon nitride waveguides and gratings for optogenetic applications,” Sci Rep, vol. 6, p. 22693, 3 4 2016. [PubMed: 26941111]

[32]. Reddy JW. and Chamanzar M, “Low-loss flexible Parylene photonic waveguides for optical implants,” (in English), Optics Letters, vol. 43, no. 17, pp. 4112–4115, 9 1 2018. [PubMed: 30160729]

[33]. Klein E, Gossler C, Paul O, and Ruther P, “High-Density muLED-Based Optical Cochlear Implant With Improved Thermomechanical Behavior,” Front Neurosci, vol. 12, p. 659, 2018. [PubMed: 30327585]

[34]. Wu F, Stark E, Ku PC, Wise KD, Buzsaki G, and Yoon E, “Monolithically Integrated muLEDs on Silicon Neural Probes for High-Resolution Optogenetic Studies in Behaving Animals,” Neuron, vol. 88, no. 6, pp. 1136–1148, 12 16 2015. [PubMed: 26627311]

[35]. Vasquez-Lopezet SA et al., “Subcellular spatial resolution achieved for deep-brain imaging in vivo using a minimally invasive multimode fiber,” Light Sci Appl, vol. 7, p. 110, 2018. [PubMed: 30588295]

[36]. Li ZY, Butun S, and Aydin K, “Large-Area, Lithography-Free Super Absorbers and Color Filters at Visible Frequencies Using Ultrathin Metallic Films,” (in English), Acs Photonics, vol. 2, no. 2, pp. 183–188, 2 2015.

[37]. Sivaramakrishnan S, Wang A, Gill P, and Molnar A, “Design and Characterization of Enhanced Angle Sensitive Pixels,” (in English), Ieee T Electron Dev, vol. 63, no. 1, pp. 113–119, 1 2016.

[38]. Bruckstein AM, Donoho DL, and Elad M, “From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images,” (in English), Siam Rev, vol. 51, no. 1, pp. 34–81, 3 2009.

[39]. Chen SB and Donoho D, “Basis Pursuit,” (in English), Conf Rec Asilomar C,pp. 41–44, 1994.

[40]. Gill PR, Wang A, and Molnar A, “The In-Crowd Algorithm for Fast Basis Pursuit Denoising,” (in English), Ieee T Signal Proces, vol. 59, no. 10, pp. 4595–4605, 10 2011.

[41]. Amelunxen D, Lotz M, McCoy MB, and Tropp JA, “Living on the edge: Phase transitions in convex programs with random data,” arXiv e-prints, Accessed on: March 01, 2013[Online]. Available: https://ui.adsabs.harvard.edu/abs/2013arXiv1303.6672A

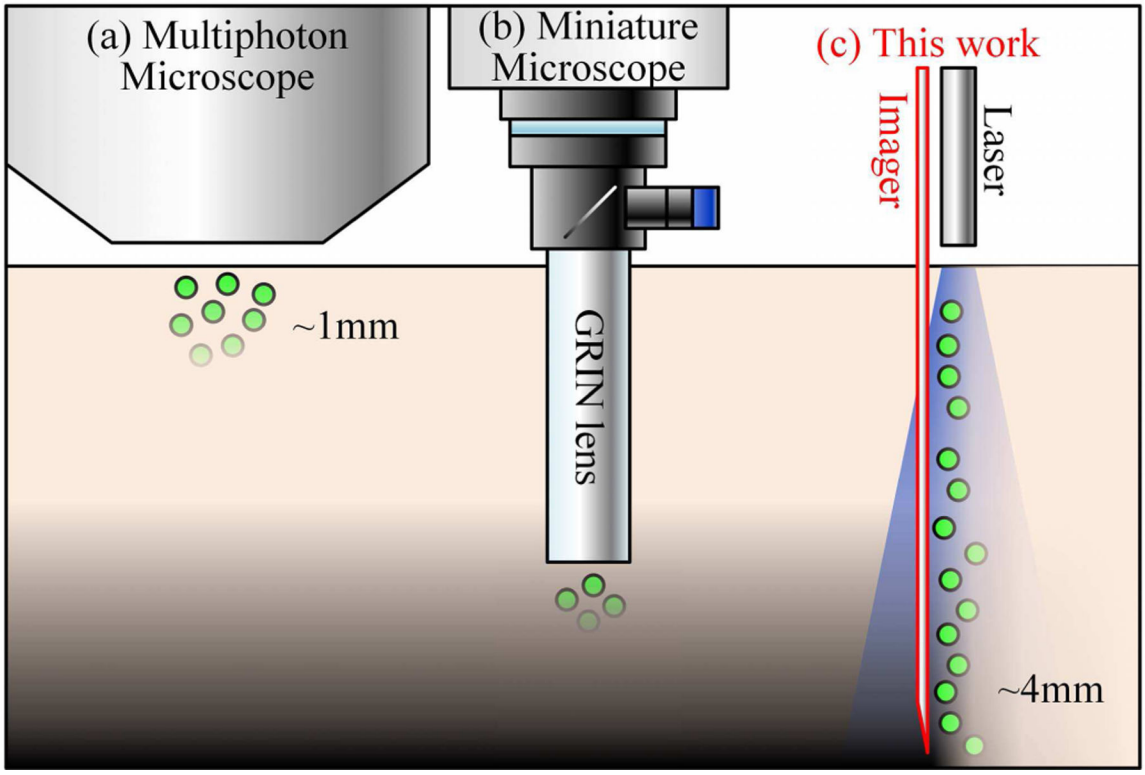

Fig. 1.

Size and imaging depth comparison of (a) multi-photon microscope, (b) implantable microscope with inserted GRIN lens, and (c) imaging probe resolving 4.1 mm-deep in brain tissue with an external excitation laser. Fluorescent emission is collected by 512 SPAD pixels located on two shanks.

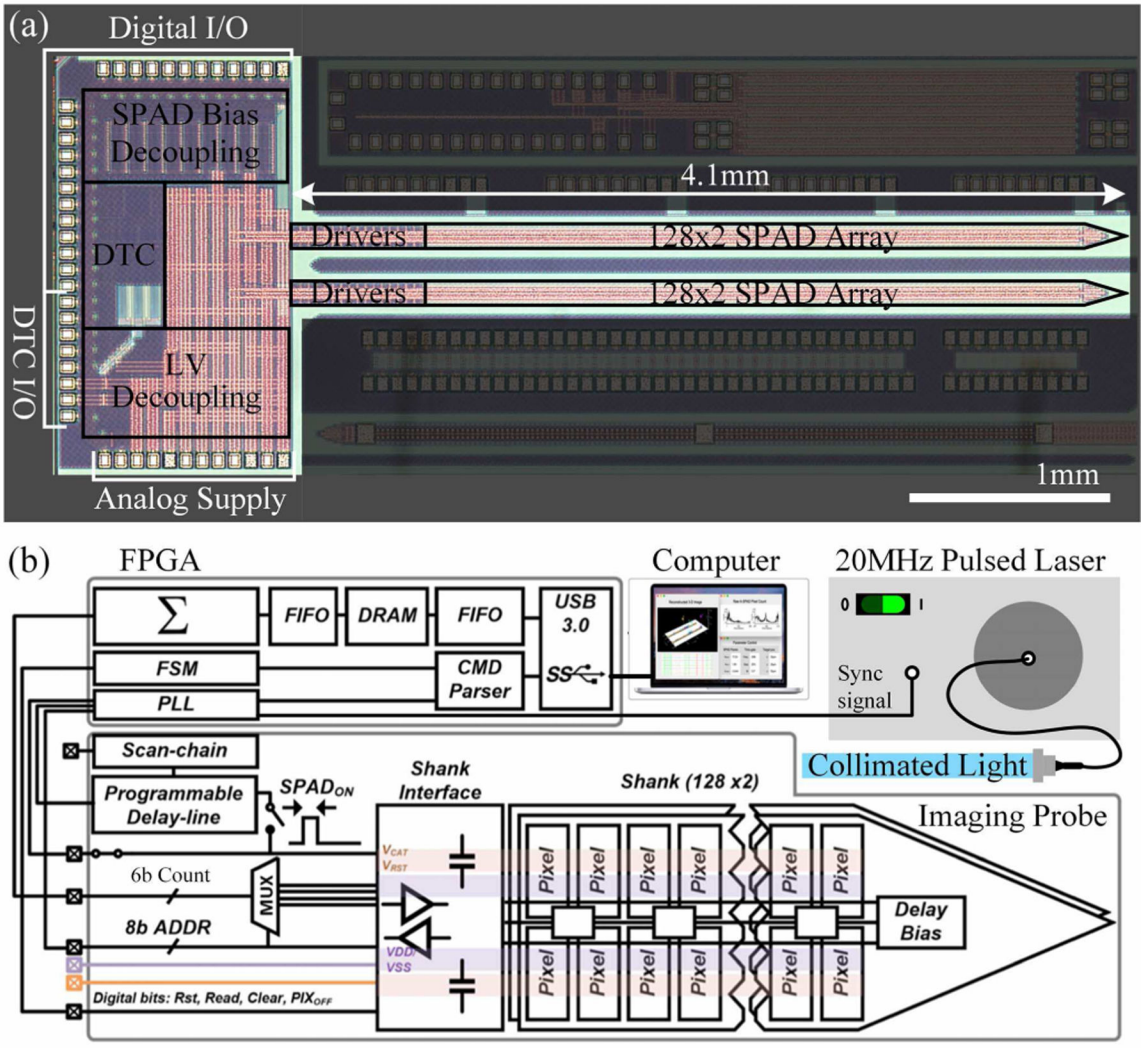

Fig. 2.

(a) Chip micrograph with shank IC highlighted. (b) System architecture consisting of optoelectronic IC, read-out FPGA and computer for image reconstruction.

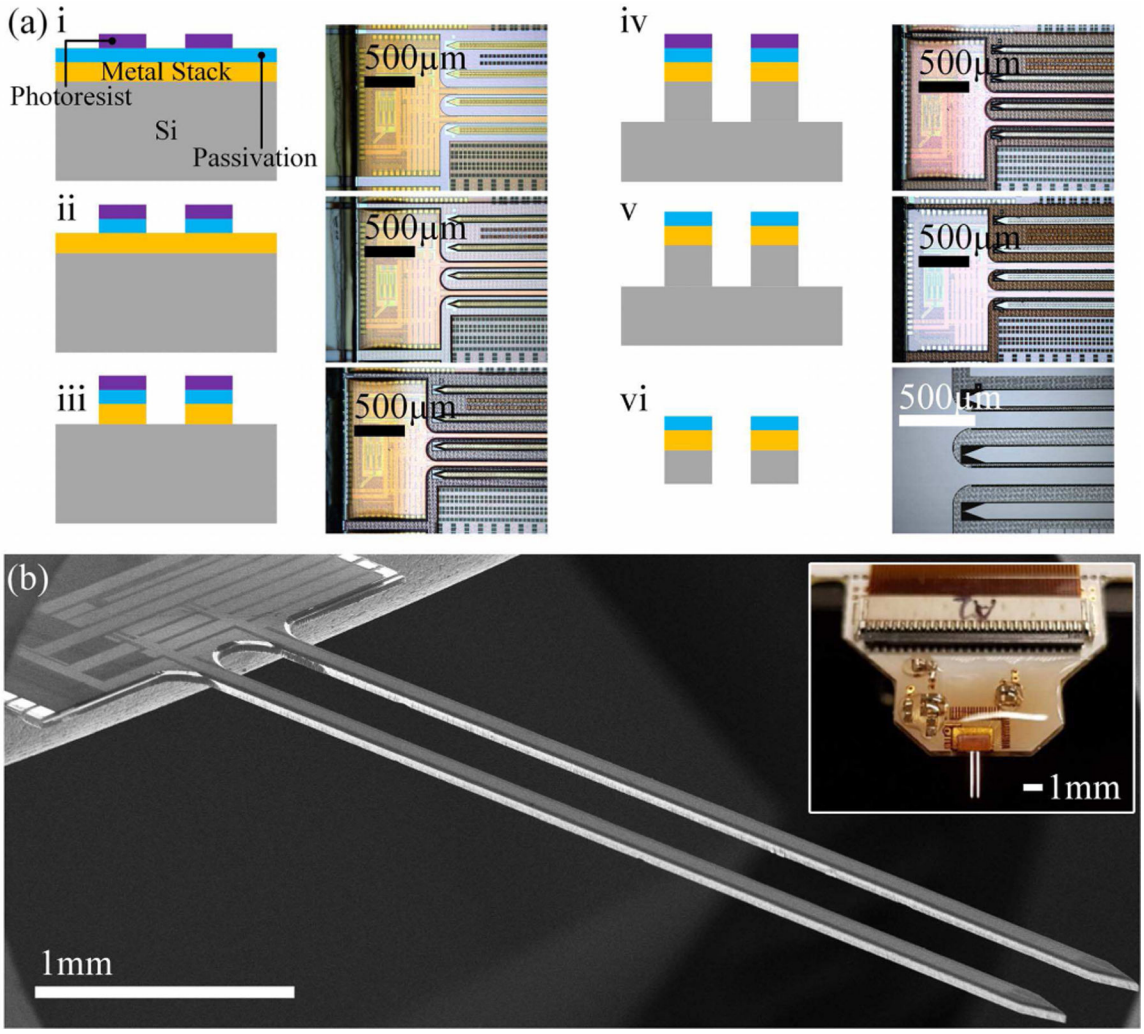

Fig. 3.

(a) Post-processing of CMOS die into implantable imaging probe. (b) Scanning electron microscopy image of post-processed imaging probe with 120 μm width and 40 μm thickness, packaged for insertion (inset).

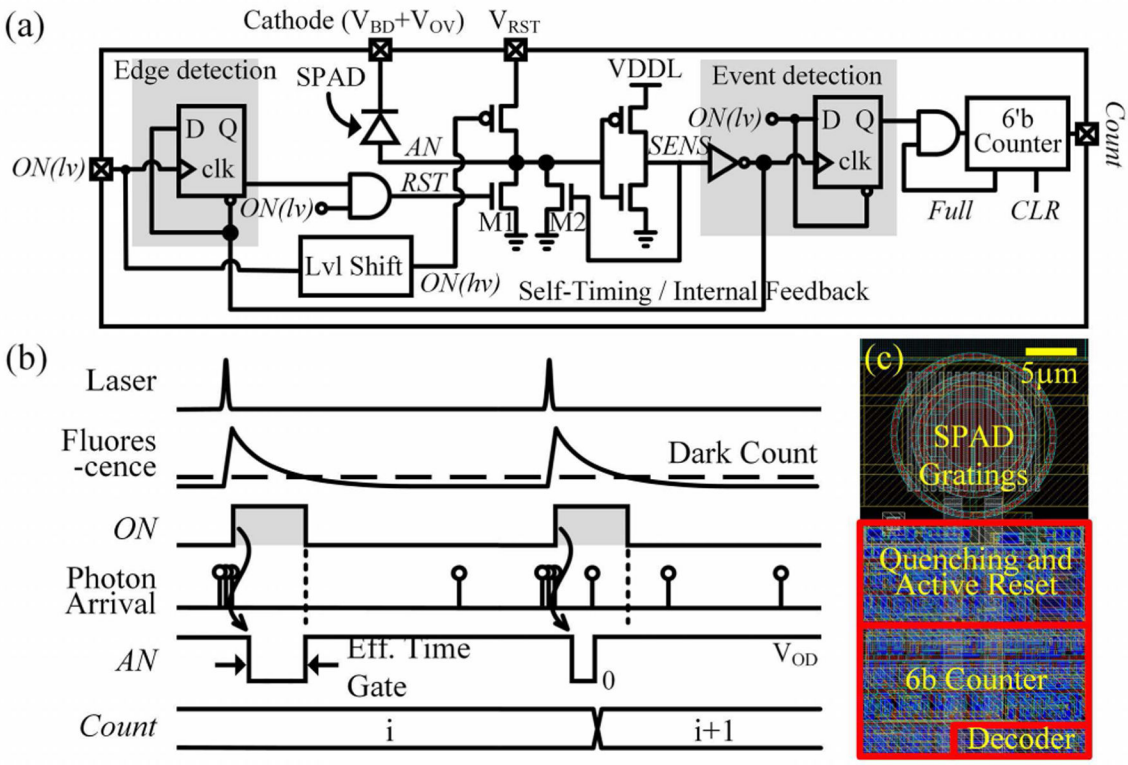

Fig. 4.

(a) Circuit diagram of active-reset quenching circuit. (b) Timing diagram of laser-synced time-gate operation. (c) Layout of single pixel.

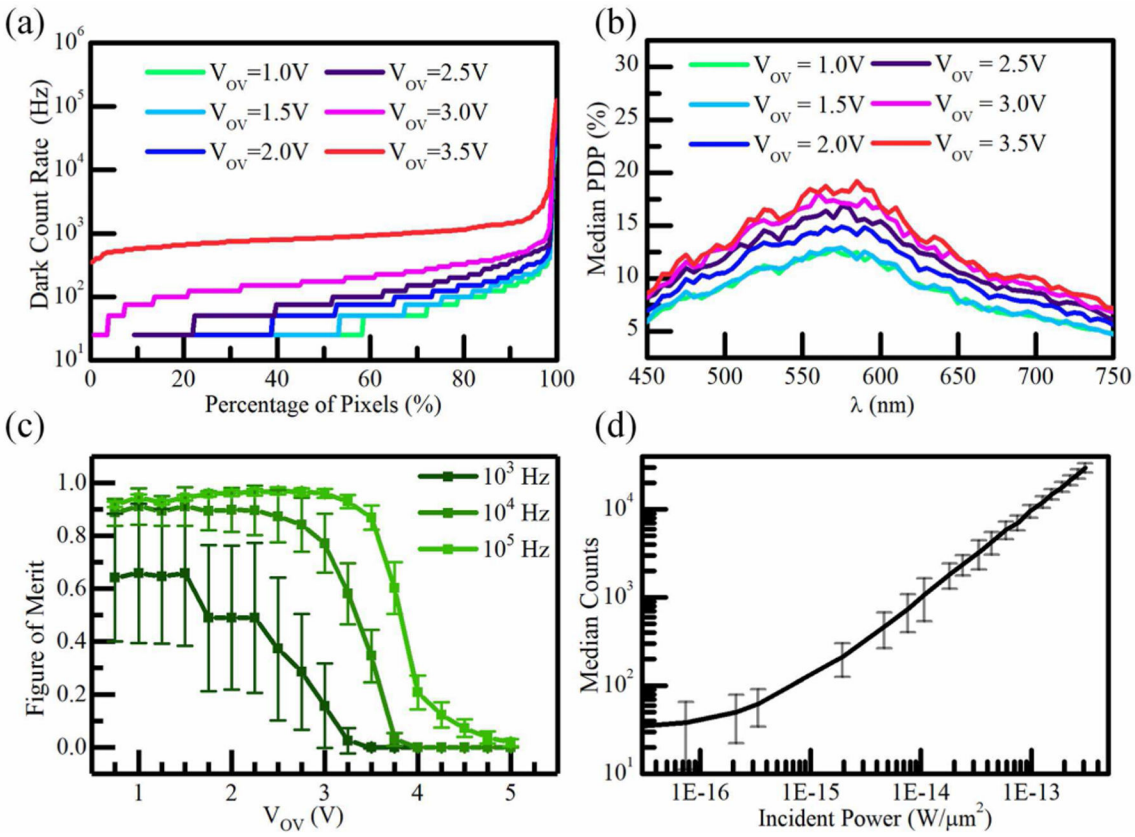

Fig. 5.

(a) Dark count rate of 512 pixels. (b) Median PDP versus incident wavelength for 350 nW isotropic illumination. (c) FoM [12] at three different fluences plotted versus V_{OV} (d) SPAD count linearity versus incident light intensity. Error bars indicate the standard deviation over 512 pixels.

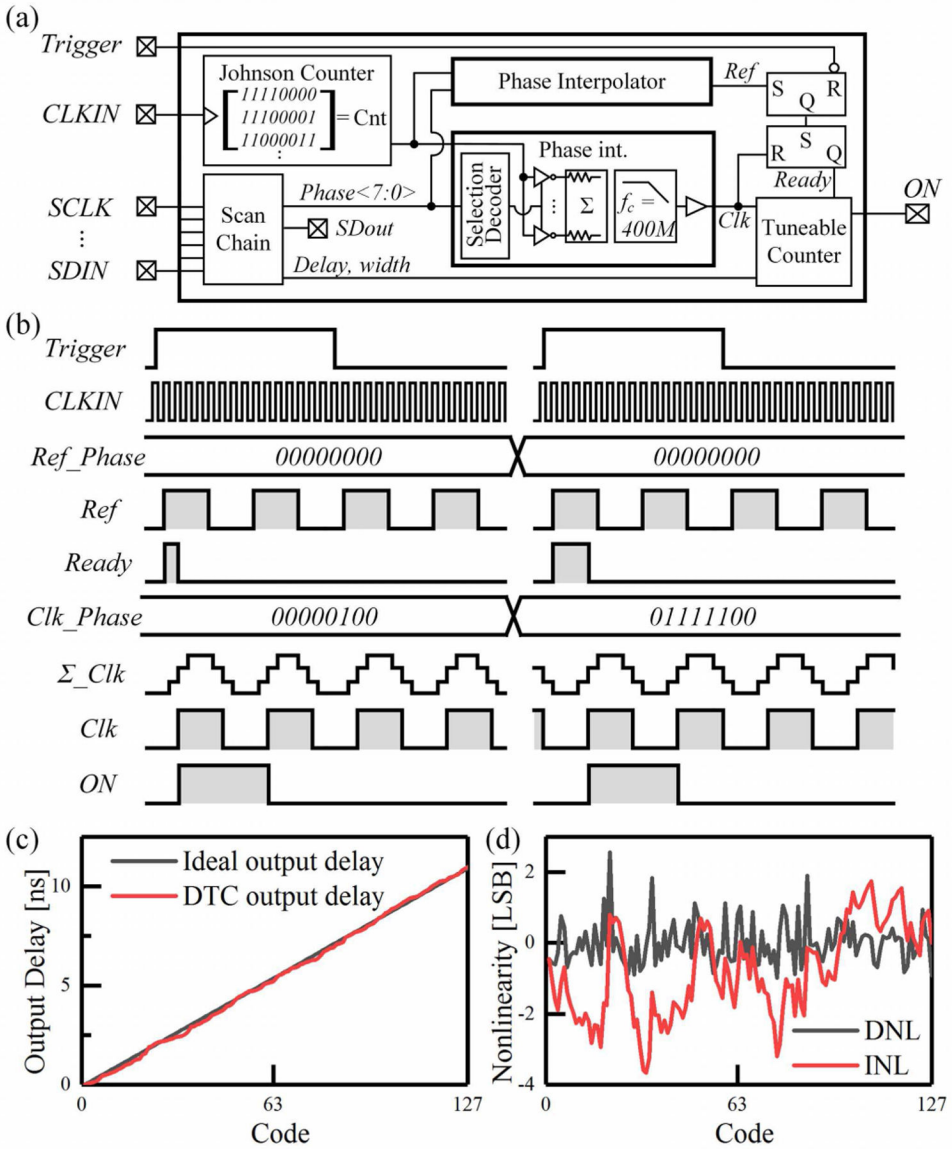

Fig. 6.

(a) Circuit diagram of DTC. (b) Timing diagram with and tuneable counter width=1 and delay=0. (c) Ideal and synthesized *ON* signal delays and (d) conversion nonlinearity.

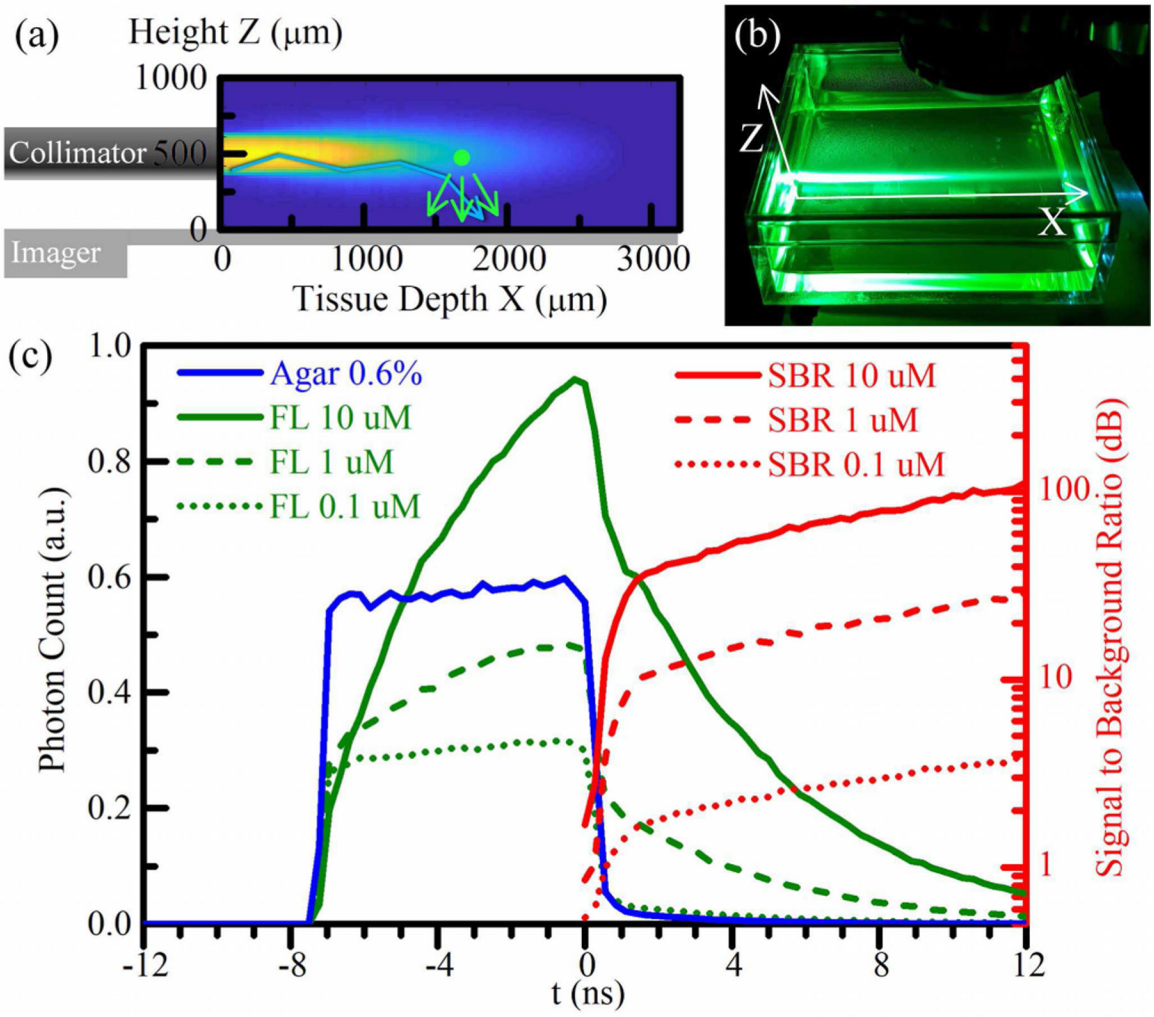

Fig. 7.

(a) Diagram of fluorescence measurement setup showing fluorescent and scattered excitation photons reaching the SPAD array located along the x axis. (b) Photo of 10 μM fluorescein in 0.6% agarose, excited by a collimated 488 nm pulsed laser. (c) Time-gated photon counts synchronized to the laser pulse at t=0 s for pure agarose (blue) and fluorescein in agarose (*FL*, green) plotted against the left axis; respective SBR (Fluorescein:Agarose) is plotted on the right axis (red)

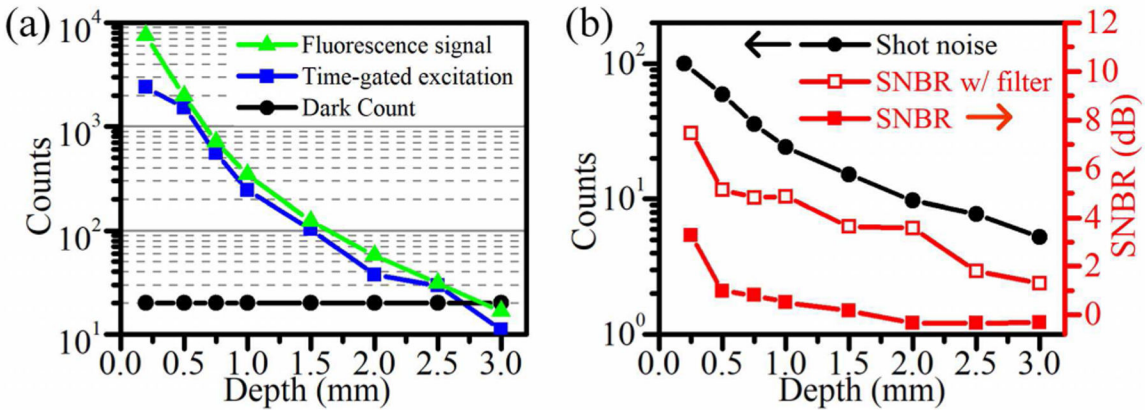

Fig. 8.

(a). Fluorescence emission of EGFP-expressing somata in model tissue compared with background and dark count as a function of tissue depth. (b) Combined photon shot noise and dark count shot noise, plotted with SNBR as a function of tissue depth.

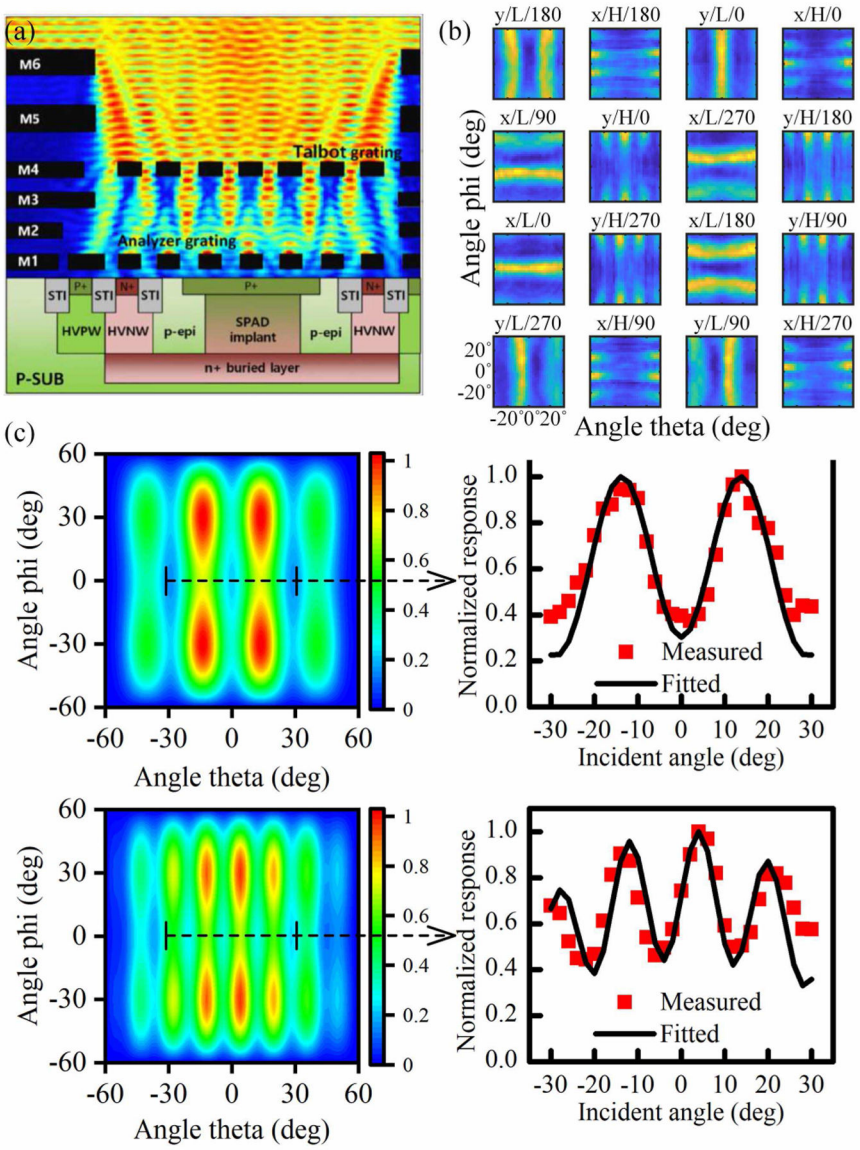

Fig. 9.

(a) Illustration of an angular sensitive SPAD implemented in CMOS. (b) Measured angular modulation of sixteen different grating structures in direction (x, y), frequency (High, Low), and quadrature phase (0°, 90°, 180°, 270°), for 520 nm illumination (c) Simulated angular responses of grating combinations y/L/180° (top left) and y/H/90° (bottom left), with respective φ =0° cross sections measured and parametrically fitted (right).

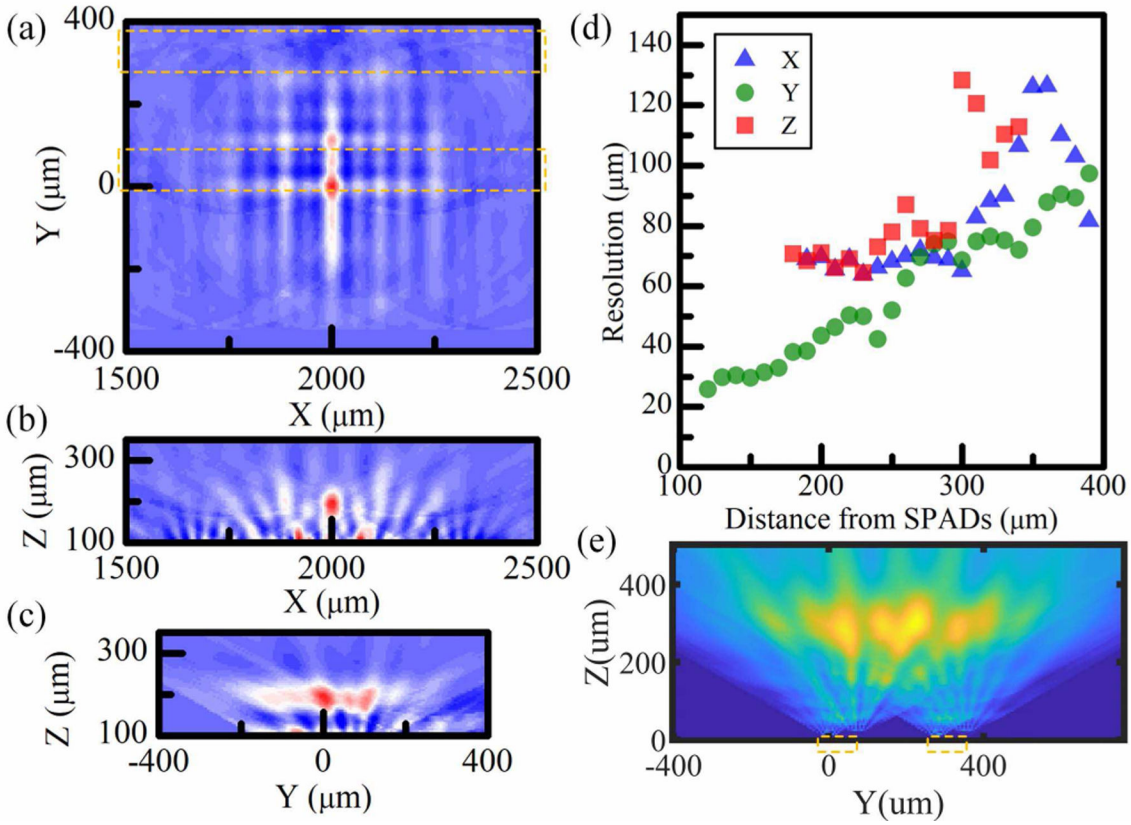

Fig. 10.

PSF of a single point source located at coordinates [2000 μm, 0 μm, 200 μm] portrayed as (a) xy-, (b) xz-, (c) yz-plane cross sections of a 3D pseudoinverse backprojection; voxels are 5 μm across all dimensions; (d) PSF FWHM in x, y, z with respect to with increasing distance of point source from imager; (e) xz-plane cross section of imager FoV. Dotted yellow lines in (a) and (e) mark the location of the imager.

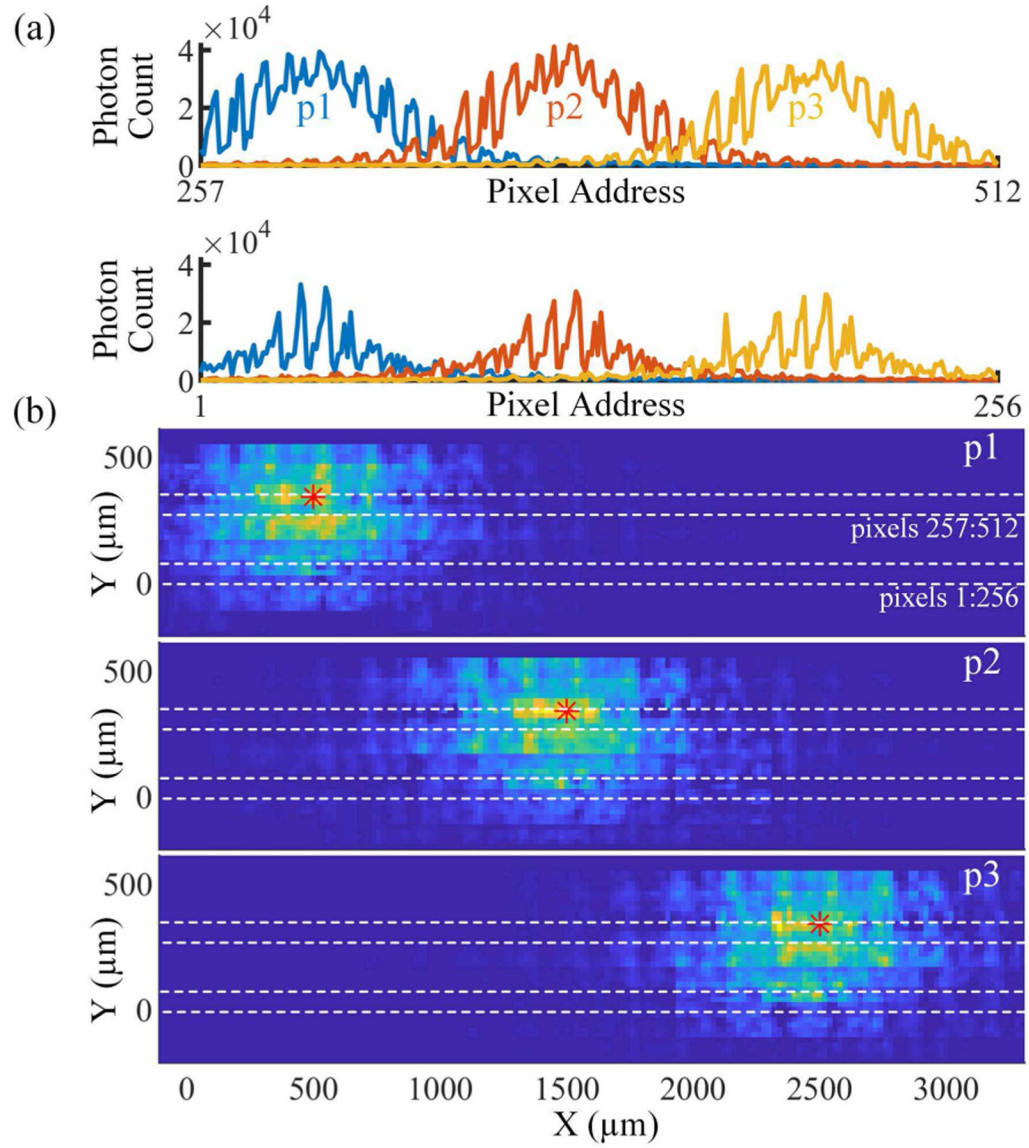

Fig. 11.

A monochromatic light source hovering 200 μm above the imager was placed at x locations of 500 μm (p1), 1500 μm (p2), and 2500 μm (p3). (a) Raw 512-pixel counts along two shanks and (b) pseudoinverse backprojection in plane z=400 μm with source locations (*) and probe (–) overlaid.

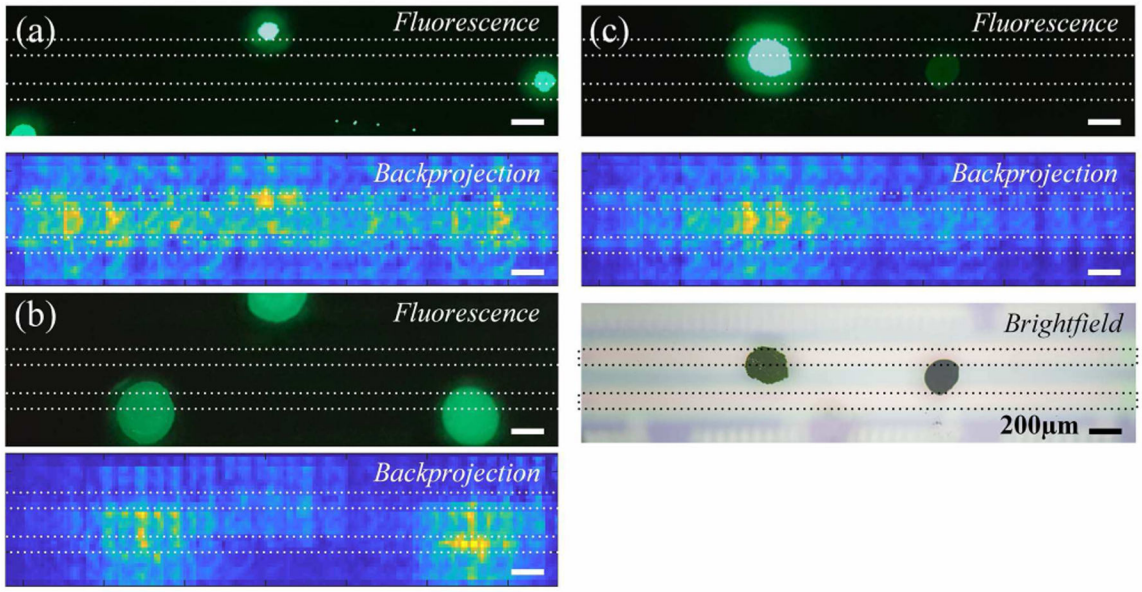

Fig. 12.

(a) 10 μm-diameter microsphere clusters deposited on cover glass placed 400 μm above imager (dotted line overlay); epifluorescence microscopy image(top) and pseudoinverse backprojection(bottom). (b) Same with larger clusters. (c) Two sphere types, fluorescent (left) and scattering (right), showing selective imaging of fluorescent target; top-fluorescent, middle-brightfield, bottom-backprojection. All scale bars are 200 μm.

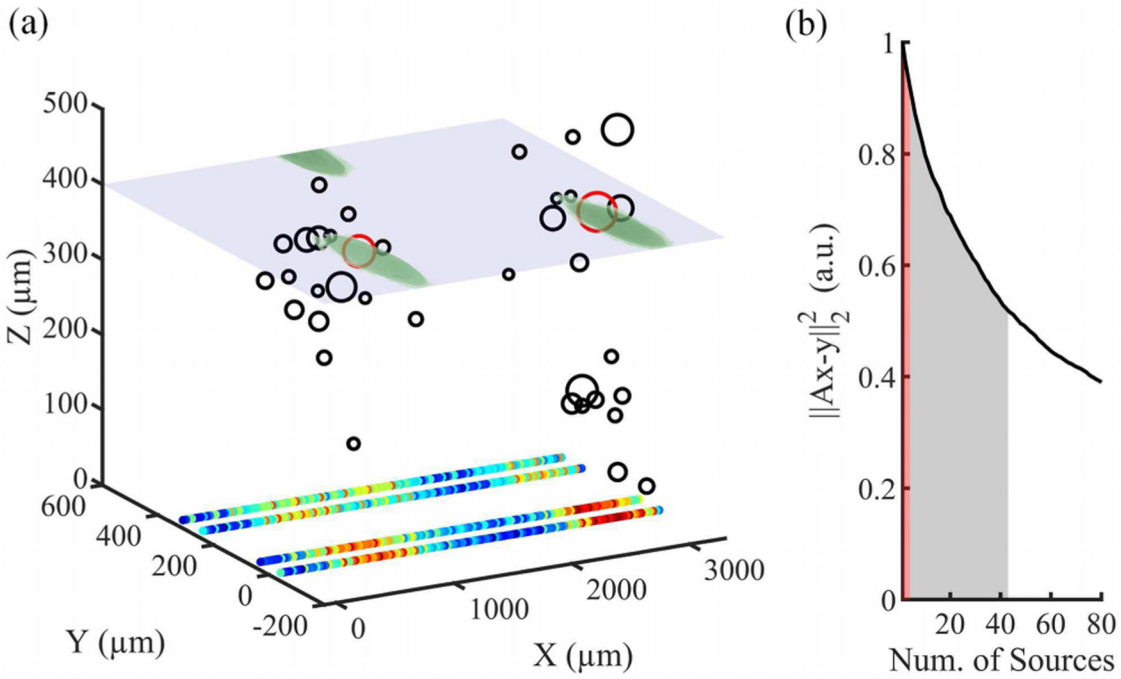

Fig. 13.

(a) 3D reconstruction of microsphere clusters in Fig. 12b; two most significant solutions are marked in red circles and the following 40 in black, with diameters depicting amplitude estimate; the overlaid ground truth image indicates the location of microspheres 400 μm above the chip. (b) L2-norm of residual error decreasing with added sources, red and gray corresponding to first 2 and next 40 voxels in order of highest amplitude estimate.

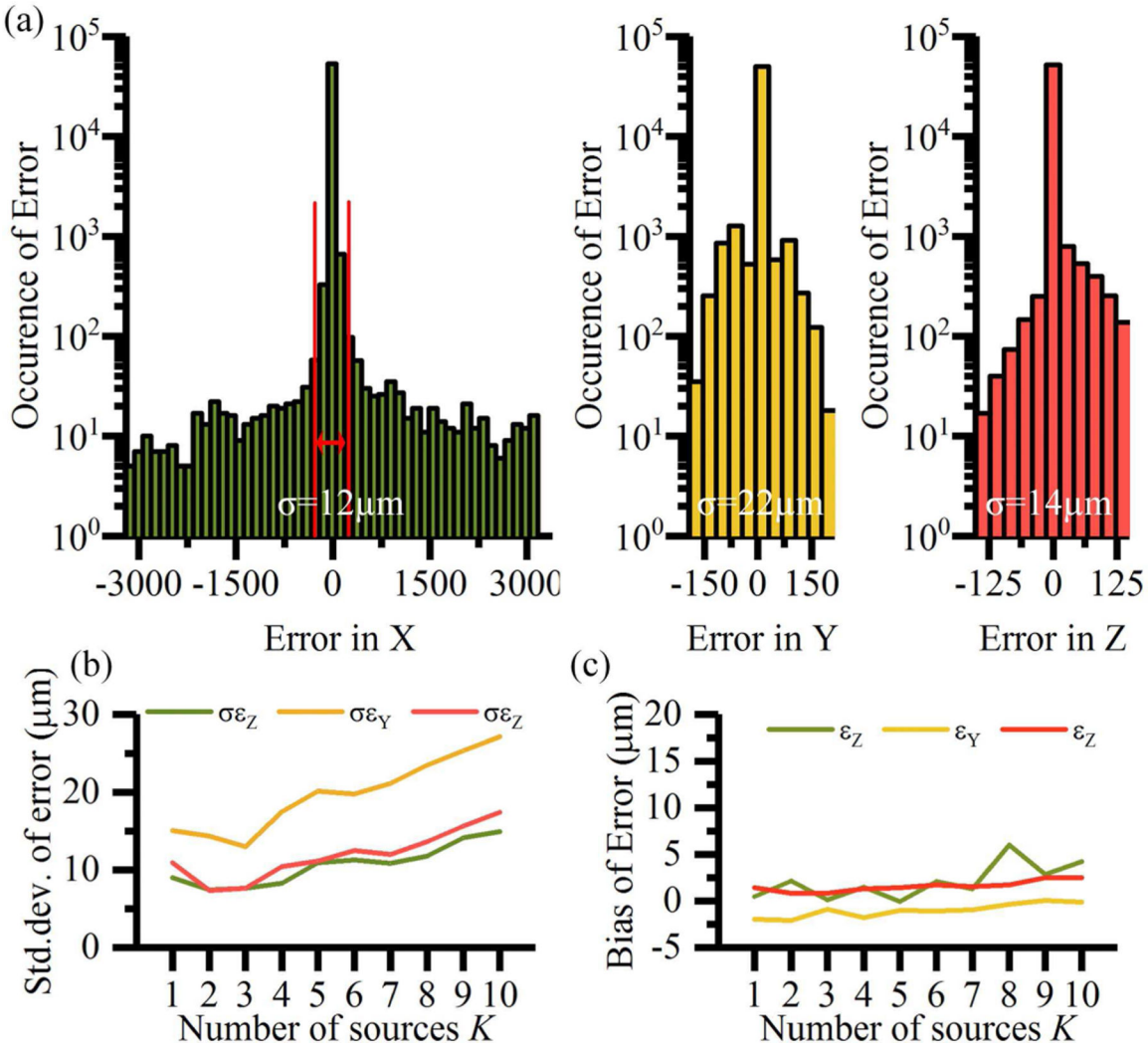

Fig. 14.

(a) Distribution of localization error in spatial coordinates. (b) Standard deviation when accounted for 2% false positive / false negative outliers. (c) Bias of localization error.

TABLE I

COMPARISON TABLE WITH PRIOR SPAD IMAGERS

	Schwartz [14]	Lee [15]	Perenzoni [16]	Field [17]	This work
Technology	0.35 µm	0.18 µm	0.35 µm	0.13 µm	0.13 µm
Pixel Pitch	40 µm	35 µm	50 µm	48 µm	25.3 µm
Pixel Count	64 × 64	72 × 60	160 × 120	64 × 64	4 × 128
Power	-	83.8 mW	157 mW	26 W	6.24 mW
Temporal Resolution	350 ps	71 ps	194 ps	62.5 ps	137 ps
DCR Median	1059 Hz @ 2.7 V	400 Hz @ 1.2 V	580 Hz @ 3 V	544 Hz @ 1.5 V	40 Hz @ 1 3 V
Max PDP	4.8% @ 2.7 V	2.7% @ 1.2 V	-	30% @ 1.5 V	12.4% @ 1 V
Fill Factor	14%	9.6%	21%	0.77%	6.3%
Frame Rate	1144 fps	-	486 fps	100 fps	51 kfps

*PDP and DCR are reported at respective excess bias voltages. PDP is reported as the maximum value over all wavelengths.