A Sensitive Dynamic and Active Pixel Vision Sensor for Color or Neural Imaging Applications

Diederik Paul Moeys, *Student Member, IEEE*, Federico Corradi, *Member, IEEE*, Chenghan Li, Simeon A. Bamford, Luca Longinotti, Fabian F. Voigt, Stewart Berry, Gemma Taverni, Fritjof Helmchen, and Tobi Delbruck, *Fellow, IEEE*

***Abstract*—Applications requiring detection of small visual contrast require high sensitivity. Event cameras can provide higher dynamic range (DR) and reduce data rate and latency, but most existing event cameras have limited sensitivity. This paper presents the results of a 180-nm Towerjazz CIS process vision sensor called SDAVIS192. It outputs temporal contrast dynamic vision sensor (DVS) events and conventional active pixel sensor frames. The SDAVIS192 improves on previous DAVIS sensors with higher sensitivity for temporal contrast. The temporal contrast thresholds can be set down to 1% for negative changes in logarithmic intensity (OFF events) and down to 3.5% for positive changes (ON events). The achievement is possible through the adoption of an in-pixel preamplification stage. This preamplifier reduces the effective intrascene DR of the sensor (70 dB for OFF and 50 dB for ON), but an automated operating region control allows up to at least 110-dB DR for OFF events. A second contribution of this paper is the development of characterization methodology for measuring DVS event detection thresholds by incorporating a measure of signal-to-noise ratio (SNR). At average SNR of 30 dB, the DVS temporal contrast threshold fixed pattern noise is measured to be 0.3%–0.8% temporal contrast. Results comparing monochrome and RGBW color filter array DVS events are presented. The higher sensitivity of SDAVIS192 make this sensor potentially useful for calcium imaging, as shown in a recording from cultured neurons expressing calcium sensitive green fluorescent protein GCaMP6f.**

***Index Terms*—CMOS image sensors, biomedical optical imaging, image color analysis, neuromorphics, machine vision, discrete-event systems.**

I. Introduction

THE interest in bio-inspired architectures has led to the development of event-based neuromorphic systems to achieve faster and lower power sensing and computing by using the brain's sparse, asynchronous digital communication architecture [1]–[3]. Neuromorphic sensors have been developed, in particular, the Dynamic Vision Sensor (DVS) [4]. Each DVS pixel reports if there has been a positive or negative brightness (log intensity) change since the last event. ON and OFF event coordinates are thus asynchronously generated and communicated off chip through the Address Event Representation (AER) protocol [3], [5]. This sparse information about scene brightness changes allows lower data-rate and latency, which in turn means lower power consumption and higher speeds.

To enable compatibility with conventional frame-based cameras the Dynamic and Active Pixel Vision Sensor (DAVIS) was developed in [6], combining DVS and Active Pixel Sensor (APS) readout. A color DAVIS, consisting of single DAVIS pixels surrounded by three 5-transistor global-shutter pinned photodiode APS pixels was also built in [7]. Another method of combining light intensity readout with DVS is the ATIS [8]; it offers high Dynamic Range (DR) intensity but large pixel size.

To satisfy the need for more sensitive dynamic vision sensors to serve application areas such as bolometry, fluorescence microscopy and fluid dynamics, pixel variants were designed to achieve higher contrast sensitivity by increasing photoreceptor front-end gain. These areas require sensitivity to contrasts below 10%. Microbolometers require sensitivities of around 1% to be sensitive to 1°K temperature changes, because the temperature coefficients of the micromachined devices are roughly around 1.5%/°K [9].

Of particular interest within fluorescence microscopy is the recording of functional signals in living cells and tissues, for example by imaging calcium or voltage changes in neurons using fluorescent indicators. The newest genetically encoded calcium indicators have changes of fluorescence over the baseline ΔF/F of 20% for a single action potential in cultured neurons and rise times of 50–150 ms [10].

Sensing low contrast is difficult because noise in the signal can obscure the small relative variation of the signal and transistor mismatch makes it difficult to set low thresholds. The fundamental shot noise variance in a collected signal of N photoelectrons is $\sigma^2 = N$, and reliably sensing a fractional deviation $C = \Delta N/N$ (single pixel contrast) requires a signal that is a multiple of the noise. If we take m standard deviations ($m\sigma$) as our required signal to noise ratio (SNR), then we can compute

Manuscript received March 31, 2017; revised July 23, 2017; accepted September 30, 2017. Date of publication November 3, 2017; date of current version January 26, 2018. This work was supported in part by the European Commission Project VISUALISE (FP7-ICT-600954), SeeBetter (FP7-ICT-270324), in part by the Swiss National Science Foundation (31003A_149858; F.H.), in part by the U.S. NIH BRAIN Initiative (1U01NS090475-01, F.H.), and in part by the Samsung Electronics Corporation. This paper was recommended by Associate Editor A. Bermak. *(Corresponding author: Tobi Delbruck.)*

D. P. Moeys, C. Li, G. Taverni, and T. Delbruck are with the Institute of Neuroinformatics, University of Zürich and ETH Zürich, Zürich 8057, Switzerland (e-mail: diederikmoeys@live.com; li3.cheng2han4@gmail.com; gemmataverni@gmail.com; tobi@ini.uzh.ch).

F. Corradi, S. A. Bamford, and L. Longinotti are with iniLabs GmbH, Zürich 8057, Switzerland (e-mail: federico.corradi@gmail.com; simbamford@gmail.com; luca.longinotti@inilabs.com).

F. F. Voigt, S. Berry, and F. Helmchen are with the Brain Research Institute, University of Zürich, Zürich 8006, Switzerland (e-mail: voigt@hifo.uzh.ch; berry@hifo.uzh.ch; helmchen@hifo.uzh.ch).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBCAS.2017.2759783

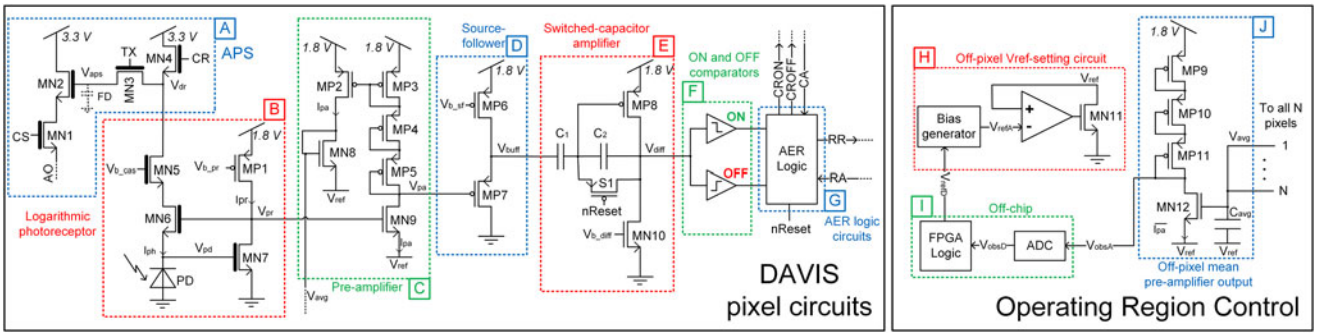


Fig. 1. SDAVIS192 pixel, thick gate lines denote 3.3 V transistors. A: APS, B: logarithmic photoreceptor, C: pre-amplifier stage, D: p-type source-follower, E: difference amplifier, F: ON and OFF comparators, G: in-pixel AER logic circuits, H: off-pixel circuit for setting $\mathrm{V_{ref}}$ and bias generator, I: off-chip FPGA state machine and ADC, **J**: off-pixel mean pre-amplifier output circuit.

the required N [11]:

$$\mathrm{C} = \frac{\Delta N}{N} > m\sigma = \frac{m}{\sqrt{N}}$$

$$N > \frac{m^2}{C^2} \tag{1}$$

E.g. if $\mathrm{C} = 0.01$ and $m = 3$, then $N = 9 \times 10^4$. Thus a 3-sigma detection of a 1% contrast requires collecting 90k e$^-$, which exceeds the typical 10k e$^-$ Full Well Capacity (FWC) of conventional CMOS image sensors by a factor of ten [9].

In the previously mentioned application areas, the collected light is often the result of a secondary process such as fluorescence, where the primary exciting light is blocked by a filter, and the emission is a small fraction of the excitation, typically between $10^{-5} - 10^{-6}$ [12]. The level of excitation illumination is limited by the small molecular cross sections for absorption [13] and by the low concentrations of fluorophores commonly used. The above calculation also ignores other noise sources such as 1/f noise in the readout. This discussion assumes single pixel reliable detection of given contrast, but area can trade off for contrast, as explained in [11]. Detecting a smaller contrast trades off linearly with increasing the linear dimension of the feature to be detected.

Past DVS developments aimed towards higher sensitivity include a bolometer DVS [14] that used two switched capacitor amplifiers in series to achieve gain with modest capacitor ratios; the color-change DVS in [15] that used a sample and hold with capacitive amplification in the front-end stage; the sensitive DVS in [16] that used a stacked diode voltage amplifier; and the delta-modulator DVS in [17] that used a programmable-gain operational amplifier. We adopted the practical approach pioneered in [16]. Although it sacrifices intrascene DR, it leads to a simple and compact design. However, since the circuit in [16] consumes the photocurrent, it is incompatible with the DAVIS sensor since the photocurrent is also needed for integration by the APS circuit. To overcome this problem, the SDAVIS192 was designed and fabricated. SDAVIS192 simplifies the gain stage of [16] and moves the operating region control outside the IC, where it can be controlled more flexibly by a designed control policy, rather than a fixed analog feedback loop.

One of the main additional contributions of this paper is to improve measurement of neuromorphic vision sensor specifications towards establishment of standards. Specifically, this paper reports our methodology for measuring the temporal contrast threshold and suggests a measure of signal to noise ratio (SNR) for the temporal contrast. We use this method to measure the temporal contrast threshold of SDAVIS192. The rest of this paper is organized as follows: Section II describes the design of SDAVIS192. Section III reviews characterization methods in the literature and presents setup and protocols used for the characterization of SDAVIS192 and the measurement results. Finally, Section IV shows an application of SDAVIS192 in the field of neural imaging.

II. Design

This section describes the design of SDAVIS192, including the operating point control and IC layout.

A. Concept and Preamplifier Design

The schematic of the SDAVIS192 pixel is shown in Fig. 1. The pixel improves on the original DAVIS [6], which in turns improves the original DVS pixel (to detect logarithmic brightness changes over time) of [4] by adding the well-known APS part to obtain signal intensity frames (box **A** in Fig. 1). The APS readout shares the same photodiode of the DVS circuit and it is isolated from it by the cascode transistor MN5. The APS readout is reset when MN4 is activated through the Column Reset (CR) signal. The reset charge, and alternatively, the illumination-dependent integrated charge, are integrated on the Floating Diffusion (FD) node Vaps when the transmission gate MN3 has been connected (integration starts) and disconnected (integration end), depending on the signal TX. Once the pixel's column is selected through the Column Select (CS) signal through the transistor MN1, the integrated FD value is read through the source-follower pair composed by MN2 and a column-parallel source follower bias transistor.

The DVS circuitry of SDAVIS consists of the logarithmic photoreceptor with output $\mathrm{V_{pr}}$ (box **B**) which compresses the input photo-signal, a preamplification stage with output $\mathrm{V_{pa}}$ (box **C**) to boost the signal, a p-type source-follower buffer to

avoid signal coupling with output V_{buff} (box **D**), a switched capacitor amplifier with output V_{diff} (box **E**), and two comparators to generate ON and OFF events (box **F**). Future designs could probably eliminate the source-follower stage since 2 columns of test pixels indicate that coupling back to the high impedance photoreceptor is already minimized by the intervening gain stage. Then AER communication circuits (box **G**) do a four-phase handshake by generating the pixels' Row Request (RR), Column Request OFF (CROFF) and Column Request ON (CRON) signals and receive the Row Acknowledge (RA) and Column Acknowledge (CA) signals.

The main improvement over previous DAVIS sensors is the introduction of a preamplification stage (box **C**), adapted from [16] to the DAVIS pixel circuit. It consists of a low-mismatch subthreshold voltage amplifier. While in [16] the amplifier is composed by a load of 4 series diode-connected nFET transistors driven by a pFET transistor, the proposed implementation uses only 3 diode-connected pFETs (transistors MP3-5 driven by the nFET MN9) to match the 1.8 V supply. The preamplifier of [16] was coupled to the photoreceptor through a source-driven active current-mirror [18] which is incompatible with the DAVIS APS readout. The SDAVIS192 unifies the preamplifier design of [16] with the front-end photoreceptor of [4], [6] to achieve a DAVIS pixel with higher front-end gain. The preamplifier, whose input is V_{pr} and whose output is V_{pa}, works on the principle of a subthreshold common source amplifier with diode-connected load, as explained in [18]. A pFET active-load would have too much uncontrolled gain, together with an additional bias line. A very short pFET to get low enough gain would have a very process-dependent gain dependent on the Early voltage. The minimum size diode-connected transistors occupy less than 5% of the total pixel area. The gain A_1 of the preamplifier can be derived by equating the current in MP3-5 and MN9 [16]:

$$A_1 = -\frac{\kappa_n}{\kappa_p}\left(1+\frac{1}{\kappa_p}+\frac{1}{{\kappa_p}^2}\right) \tag{2}$$

where κ_p and κ_n are the back gate coefficients of p and n FETs, respectively. For $\kappa_p = 0.7$ and $\kappa_n = 0.8$, $A_1 = -5.1$. A detailed analysis on the low gain mismatch of this circuit, which mainly depends on body effect mismatch, is reported in [16]. The relationship between the change in logarithmic intensity $\Delta\ln(I)$ and V_{diff} is given by:

$$\begin{aligned}\Delta V_{diff} &= -A_2\Delta V_{buff} = -A_2\kappa_p\Delta V_{pa}\\ &= -A_1A_2\kappa_p\Delta V_{pr} = \frac{-U_T A_1 A_2\kappa_p}{\kappa_n}\Delta\ln(I)\end{aligned} \tag{3}$$

where A_2 is the closed-loop gain of 30 of the difference amplifier corresponding to the capacitive ratio C_1/C_2 (its simulated open-loop gain is 251), and U_T the thermal voltage.

A detailed analysis on noise was presented in [17]. The input-referred noise introduced by the pre-amplifier and all other cascaded circuits is negligible compared to the amount introduced in the photoreceptor front-end. Furthermore, because of the relatively small size of the transistors, flicker noise from the transistors MN3-6 in the photodiode current branch dominates (as verified in simulation).

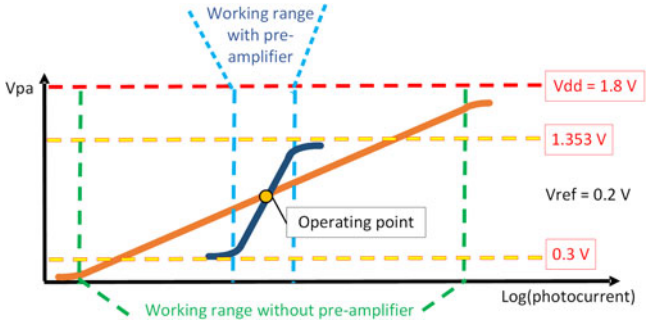


Fig. 2. Dynamic range reduction with increasing gain (conceptual plot).

B. Operating Region Control

The additional gain introduced by A_2 reduces the DR of the pixel and therefore leads to possible clipping. In high illumination conditions, V_{pa} saturates at V_{ref} plus the saturation voltage of MN9 ($4U_T$). For low illuminations, V_{pa} will begin to saturate at the saturation voltage of MP3 plus the two drain-to-source voltages of MP4 and MP5, which are larger because of bulk effect. Fig. 2 shows the range of V_{pa}:

$$V_{ref} + 4U_T < V_{pa} < V_{dd} - 4U_T\left(1+\frac{1}{\kappa_p}+\frac{1}{{\kappa_p}^2}\right) \tag{4}$$

For $V_{ref} = 0.2\,V$ and $\kappa_p = 0.7$, the possible swing is 1.053 V.

The operating points of the preamplifiers can be globally adjusted through the V_{ref} common source. V_{ref} control is based on [16], which faced the same saturation problem with the preamplifier. In [16], the operating point was adapted by changing the gate voltage of the preamplifier rather than the source, since the DVS circuit is coupled to the photoreceptor by a source-driven active current mirror. The analog adaptation circuit continuously adapted the reference gate voltage depending on the off-pixel recreated average photocurrent. A digital control approach is amenable to software design, and was chosen to replace the analog feedback of [16] that is prone to oscillations (personal communication from Linares-Barranco and previous experience with analog feedback [17]). This concept is illustrated in Fig. 3. A replica-biasing circuit external to the pixel was therefore designed to set the source reference bias voltage V_{ref} (hence the operating point) of the preamplifier of every pixel depending on the average scene illuminance.

The aim of the adaptation replica biasing circuit is to check if the preamplifier's outputs are saturating at either one of their limits, and if they do so, to digitally control V_{ref} to bring back the signals to their correct operating range. As seen in Fig. 1, the adaptation circuit (box **H**, **I**, **J**) makes a copy of the preamplifier's current, by mirroring it with the pFET MP2 and then averaging it onto the diode-connected nFET MN8 in parallel with the shared 8.8 pF C_{avg} pMOS capacitor (box **J**) to smooth this value (V_{avg}) with a time constant dependent on the average preamplifier current $\overline{I_{pa}}$ of the N pixels. V_{avg} is then fed into the preamplifier replica (MP9-11 and MN12) to re-create the mean

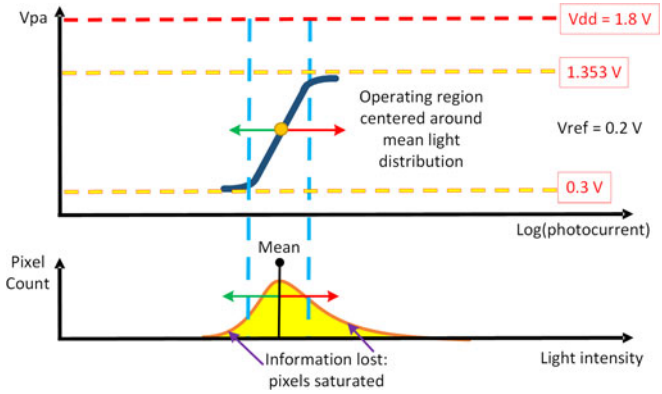


Fig. 3. Operating region shift calculated from the mean intensity distribution.

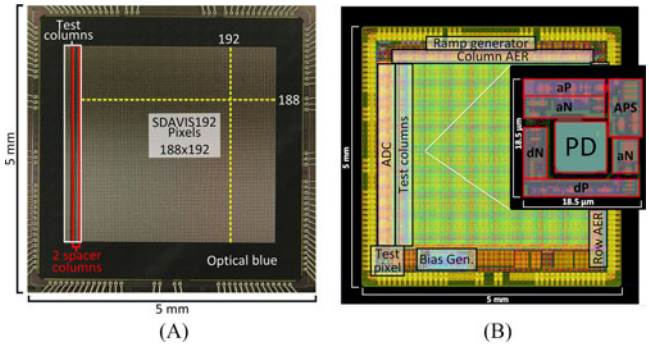


Fig. 4. (A) SDAVIS192 microphotograph. (B) SDAVIS192 IC layouts (showing active, poly and M1). In single pixel: PD denotes the photodiode and aP/aN and dN/dP denote analog and digital n or pFETs respectively.

pixel condition. V_{avg} is given by:

$$V_{avg} = \frac{U_T}{k_n}\left(\frac{v_{ref}}{U_T} + \ln\left(\frac{\overline{I_{pa}}}{Io}\right)\right) = \frac{u_T}{k_n}\left(\frac{v_{ref}}{U_T} + \ln\left(\frac{\overline{\sum_{n:1}^{N} I_{pa_n}}}{IoN}\right)\right)$$
$$= \frac{1}{k_n}\left(v_{ref} + U_T \ln\left(\left(\sum_{n:1}^{N} e^{\ln(I_n) - v_{ref}/U_T}\right)\right)/N\right)$$
$$= \frac{U_T}{k_n}\ln\left(\sum_{n:1}^{N} I_n/N\right) \quad (5)$$

where I_0 is the subthreshold off current, V_{pa_n} is the n'th preamplifier output voltage and I_n is the n'th photocurrent. V_{avg} is then roughly proportional to the logarithm of the photocurrent. V_{obsA} is then the average of the array's V_{pa}.

V_{obsA} is connected to an off-chip 10-bit Analog-to-Digital Converter (ADC), whose output V_{obsD} connects to the FPGA. The latter, according to a Look-Up table, chooses the appropriate action to bring back the preamplifier to its correct operating range. This in practice means sending out the desired V_{refD}, which is then in turn converted into an analog value V_{refA} by the on-chip coarse-fine bias generator of [19] (also used to provide all other biases). The ADC sample rate is adjustable up to 10 MHz. A row-parallel distributed low-dropout voltage regulator drives V_{refA} to the array V_{ref}, by sinking as much current as needed through MN11 [19].

C. Layout

The pixel size is $18.5 \times 18.5\ \mu m^2$ and the transistor count is 47. Digital and analog signals were carefully separated at the pixel level in order to avoid the crosstalk between APS and DVS that also affected [6]. Where possible, metal layers connected to the analog supply rails isolate digital and analog signals. Unfortunately, the weak ground connection of the grounded metal sheet shielding the CR signal from the other analog circuits causes ground fluctuations at every reset of the APS array (CR high), creating flashes of DVS events. This makes DVS and APS unusable simultaneously.

The photodiode has a 21.2% Fill Factor (FF), and is surrounded by a shallow active guard ring. A deep p-implant under the n-well prevents the n-well from stealing photo-generated carriers from the photodiode. A Buried Photodiode (BPD) is used. In a BPD, the junction is buried beneath the silicon surface by a shallow p-type implant that reduces SiO_2 interface trap-states and therefore the dark current of the photodiode. An ohmic n-type contact penetrates the p-type implant. Salicidation was blocked on the source and drain of the APS structure to further reduce dark current. The reset transistor S1 of the DVS is also non-salicided. The redesign of the pixel allowed the resizing of C_1, to increase the gain of the difference amplifier from 20 to 30 with respect to [6]. The overall added gain in the amplifier's passband is a factor of 6.

The array size of sensitive pixels is 188×192, which is less than the effective size of the array (208×192) because 20 columns of test pixels are present on the left-hand-side. Fig. 4 shows the layout of the SDAVIS192 chip on the Towerjazz 180 nm 6M1P technology. In the chip microphotograph, only the optical array is visible; in the RGBW split (see next paragraph) the rest of the chip is covered with a red+blue optical filter, to shield parasitic photodiodes from visible light.

A wafer split of the dedicated mask run included an RGBW (Red, Green, Blue, White) Color Filter Array (CFA) to add DVS color sensitivity as reported in [20]. Also, an SDAVIS192 test pixel is connected to 3 analog multiplexers to observe most internal nodes. APS readout uses the same internal 10-bit column-parallel ADC of [7].

III. Characterization

Unless otherwise specified, all characterization results refer to the monochrome SDAVIS192 chip.

A. Measurement Setup

To perform the measurements in which a uniform light source sweep is required, the setup of Fig. 5 was used. The aim of this section is to publish the detailed schematic of our LED driver, which we believe is a practical means of generating linear LED

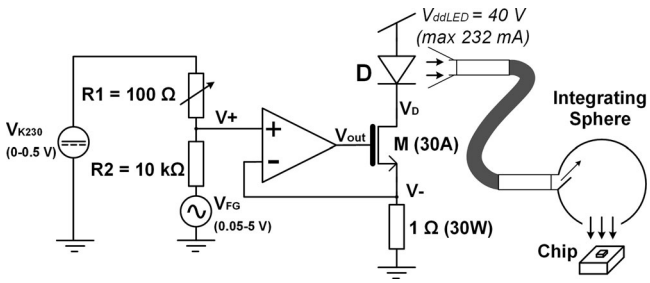


Fig. 5. Testing setup for uniform light intensity sweep.

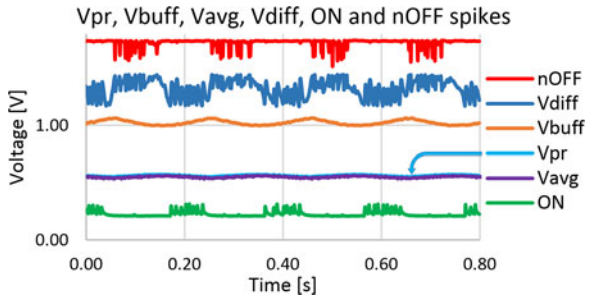


Fig. 6. Test pixel oscilloscope traces.

power output from a waveform generator in our testing, and thereby save development effort by other groups. This setup is an improvement on the set-ups used for testing [6], [17], [21]. An operational amplifier in feedback configuration was used to change the intensity of the light linearly with a function generator. This circuit forces the current through the high-brightness LED source to follow Ohm's law as the voltage on the 1Ω resistor changes. DC and AC components can be applied to the light whose illuminance at the integrating sphere output port can linearly span 10 to 3500 lux. Neutral Density Filters (NDF) were used to attenuate this range by various factors. To stimulate only a small part of the pixel array (to not saturate the AER bandwidth), the integrating sphere is removed. An interposing surface with a hole of diameter 0.5 mm only lets a small fraction of the fiber's output reach the sensor, which is mounted with a lens (1/2″ 4.5 mm f/1.4).

B. Test Pixel and Preamplifier Characterization

The transient response of the preamplifier was obtained from the test pixel that instruments pixel internal nodes. Fig. 6 illustrates the circuit response to a 5 Hz, 1.67 contrast ($\mathrm{I_{ACpp}/I_{DC}} = 50\%$)., 1 klux light oscillation with the setup of Fig. 5. The photoreceptor output $\mathrm{V_{pr}}$ responds with a 23 mV peak-to-peak oscillation. $\mathrm{V_{pa}}$ was not observable, but $\mathrm{V_{buff}}$ corresponds to $\mathrm{V_{pa}}$ with a negative offset of roughly a threshold voltage. $\mathrm{V_{pa}}$'s inferred value matches the observed behavior of $\mathrm{V_{obsA}}$ (observed by exposing the entire array at different illuminations with various bias settings). For a different $\mathrm{V_{ref}}$ bias, the transistor MN9 of Fig. 1 is in a different operating region and the transistor's κ_n changes. The $\mathrm{V_{buff}}$ to $\mathrm{V_{pr}}$ voltage gain ranges from 0.5 to 3.15 depending on $\mathrm{V_{ref}}$. In the region 200–400 mV the gain is about constant. The gain is lower than

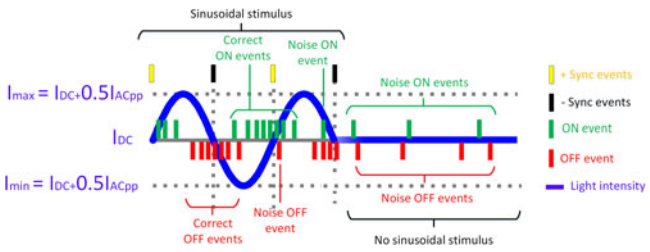


Fig. 7. Typical noisy DVS activity for sinusoidal and no sinusoidal input.

the estimated preamplifier gain because of the source-follower's gain κ_p lower than unity. A small ripple on every waveform is due to back-coupling of the ON and OFF events generated in the comparators' stage, though it is insufficient to generate further events. In Fig. 6, $\mathrm{V_{avg}}$, which represents the average $\mathrm{V_{pr}}$ value of the array, is identical to $\mathrm{V_{pr}}$ (all pixels see the same scene). $\mathrm{V_{diff}}$, the output of the differential amplifier, amplifies changes in $\mathrm{V_{pa}}$ by about 30 times. ON and active-low nOFF events appear in the corresponding rising and falling parts of the light cycle although their full amplitudes are not visible due to the low timescale resolution set on the oscilloscope.

C. DVS Contrast Threshold, SNR and FPN

Contrast sensitivity characterization estimates the sensor's temporal contrast threshold, which is defined as the minimum percentage temporal contrast necessary to trigger either an ON or OFF DVS event. In the ideal case, for a sinusoidal input, the sensor produces exactly $\mathrm{s_{ON}}$ events when the light intensity increases and exactly $\mathrm{s_{OFF}}$ events when it decreases and no events should be generated for no temporal contrast. In reality, $\mathrm{n_{ON}}$ noise events can appear during the negative part of the cycle, where only OFF events should be present and vice versa for $\mathrm{n_{OFF}}$ events, as Fig. 7 shows. More generally, the expected cyclic pattern of events, which could ideally be used to reconstruct the signal, can show jitter and corruption. This can be due to oscillations and coupling in the pixel but also to a temporal noise activity. This activity can be due to the leak in the reset switch [4], or, especially in lower lighting, it can be due to temporal noise at the photodiode exceeding the small thresholds and triggering events.

In order to achieve a higher contrast sensitivity, the ON and OFF thresholds of the pixel must be lowered to the minimum possible. Due to mismatch in the pixels, some end up having their $\mathrm{V_{diff}}$ passing threshold at reset making those pixels fire all the time (so-called 'hot pixels'). If contrast thresholds are computed without taking into account this type of fixed-pattern noise, it becomes hard to outperform the high contrast sensitivities of the state-of-the-art.

The single test pixel of [15] achieved a 50X reduction in contrast threshold from 15% by greatly increasing photoreceptor gain, but only by also running at very low bandwidth; additionally the continuous-time photoreceptor adaptation caused additional events that are difficult to interpret. Lichtsteiner *et al.* [4] quantified threshold variation by counting events produced by the passage of a stimulus of large 15:1 contrast. This method focused on measuring fixed pattern variation, while temporal

variation was addressed by a separate experiment quantifying latency jitter. Posch *et al*. [14] quantified contrast threshold in terms of response probability for a given step change stimulus, but did not characterize variation between pixels whether temporal or fixed-pattern and thereby gave no information about the usability of the sensor as sensitivity is increased. Lenero-Bardallo *et al.* [22] and Serrano-Gottaredona and Linares-Barranco [16] followed the method of Lichtsteiner *et al.* [4] but used a TFT monitor stimulus with a large contrast stimulus of 400%. Unlike Lichtsteiner *et al.* [4], Gottaredona and Linares-Barranco [16] did not establish the asymptote at which contrast threshold mismatch becomes independent of contrast threshold; rather, they justified a 1.5% contrast threshold at 0.9% fixed-pattern noise by including an image which shows a large edge visible to a human. They also stated that pixels could self-oscillate for very low thresholds (since, due to comparator mismatch, some thresholds will always be crossed [22]), but did not quantify the problem. Brandli *et al.* [6] and Yang *et al.* [17] used a modulated LED stimulus and integrating sphere to improve homogeneity of stimulus, (as proposed in [21]). Yang *et al.* [17] rightly pointed out that in order for the sensor to be performant, with low contrast threshold, the pixel response needs to be appropriately band limited and the illumination must be sufficiently high to reduce temporal noise. However, the 1% contrast threshold of the sensor was justified by placing an arbitrary cut-off coefficient of variation of contrast threshold of 35%.

The method proposed here estimates noise activity and the quality of the signal through a Signal-to-Noise Ratio (SNR). By so doing we characterize the usability of the sensor at various contrast thresholds and in varying conditions and bias levels. Counting all events of the same type in the correct part of the cycle (ON in upswing and OFF in downswing) and dividing by the number of periods, gives the number of events generated per cycle by the sensor, $\mathrm{s_{ON}}$ and $\mathrm{s_{OFF}}$. In all measurements, it was assured that no coupling events would appear in the wrong part of the cycle (ON events in the downswing and OFF events in the upswing). Recording the sensor's output for no contrast present after the AC stimulation allows estimation of the noise activity present in the time corresponding to a cycle in the AC stimulation: $\mathrm{n_{ON}}$ and $\mathrm{n_{OFF}}$. These may not correspond to the actual noise during AC stimulation, since all pixels are being stimulated and signal overtakes noise (also thanks to the refractory period). However, the influence of noise is still present in the cycle and does affect sensitivity and must be taken into account. $\mathrm{n_{ON}}$ and $\mathrm{n_{OFF}}$ constitute the noise floor and without prior knowledge of the signal, they would not allow understanding whether recorded events constitute signal or noise. As observed from measurements, $\mathrm{n_{ON}}$ and $\mathrm{n_{OFF}}$ do not increase over time but rather depend only on the specific bias settings and illumination. For their measurement, we wait two seconds after the end of the AC stimulation, to let the pixels recover from possible refractory period effects. Subtracting the $\mathrm{n_{ON}}$ and $\mathrm{n_{OFF}}$ from the $\mathrm{s_{ON}}$ and $\mathrm{s_{OFF}}$, gives the actual signals $\mathrm{s_{ON}} - \mathrm{n_{ON}}$ and $\mathrm{s_{OFF}} - \mathrm{n_{OFF}}$. Knowing these and the contrast of the light intensity gives the sensitivity of the sensor.

If a sinewave of DC light intensity $\mathrm{I_{DC}}$ is supplied with an AC peak to peak oscillation equal to $\mathrm{I_{ACpp}}$, the contrast in the upward phase of the cycle can be defined as $\mathrm{TCON_{ON}}$, which is the maximum value of light intensity $\mathrm{I_{max}}$ over the minimum $\mathrm{I_{min}}$, assuming a monotonic change:

$$\mathrm{TCON_{ON}} = \frac{\mathrm{I_{max}}}{\mathrm{I_{min}}} = \left(\frac{\mathrm{I_{DC}} + 0.5\mathrm{I_{ACpp}}}{\mathrm{I_{DC}} - 0.5\mathrm{I_{ACpp}}}\right) \tag{6}$$

In the case of the contrast in the downward phase of the cycle ($\mathrm{TCON_{OFF}}$), the ratio is reversed. To convert from a fractional increase ε (for $\varepsilon \ll 1$) to a percentage contrast threshold [4] for $\mathrm{s_{ON}} - \mathrm{n_{ON}}$ or $\mathrm{s_{OFF}} - \mathrm{n_{OFF}}$ events recorded (noise subtraction), the logarithmic threshold of the DVS can be approximated as in (7) and (8):

$$\theta_{\mathrm{ON}}\,[\%] = \frac{\ln\left(\mathrm{TCON_{ON}}\right)}{\mathrm{s_{ON}} - \mathrm{n_{ON}}} = \frac{\ln\left(1+\varepsilon\right)}{\mathrm{s_{ON}} - \mathrm{n_{ON}}} \approx \frac{\varepsilon}{\mathrm{s_{ON}} - \mathrm{n_{ON}}} \tag{7}$$

$$\theta_{\mathrm{OFF}}\,[\%] = \frac{\ln\left(\mathrm{TCON_{OFF}}\right)}{\mathrm{s_{OFF}} - \mathrm{n_{OFF}}} \tag{8}$$

This approximation is valid for $\varepsilon < 0.2$ (as it is the case for SDAVIS192) for an error within 3%. θ_{ON} and θ_{OFF} are computed from the median event count per cycle minus the median noise ($\mathrm{s_{ON}} - \mathrm{n_{ON}}$ and $\mathrm{s_{OFF}} - \mathrm{n_{OFF}}$) of all pixels of the array. The median count is used to filter out outlier hot-pixels. This effect was observed in measurements when the mean contrast threshold would be 2% and the median 30%, highlighting the masking effect of hot pixels.

With the setup of Fig. 5, for a particular homogenous illumination, a 1.67 $\mathrm{TCON_{ON}}$ contrast is applied with a 1 Hz sinusoid and events are recorded from the sensor with the APS turned off. The frequency is low enough to not saturate the output bus (at 15 Hz, 20 eps per pixel occupy the entire 11 Meps bandwidth), and at the same time higher than the mean firing of the background leak activity (typically 0.1 Hz). Thresholds are set initially once to a good sensitivity (and the sensor produces correlated activity) and then bias, contrast and background illumination sweeps are performed. The half cycles of the sinewave in which count events are timestamped through sync events gathered by the FPGA, from the $\mathrm{V_{FG}}$ function generator. At the end of 10 cycles, an extra 10 s are recorded with DC input count $\mathrm{n_{ON}}$ and $\mathrm{n_{OFF}}$. For every sensitivity level, a Signal-to-Noise Ratio (SNR) is computed as:

$$\mathrm{SNR_{ON}}\,[\mathrm{dB}] = 20\log_{10}\left(\frac{\mathrm{s_{ON}} - \mathrm{n_{ON}}}{\mathrm{n_{ON}}}\right) \tag{9}$$

$$\mathrm{SNR_{OFF}}\,[\mathrm{dB}] = 20\log_{10}\left(\frac{\mathrm{s_{OFF}} - \mathrm{n_{OFF}}}{\mathrm{n_{OFF}}}\right) \tag{10}$$

The SNR ratios were taken as signal minus noise over noise, because the minimum $\mathrm{s_{ON}}$ or $\mathrm{s_{OFF}}$ which can be recorded during AC stimulation is $\mathrm{n_{ON}}$ or $\mathrm{n_{OFF}}$ (just pure thermal or leakage noise). If SNR was taken as $\mathrm{s_{ON}}/\mathrm{n_{ON}}$ or $\mathrm{s_{OFF}}/\mathrm{n_{OFF}}$, then in the worst case scenario, the SNRs would just be 0 dB for $\mathrm{s_{ON}} = \mathrm{n_{ON}}$ or $\mathrm{s_{OFF}} = \mathrm{n_{OFF}}$. With our SNR method, SNRs fall below zero dB for $\mathrm{s_{ON}} < 2\mathrm{n_{ON}}$ or $\mathrm{s_{OFF}} < 2\mathrm{n_{OFF}}$. Finally, for $\mathrm{s_{ON}} = \mathrm{n_{ON}}$ or $\mathrm{s_{OFF}} = \mathrm{n_{OFF}}$, SNR is equal to minus infinity dB (no input-dependent signal).

To test how far the sensitivity of SDAVIS192 can be pushed while still maintaining reliability, ON and OFF threshold biases

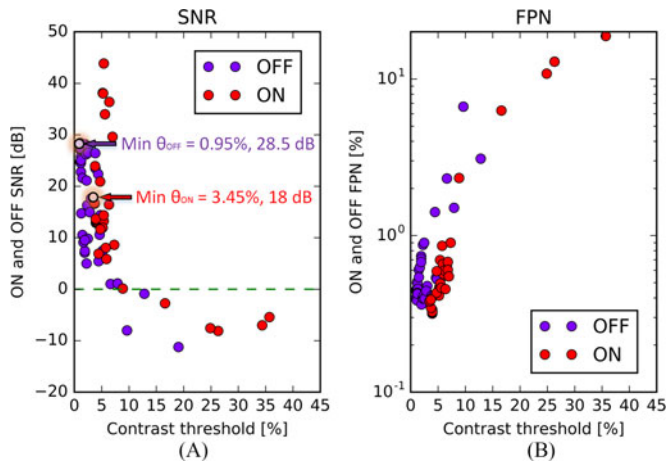


Fig. 8. SDAVIS192 SNR_{OFF} and SNR_{ON} (A) and percentage FPN_{OFF} and FPN_{ON} (B) vs the respective contrast thresholds at 1 klux, with IR cut filter.

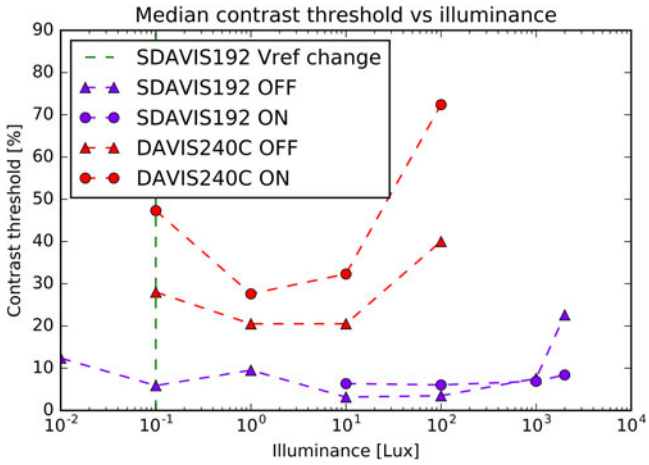


Fig. 9. SDAVIS192 and DAVIS240C ON and OFF contrast thresholds vs illumination level. No point at a particular illumination means SNR<0.

were swept with a sensor illumination of 1 klux in various combinations. Fig. 8(a) plots the resulting $\mathrm{SNR_{ON}}$ and $\mathrm{SNR_{OFF}}$ versus θ_{ON} and θ_{OFF}. The lower the contrast threshold, the more the signal with respect to the temporal noise (and therefore the higher the SNR). If contrast threshold is lowered further, however, the pixels start to self-oscillate and the SNR abruptly drops to minus infinity dB ($\mathrm{s_{ON}} = \mathrm{n_{ON}}$ or $\mathrm{s_{OFF}} = \mathrm{n_{OFF}}$). Minimum contrast thresholds were established by the lowest contrast threshold achievable with SNR > 0. These are, at 1 klux for SDAVIS192, $\theta_{\mathrm{OFF}} = 0.95\%$ with $\mathrm{SNR_{OFF}} = 28.5$ dB and $\theta_{\mathrm{ON}} = 3.45\%$ with $\mathrm{SNR_{ON}} = 18$ dB. Approaching the limit of contrast sensitivity, more OFF than ON events are generated and therefore OFF has a lower contrast threshold. This is the case because the two different bias current references of the comparators determine their different speeds, which limit the encoding favoring one type of event [17]. For prior DAVIS sensors it is the opposite, due to the SDAVIS192 inverting preamplifier. In the case of DAVIS240C [6], minimum contrast thresholds coincide with the ones of Fig. 9 at 1 lux: $\theta_{\mathrm{ON}} = 27.6\%$ with $\mathrm{SNR_{ON}} = 19$ dB, $\theta_{\mathrm{OFF}} = 20.5\%$ with infinite $\mathrm{SNR_{OFF}}$ (no $\mathrm{n_{OFF}}$).

Fixed Pattern Noise (FPN) can be estimated from the pixels' mismatch (spread) in firing rates, as done by Lichtsteiner *et al*. [4]. Plotting an event histogram, however, shows that the distribution of firing rates is not normal and that there are a certain amount of kurtosis and skew. Standard deviation (normally used to compute FPN) is a summary statistic of a more complex underlying distribution. The Cumulative Distribution Function was used to provide lower and upper percentiles of number of events per cycle, corresponding to one standard deviation in a normal distribution (15.8%–84% percentiles divided by 2). These can then be used to obtain $\mathrm{FPN_{ON}}$ and $\mathrm{FPN_{OFF}}$ expressed as percentage contrast [4]. For the same set of biases of Fig. 8(a), $\mathrm{FPN_{ON}}$ and $\mathrm{FPN_{OFF}}$ are plotted in Fig. 8(b) as a function of contrast threshold. The higher the contrast threshold, the higher the FPN: [4] attributed this trend to amplifier gain mismatch. For the sets of biases to achieve the lowest θ_{ON} and θ_{OFF}, $\mathrm{FPN_{ON}} = 0.2 - 1\%$ contrast and $\mathrm{FPN_{OFF}} = 0.8 - 0.9\%$ contrast for SDAVIS192. In comparison, for DAVIS240C, $\mathrm{FPN_{ON}} = 7.5 - 15\%$ contrast and $\mathrm{FPN_{OFF}} = 5.5 - 8.5\%$ contrast.

As sensitivity increases, temporal noise becomes an important factor since the number of "hot-pixels" firing increases. Due to their low threshold, these pixels lock themselves in a state of high firing of one type. If a pixel is considered 'hot' if it fires more than 200 times per cycle, the number of ON hot-pixels for SDAVIS192 is almost zero, while OFF hot-pixels reaches 0.5% for $\theta_{\mathrm{OFF}} < 1.2\%$. Given a 11 Meps system bandwidth, this hot pixel activity consumes $0.005 * 188 * 192 * 200/11\mathrm{M} = 0.3\%$ of the total bandwidth.

D. DVS Dynamic Range

In [4], [6], [16], [23] and [17] DR was obtained by observing the sensor under extremely low and high light conditions by 'visually recognizing' a correlation in the events which would match the stimulus. However, this is not case of this work because, although the sensor still shows some correlation with the input signal (a 1 Hz flashing can be recognized even beyond the given DR of SDAVIS192), noise is dominant and SNR<0. If the noise is larger than the signal, then the sensor is difficult to use in practical applications. Prior work also made no distinction between ON and OFF events.

Here the illumination was varied in the range 0.01 to 2 klux. An IR-cut filter was used to ensure correct NDF attenuation. Fig. 9 illustrates contrast thresholds versus illuminance level measured at the sensor for a single set of OFF and ON biases, for both DAVIS240C and SDAVIS192, which work best at most illuminations. It therefore does not mean that the sensor cannot work outside the range shown in Fig. 9, however, the sensor exhibits its larger DR in this range. In the case of DAVIS240C, for the selected set of biases giving SNR>0, the DR is 60 dB for ON and OFF. For SDAVIS192, the global operating range bias was changed only when necessary: the same set of biases, which would still give SNR > 0 and contrast thresholds lower

than 100% was kept whenever possible (this change is marked with the green vertical line occurring at 0.1 lux). The intra-scene DR of SDAVIS192 (for the same set of biases) is at least 50 dB for ON and at least 70 dB for OFF. "At least" means that the DR could be higher, but the uniform light intensity of the setup of Fig. 5 cannot reach 10 klux to verify this claim. The relatively small DR of SDAVIS192 comes from the fact that the sensor has high gain. The DR of DAVIS240C is however similar, although the smaller gain, because of the smaller s_{ON} and s_{OFF} with respect to n_{ON} and n_{OFF} does not guarantee SNR>0 on a larger illumination range. Thus SDAVIS192 has more points in the graph of Fig. 9 where SNR>0 and contrast threshold is less than 100%. The parabolic shape of Fig. 9 shows that the event count decreases as the sensor reaches either ends of its intra-scene DR. Thanks to the global V_{ref} bias sensitive to mean illumination, the total operating range of the DVS sensor spans at least 110 dB for OFF for the same set of biases. For the same ON bias, no extension of the DR can be made and this remains at least 50 dB (similarly to the 54 and 60 dB of [22] and [16]).

The ON pathway has a lower DR than the OFF pathway from two phenomena that take place at low and high intensities. At low light levels, the photodiode dark current decreases the signal contrast in the photocurrent. At the same time, junction leakage in the S1 reset transistor continually produces a fictitious OFF temporal contrast. The combination of these two effects means that at low intensity, the sensor favors making OFF events. The APS dark current (estimated at 16 k e^-/s for SDAVIS192, in Section III-F) reduces the photocurrent contrast. At 0.01 lux the estimated photon contrast with respect to the estimated dark current is less than 2%. However, across the entire array, pixels that are near threshold respond synchronously, resulting in a median contrast sensitivity of 10%. At high intensities, parasitic photocurrent in the S1 reset transistor source-drain junction generates an increasing number of OFF events, again producing a fictitious OFF temporal contrast. This parasitic photocurrent problem was already known from [4] and [6] and therefore particular care was taken to shield the reset transistor with metal and an active guard ring. However, the shielding is not perfectly effective. SDAVIS192 functions at 3.5 klux, while DAVIS240C, stops working at 2 klux as parasitic photocurrents in the insufficiently shielded bias generator deadlock the chip for this set of biases. (DAVIS240C has no optical shield on peripheral circuits.)

E. DVS Pixel Bandwidth

The Frequency Response (FR) of the single pixel was first obtained by Lichtsteiner *et al*. in [4] together with the second-order filter behavior of the pixel. The pixel FR can be measured using the same setup as for contrast threshold, but with only a few pixels being stimulated, and using a sinewave as the modulating signal. We use a lens to image a pinhole. Knowing the diameter of the pinhole and the circular field of view of the Tektronix J17 luminance meter, we can measure the pinhole luminance, which could be set to a maximum of 30 cd/ m^2. Geometrical optics is then used to compute the sensor illuminance. By

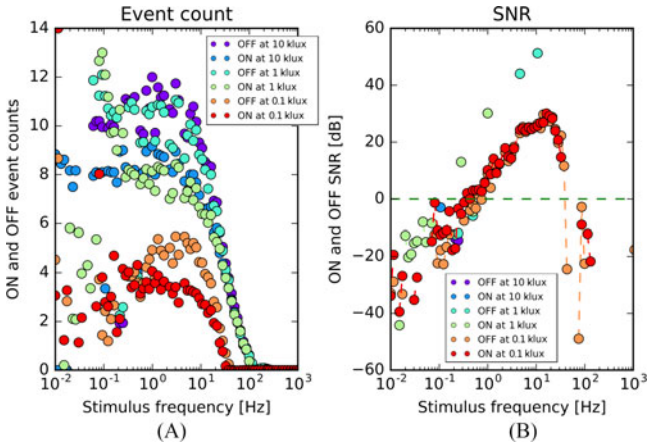


Fig. 10. (A) Pixel bandwidth of SDAVIS192 as a function of frequency (with contrast of 1.67) for bias settings giving good contrast sensitivity. 10 events per cycle corresponds to a 5% temporal contrast threshold. (B) Corresponding ON and OFF SNRs (SNRs at 1 and 0.1 klux are not visible since infinite). Illuminance values are at the sensor focal plane.

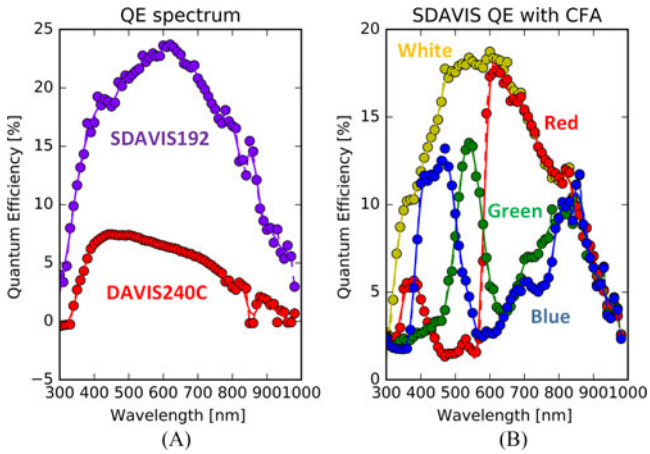


Fig. 11. (A) External Quantum Efficiency (QE) spectrum of Active Pixel Sensors of SDAVIS192 and DAVIS240C. (B) Measured external QE spectrum of the RGBW channels of SDAVIS192. Results obtained at IMEC.

sweeping the 1.67 $TCON_{ON}$ contrast sinewave frequency and measuring the number of events produced per half cycle, with noise subtracted, it is possible to obtain a pixel FR. Since high-bandwidth biases are not the purpose of SDAVIS192, but rather high-contrast sensitivity biases with good SNR, only these are presented. Fig. 10(A) illustrates ON and OFF FRs with a contrast of $TCON_{ON} = 1.67$. θ_{OFF} and θ_{ON} are both about 5%. With these biases, the 3 dB cutoff is about 20 Hz and the absolute cut-off is about 100 Hz. The preamplifier, weakly biased to achieve higher sensitivity, is the main reason for this FR that hardly changes with intensity. The roll-off is approximately that of a second-order lowpass filter and no ringing was observed.

Noise in the measurement is considerable below 0.1 Hz since recordings of 10 cycles of activity can take minutes. During these long recordings any sudden source of noise can be can be falsely detected as signal. The corner frequency of the pixel is thus uncertain and can only be defined by looking at the specific SNR. In [4], no lower cut off was measured. Below the

TABLE I
SPECIFICATION TABLE

	This work: **SDAVIS192**	Brandli *et al*. [16], DAVIS240C, measured with current setup	Serrano-Gotarredona and Linares-Barranco [16]	Yang *et al*. [17]
Functionality	Asynchronous temporal contrast + APS	Asynchronous temporal contrast + APS	Asynchronous temporal contrast	Asynchronous temporal contrast
CMOS technology	180 nm 1P6M MIM CIS	180 nm 1P6M MIM CIS	350 nm 2P4M OPTO	180 nm 1P6M
Chip size mm^2	5×5	5×5	4.9×4.9	3.2×1.6
Array size	192×188	240×180	128×128	60×30
Pixel size μm^2	18.5×18.5	18.5×18.5	30×31	31.2×31.2
Fill factor	21.2%	22%	10.5%	10.3%
Peak absolute QE	24%	7%	N.A.	N.A.
Pixel complexity	47 transistors, 2 MIM caps, 1 MOS cap, 1 BPD, 1 microlens	48 transistors, 2 MIM caps, 1 MOS cap, 1 SPD	N.A.	N.A.
DVS interface	**8-bit word-serial AER**	**8-bit word-serial AER**	Word-parallel AER	6-bit word-serial AER
Supply voltage	3.3/1.8 V	3.3/1.8 V	3.3 V	1.8 V
Power consumption	DVS only (best): 15 mW (100k eps), 20 mW (2M eps) @ 100 lux. Normally 60 mW. APS only: 50 mW (1 fps), 60 mW (30 fps) @ 100 lux	Minimum (unconfirmed) at 4 mW (low activity), 15 mW (high activity)	2.6 mW (1k eps) 4 mW (100k eps) 95 mW (20M eps)	0.72 mW (10k eps)
DVS dynamic range	ON: at least 50 dB intrascene. OFF: at least 70 dB intrascene and at least 110 dB overall. With SNRs > 0	ON: 60 dB intrascene OFF: 60 dB intrascene. With SNRs > 0	60 dB intrascene, 120 dB overall[a]	130 dB[a], obtained by waiving a hand in front of the sensor and arbitrarily recognizing it by eye
DVS minimum contrast sensitivity	ON: 3.45% (28.5 dB SNR)[b] OFF: 0.95% (18 dB SNR)[b]	ON: 27.6% (19 dB SNR)[c]. OFF: 20.5% (Inf dB SNR)	1.5%[a]	1% for 35% coefficient of variation[a]
Max bandwidth	54 Meps (self-ack) 11 Meps (FPGA), 50 fps	54 Meps (self-ack) 11 Meps (FPGA), 50 fps	20 Meps	10 Meps
DVS optimized minimum latency	ON: 35 μs @ 30 kcd/m^2 OFF: 10 μs @ 30 kcd/m^2	ON: 15 μs @ 30 kcd/m^2 OFF: 20 μs @ 30 kcd/m^2	3.2 μs @ 2 klux[a]	N.A.
DVS normal latency	0.1-2 ms[b]	0.1-2 ms[b]	Up to 6 ms @ 0.2 lux	N.A.
DVS FPN	ON: 0.2–1% contrast OFF: 0.8–0.9% contrast	ON: 7.5–15% contrast OFF: 5.5–8.5% contrast	0.9% contrast[a]	35% coefficient of variation at 1% contrast[a]
DVS bandwidth	ON, OFF: 0.1–100 Hz[d]	ON, OFF: 0.1–1k Hz[d]	N.A.	N.A.
APS dynamic range	56.12 dB	52.57 dB	-	-
APS FPN	3.39%	0.97%	-	-
APS DSNU	25.79 e^-	21.38 e^-	-	
APS dark signal	16k e^-/s (0.748 nA/cm^2)	18k e^-/s (0.858 nA/cm^2)	-	-
APS readout noise	61 e^-	57 e^-	-	-
APS conv. gain	22 $\mu V/e^-$	24 $\mu V/e^-$	-	-

[a]ON and OFF not specified separately. [b]At 1 klux. [c]At 1 lux. [d]At 10 klux chip Illumination (30 cd/m^2 source).

corner frequency, $s_{ON} = n_{ON}$ and $s_{OFF} = n_{OFF}$. In this work, $s_{ON} - n_{ON}$ approaches zero around DC as background noise events take over and hide the signal. As illuminance decreases, dark current contrast reduction reduces the event count, but the lower cut-off frequency remains very similar. As can be seen in Fig. 10(B), SNRs plummet as soon as the event count decreases and the fixed background noise takes over. For these biases, noise is almost non-existent at 10 and 1 klux chip illuminations (very large SNRs). At 10 lux, noise becomes visible and SNRs appear in the range from -40 to 30 dB. In the work of [22], [16] and [17], unrealistic illuminations of up to 50 klux were used to estimate FR.

F. APS Characterization and Dark Current

APS was characterized using the Photon Transfer Curve (PTC) measurement procedure of the EMVA1228 standards

[24]. All results of APS characterization are reported in Table I. The 3.39% APS FPN of SDAVIS192 is about three times the 0.97% of DAVIS240C because the latter uses an external ADC rather than the on-chip column parallel ADC of SDAVIS192. DAVIS240C, has an 18 k e^-/s APS dark current (0.858 nA/ cm^2) at 25°C. The SDAVIS192 APS dark current should ideally be one third of such value for similar transistor and photodiode sizing, since it uses BPD instead of Surface Photodiode (SPD), but is only 21% smaller: 16 k e^-/ s (2.6 fA or 0.748 nA/ cm^2). This high leakage is close to vanilla n-well photodiode performance (1 nA/ cm^2). This values are also worse than the 4 fA measured by Lichtsteiner *et al*. [4] on the larger nwell photodiodes manufactured with a non-CIS process. The reasons for this problem might be large transistor junction leakage in the APS readout pathway. Read noise σ_{read} (sensitivity) are less than 1 Digital Number (DN, 1.26 mV for SDAVIS192 and 1.47 mV for DAVIS240C) and are therefore limited by ADC quantization: 61 e^- and 57 e^- respectively.

G. External QE Characterization

Measurements of absolute external Quantum Efficiency (QE) were performed at the laboratory of Interuniversity Micro Electronics Center (IMEC), Belgium. External QE is the ratio of collected charge carriers over incident photons at a specific wavelength. The setup consisted of a Xenon-based white light source whose light passes through a monochromator, which, through a series of mirrors and shutters, only lets through a specific wavelength. PTC spectral measurements were performed at each wavelength. Comparing the photo-generated current to the current measured from a Newport 918 series reference photodiode with Newport 1936 power meter, the results shown in Fig. 11(A) were obtained for the monochrome sensors. The external QEs of DAVIS240C and SDAVIS192 can be compared fairly since their pixels have the same size and similar FF. The FF determines the maximum QE for each sensor if no microlens is present (as for DAVIS240C). Both sensors are front illuminated with nearly the same metal stacking. As Fig. 11(A) shows, SDAVIS192 has an overall higher QE spectrum. The thick microlenses (of size 15 × 15 μm^2 and thickness 3 μm with additional spacer), appear to be effective, since the peak QE of 24% is higher than the 21.2% FF. The diffuse incident light in the QE measurement probably underestimates the actual QE using a lens, since in that case the light would likely be better focused onto the photodiode. In comparison, the QE of the DAVIS240C (FF of 22%) peaks at 7% at 450 nm, with a magnitude 3.4X smaller than that of the SDAVIS192.

Part of the increase in external QE is also due to the use of a BPD for SDAVIS192 with respect to a SPD in DAVIS240C. The BPD almost doubles the junction area and reduces the chance that photo-generated carriers are captured in trap-states at the interface with the SiO_2. It can be seen that the deeper BPD is slightly more sensitive to longer wavelengths while the SPD is slightly more sensitive to shorter wavelengths that could be absorbed in the BPD surface cover implant. Both photodiodes have a sharp cut-off around 350 nm because UV light is reflected by the N-BK7 package cover glass. The repeatable complementary fluctuation at 850 nm is not explained. For longer wavelengths, the external QE decreases due to the shallow photodiodes and small reverse bias voltage. Overall, the thick microlens combined with BPD appears to approximately triple the external QE for this diffuse incident illumination. The QE of a color split of the SDAVIS192, with a CFA on top of the pixels with RGBW pattern, was also characterized in [20] and the QE spectrum of the single RGBW channels is shown in Fig. 11(B).

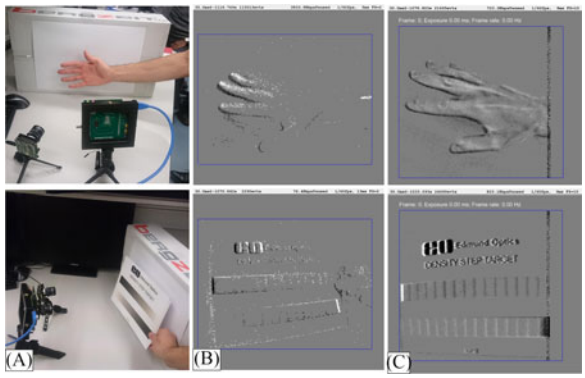


Fig. 12. (A) Moving stimuli: hand and Edmund Optics density step chart; (B) DAVIS240C DVS raw output with Full Scale (FS) of 1; (C) SDAVIS192 DVS raw output with full scale of 10 events. Both (B) and (C) show 30 ms DVS time slices. ON events are represented in white and OFF events in black.

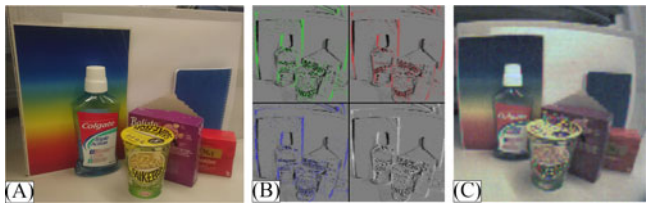


Fig. 13. (A) Ground truth (cellphone camera); (**B**) RGBW events separate by color (RGBW ON events are represented in RGBW respectively and OFF events in black); (C) Interpolated RGBW APS frame. Test pixels are ignored.

H. Sample Monochrome and RGBW Outputs

Fig. 12 compares 30 ms DVS time slices from SDAVIS192 with DAVIS240C (video available at [24]). Fig. 12(a) shows the setup for the moving hand and Edmund density step chart. The density step chart has 10 steps of 0.1 density. The contrast of each step is $10^{0.1} = 1.26$. The DVS gray scale is 10 events, i.e., from gray, 10 events ON or OFF makes the image full white or black. SDAVIS192 produces on average 10X more events revealing more gray scale and detail and better distinguishes the hand from its shadow. Fig. 13 shows results from the RGBW split of SDAVIS192 that can provide alternating RGBW frames and RGBW DVS events [20]. Fig. 13(A) shows the ground truth scene. Fig. 13(B) shows the separate RGBW DVS channels. Fig. 13(C) shows a linearly interpolated APS frame. Table I compares the design and measured specifications of SDAVIS192 with previous work.

IV. Application in Neural Imaging

The high sensitivity of SDAVIS192 makes it potentially suitable for imaging of neuronal activity using calcium-sensitive flo-

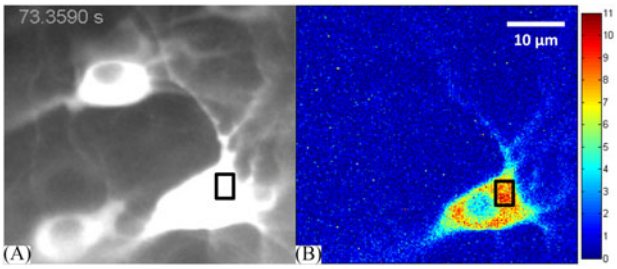


Fig. 14. (A) Orca sCMOS frame compared to (B) the corresponding 100 ms SDAVIS192 integrated DVS time slice in RGB scale, with regions of interest indicated. The two cells on the left of (A) are not visible in (B) because the long calcium reuptake (lasting a few seconds), is below the 0.1 Hz corner frequency of the detection of the DVS pixel. This long reuptake is related to the presence of bicuculline, which also makes the dendrites flicker. Since, the calcium just enters the cytoplasm and not the nucleus, the activity in the soma in (B) is low. The brightness in the nucleus of the corresponding cell in (A) is due to the very high extracellular calcium concentration, which emits light changing too slowly for the DVS to detect. Recordings: 40x-GCamp6f-fullLED-15-26-crop (APS) and 152640x (DVS).

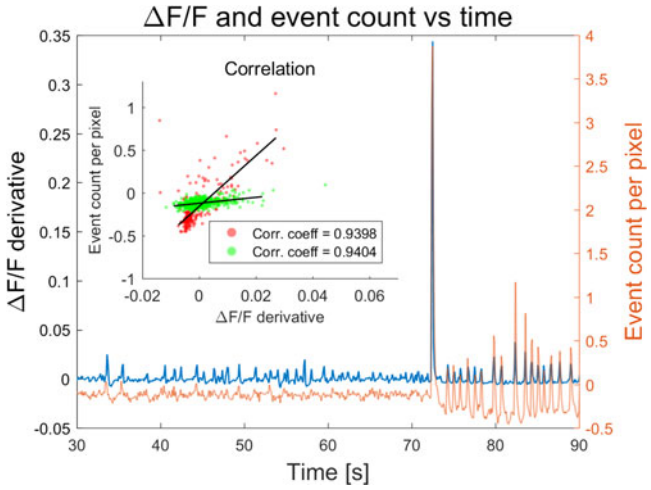


Fig. 15. Synchronized $\Delta F/F$ of the region of interest of the Orca camera frame aligned in time with the DVS binned activity corresponding to the same area. The green scatter plot corresponds to the correlation before the bicuculline's effect (up to 72 s) and the red refers to the correlation after this.

rescence. Current imaging systems use expensive frame-based CCD or scientific CMOS cameras with 75 dB or more of DR and peak QEs up to 70%, but only about 15 fps frame rate when recording at full resolution. The output of a minute of recording can fill over 40 GB of storage. A DVS sensor can potentially reduce redundancy and increase time resolution because it responds only to fluorescence changes, which streamlines recordings from large neuron populations.

As a feasibility test of this capability and the first reported test of a DVS sensor for fluorescent imaging, the SDAVIS192 sensor was connected to the output of a microscope (a Zeiss Axioscope FS with a 40x NA 0.8 water dipping objective) with a 50-50 splitter for comparison with a Hamamatsu Orca V2 sCMOS. Hippocampal organotypic slice cultures were prepared from P6 mice(29)[1] and virally transduced to express the "fast" calcium indicator GCaMP6f [10]. After 3 weeks in culture, slices were transferred to an upright microscope and perfused with aCSF containing bicuculline (50 μM), to block GABAA receptors and enhance neuronal excitation. Action potentials result in calcium influx, causing a conformational change of the genetically encoded fluorescent protein, changing its emission properties. The green fluorescence is separated from the 470 nm excitation light by an emission filter (525/39 nm). Fig. 14 shows the parallel recording of several neurons with the Orca camera and of the SDAVIS192 DVS output at a time point when the highlighted neuron is becoming brighter. The 100 ms of DVS activity clearly highlights the active cell and its dendrites compared with the relatively silent background. The frame rate of the Orca camera was limited to 10 Hz by the 100 ms exposure needed to image at the low intensity. Quantitative analysis of APS frames from SDAVIS192 show that the brightest parts of the scene produce a photocurrent of about 67k e-/s, equivalent to about 0.1 lux, which is only about 4 times the dark current.

Fig. 15 compares the 100 ms bins of DVS activity of the selected 15 $\times$ 10 pixel Region Of Interest (ROI) of Fig. 14 with the $\Delta F/F$ of the same region over 60 s of activity before and after the bicuculline effect. $\Delta F/F$ is the normalized derivative of grey scale ROI as detected by the Orca camera over time. It is the percentage change from the previous point in time of the raw ROI average grey value. The DVS activity is synchronized in time with the fast transient activity, although slow dynamics are not visible in this simple binning analysis. The inset of Fig. 15 shows that the two signals correlate with a coefficient of 0.94. Correlation was also obtained for ROIs comprising each full cell appearing in the recording and was found to be within 3-4% of this value.

The dark background is out of the intrascene DR of SDAVIS192, which results in the pixels exposed to it to be noisy and sometimes self-oscillate, occupying bandwidth and memory. This effect could be mitigated by either filtering in software the oscillation by correlation, since unstable pixels fire at a mean frequency of 10.5 Hz or, as experimented until now only with synthetic fluorescent beads moving in a solution (the experimental calibration sandbox), by increasing background illumination. This in turn reduces the contrast of the fluorescence, but the higher sensitivity of SDAVIS192 possible, since working in its best operating range, allows to pick it up. This was not attempted with the in-vitro hippocampal organotypic slices of this experiment because they were shared among other research and phototoxic effects were not known. V_{ref} in this dim illumination is non-ideal, but the signal is still correlated (Fig. 15, green scatter plot), although with smaller contrast sensitivity. V_{ref} was not adapted during the recording to avoid producing noise events due to bias switching and to leave the sensor in an operating range working well with the brightness of the neurons.

When bicuculline takes effect at 72s, the sensor begins to work in its normal operating range (red scatter plot of Fig. 15). As the average illumination of the ROI increases, so does the sensitivity of the sensor. For example, at 50 s, the contrast threshold is 13.3% on average (ON and OFF together). After the large 35% increase in illumination at 72 s, at which

[1]All animal procedures were approved by the Cantonal Veterinary Office and carried out according to the Center for Laboratory Animals guidelines of the University of Zurich.

TABLE II
SPECIFICATION TABLE COMPARING THE PERFORMANCE OF SDAVIS192 TO THE STATE-OF-THE-ART BIOIMAGING SENSORS

	This work: **SDAVIS192**	Scientific CMOS (sCMOS)	Charge-Coupled Device (CCD)	Electron Multiplication CCD (EMCCD)
Number of pixels	36k	2M-5M (ex. Hamamatsu Orca)	6M (ex. Zeiss Axiocam 506)	0.25M (ex. Andor iXon)
Frames per second	NA, pixel readout is asynchronous	From 10 fps to >100 fps for resolution <<1MP	10–15 fps	From 30 fps to 70 fps with pixels binning
Peak absolute QE	24%	65%–80%	Up to 70%	Up to 90%
Price	NA	\$ 10 k–20 k	\$ 3 k –10 k	\$ 30 k–40 k
Read noise	APS: 61e^-. DVS: NA	<2 e^-	6-7 e^-	<1 e^-
Additional cooling	No	From −30 °C to −10 °C	From −80 °C to −100 °C	From −80 °C to −100 °C
Dynamic range	DVS: ON: >50 dB intrascene. OFF: >70 dB intrascene and >110 dB overall.	Up to 90 dB	∼70 dB	∼80 dB

the contrast threshold is 7.56%, the average contrast threshold from such point onward becomes 3.45% as the background illumination is higher and matches V_{ref}. The DVS sensitivity is still less than optimal as the ON and OFF contrast thresholds were slightly increased to reduce the quantity of events due to noise in low illuminations. An FFT spectral analysis of the DVS activity at bin width from 10 ms to 200 ms shows that the fluorescence signal is limited to about 7 Hz bandwidth, confirming that in this experiment the 10 Hz Orca frame rate was sufficient. The limited signal bandwidth is due to the slow kinetics of this calcium indicator. In these conditions the SDAVIS192 has a similar performance to conventional imagers. Table II shows the comparison between SDAVIS192 the state-of-the-art technologies employed in high-end bio-imaging systems. These are mainly 3 technologies: scientific CMOS (sCMOS), Charge Coupled Devices (CCD) and Electron Multiplication CCDs (EMCCD), for single photon detection. The average specifications reported in Table II were gathered from the website of the manufacturers and from the analyses of [25] and [26]. SDAVIS192 does not outperform any of these technologies in terms of peak QE and read noise. These values, however, relate to the process technology available. The potential advantage of SDAVIS192 is the fact that it does not suffer heavily of the resolution/speed tradeoff which affects the other technologies. As a matter of fact, all three compared technologies require to cut down the resolution of the image read-out if the desired frame rate is higher than 10–30 fps. If there is no significant activity happening in the scene, the SDAVIS192 can successfully capture all information at sub-millisecond resolution, without recording enormous quantities of redundant data. The reduction in data redundancy, which will be improved by the filtering, probably on FPGA, of the synchronous noise oscillations, is one of the key advantages of SDAVIS192. Finally, SDAVIS192 does not use cooling and it is potentially less expensive than the other three technologies.

The main advantage of the SDAVIS192 sensor (its selectable artificial frame rate of event bin time slices) appears when the fluorescent signal contains higher frequency components as in the case of other calcium indicators or voltage indicators. With these indicators, not available yet at the Zurich Brain Research Institute, SDAVIS192 could capture neural responses containing frequency components up to 100 Hz (as Fig. 10 shows) with sub-millisecond latency. New developments in genetically-engineered two-component voltage sensitive dyes (VSD) result in fact in fast and high contrast indicators in selected cell populations. VSDs have the advantage of providing much faster kinetics: the optical signal recorded using VSDs in response to a single action potential last 6–7 ms. Such indicators suggest the possibility to enable the recording of high frequency spike trains with higher fidelity. DVS sensors can prove to be the right candidate for single spikes recording, due to their high temporal resolution. However, improvements are needed in order to reduce the background noise due to the low illumination working condition (20 mW/ mm^2 for VSDs). The experiment presented with the only available indicator available at the time does not show this feature yet, but it shows the feasibility of the SDAVIS pixel topology as the basis of future event-based bioimaging sensors, where ideally this higher gain DVS pixel is combined with back illumination as introduced in [27] and leakage event reduction techniques such as described in [28].

Finally, while it may seem that APS feature is redundant, it is useful for setup calibration. Since the DVS sensor does not see stationary items, focusing the lens of the sensor becomes a tedious task as moving the setup is not possible. APS helps in this regard and allows checks on the lens cleanliness, focused area and microscope correct functioning. This makes APS useful even if it does not work concurrently with DVS.

V. CONCLUSION

The design of the SDAVIS192 and its characterization show its lower contrast threshold (down to 0.95% for OFF and 3.45% for ON) and at the same time, higher average DVS SNR of 30 dB. This reduces the intrascene DR to 50 and 70 dB for ON and OFF, but the digital adaptation mechanism inspired by [16] extends it up to 110 dB. Though the sensitivity of the sensor is similar to that reported in [16] and [17], the methods described in this work introduce a quantitative estimate of the sensor's performance through SNR, to assess the boundaries of its operation capabilities (minimum contrast threshold, latency and maximum DR). The 1:1 comparison with DAVIS240C shows how the numbers first reported in [6] change depending on the criterion selected and on the desired signal quality (SNR>0). Overall, all specifications were also obtained for ON and OFF separately.

While this work also reports standard APS characterization results, it presents the first detailed DVS external quantum efficiency characterization, that was introduced in [20]. SDAVIS192 is also the first DAVIS sensor to produce, along with the color APS frames first introduced in [17], color RGBW DVS events (as used for image reconstruction in [20]).

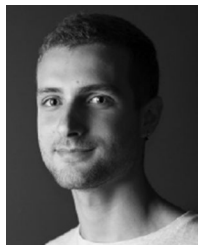
A real calcium imaging experiment, as opposed to the simulation of [17], is shown in this work. The preliminary results are promising (0.94 correlation and detection of all intensity transients), but future work will improve the quality of the SDAVIS192 sensitivity in low light conditions, which is critical in this field. The aim is to move to in-vivo imaging with VSDs to resolve single action potentials. VSDs have latencies below 0.2 ms, but also $\Delta F/F$ below 2% [29]. Attempts at in-vitro color fluorescence microscopy, where different VSDs operate at different color bands, will also be made with the color-sensitive version of SDAVIS192.

ACKNOWLEDGMENT

The authors would like to thank the Sensors Group at INI Zürich, in particular V. Villanueva for the PCB designs. They would also like to thank B. Linares-Barranco from the National Microelectronics Center for useful discussions and V. Motsnyi for his help in QE measurements. Finally, they would like to thank D. Göckeritz-Dujmovic for preparing slice cultures and Urs Gerber.

REFERENCES

[1] C. A. Mead, *Analog VLSI and Neural Systems*, 1st ed. Reading, MA, USA: Addison-Wesley, 1989.
[2] S. C. Liu and T. Delbruck, "Neuromorphic sensory systems," *Curr. Opin. Neurobiol.*, vol. 20, no. 3, pp. 288–295, Jun. 2010.
[3] S.-C. Liu, T. Delbruck, G. Indiveri, A. Whatley, and R. Douglas, Eds., *Event-Based Neuromorphic Systems*. Oxford, U.K.: Wiley, 2015.
[4] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μs latency asynchronous temporal contrast vision sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, Feb. 2008.
[5] K. A. Boahen, "Communicating neuronal ensembles between neuromorphic chips," *Neuromorphic Syst. Eng.*, vol. 447, pp. 229—259, 1998.
[6] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A 240×180 130 dB 3 us latency global shutter spatiotemporal vision sensor," *IEEE J. Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, Oct. 2014.
[7] C. Li *et al.*, "Design of an RGBW color VGA rolling and global shutter dynamic and active-pixel vision sensor," in *Proc. 2015 IEEE Int. Symp. Circuits Syst.*, 2015, pp. 718–721.
[8] C. Posch, D. Matolin, and R. Wohlgenannt, "An asynchronous time-based image sensor," in *Proc. 2008 IEEE Int. Symp. Circuits Syst.*, 2008, pp. 2130–2133.
[9] D. Matolin, C. Posch, R. Wohlgenannt, and T. Maier, "A 64 × 64 pixel temporal contrast microbolometer infrared sensor," in *Proc. IEEE 2008 Int. Symp. Circuits Syst.*, 2008, pp. 1644–1647.
[10] T.-W. Chen *et al.*, "Ultra-sensitive fluorescent proteins for imaging neuronal activity," *Nature*, vol. 499, no. 7458, pp. 295–300, Jul. 2013.
[11] A. Rose, *Vision: Human and Electronic*. New York, NY, USA: Plenum, 1973.
[12] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. New York, NY, USA: Springer, 2007.
[13] L. Kastrup and S. W. Hell, "Absolute optical cross section of individual fluorescent molecules," *Angew. Chem. Int. Ed.*, vol. 43, no. 48, pp. 6646–6649, Dec. 2004.
[14] C. Posch, D. Matolin, and R. Wohlgenannt, "A two-stage capacitive-feedback differencing amplifier for temporal contrast IR sensors," *Analog Integr. Circuits Signal Process.*, vol. 64, no. 1, pp. 45–54, Jul. 2010.
[15] T. Delbruck and R. Berner, "Temporal contrast AER pixel with 0.3%-contrast event threshold," in *Proc. 2010 IEEE Int. Symp. Circuits Syst.*, 2010, pp. 2442–2445.
[16] T. Serrano-Gotarredona and B. Linares-Barranco, "A 128 × 128 1.5% contrast sensitivity 0.9% FPN 3 μs latency 4 mW asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers," *IEEE J. Solid-State Circuits*, vol. 48, no. 3, pp. 827–838, Mar. 2013.
[17] M. Yang, S. C. Liu, and T. Delbruck, "A dynamic vision sensor with 1% temporal contrast sensitivity and in-pixel asynchronous delta modulator for event encoding," *IEEE J. Solid-State Circuits*, vol. 50, no. 9, pp. 2149–2160, Sep. 2015.
[18] T. Serrano-Gotarredona, B. Linares-Barranco, and A. G. Andreou, "Very wide range tunable CMOS/bipolar current mirrors with voltage clamped input," *IEEE Trans. Circuits Syst. Fundam. Theory Appl.*, vol. 46, no. 11, pp. 1398–1407, Nov. 1999.
[19] T. Delbrück, R. Berner, P. Lichtsteiner, and C. Dualibe, "32-bit configurable bias current generator with sub-off-current capability," in *Proc.2010 IEEE Int. Symp. Circuits Syst.*, 2010, pp. 1647–1650.
[20] D. P. Moeys *et al.*, "Color temporal contrast sensitivity in dynamic vision sensors," presented at the 2017 IEEE Int. Symp. Circuits Syst., Baltimore, MD, USA, 2017.
[21] R. Berner, P. Lichtsteiner, and T. Delbruck, "Self-timed vertacolor dichromatic vision sensor for low power pattern detection," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2008, pp. 1032–1035.
[22] J. A. Lenero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, "A 3.6 μs latency asynchronous frame-free event-driven dynamic-vision-sensor," *IEEE J. Solid-State Circuits*, vol. 46, no. 6, pp. 1443–1455, Jun. 2011.
[23] EMVA 1228 standard for characterization of image sensors and cameras release 3.0, *European Machine Vision Association*, 2010.
[24] D. P. Moeys. SDAVIS and DAVIS240C Comparison. [Online]. Available: https://www.youtube.com/watch?v=QsYeSguI0b8. Accessed Nov. 27, 2016.
[25] X. Fang *et al.*, "Fluorescence detection and imaging of cytosolic calcium oscillations: A comparison of four equipment setups," *Prog. Nat. Sci.*, vol. 19, no. 4, pp. 479–487, Apr. 2009.
[26] H. T. Beier and B. L. Ibey, "Experimental comparison of the high-speed imaging performance of an EM-CCD and sCMOS camera in a dynamic live-cell imaging test case," *PLoS One*, vol. 9, no. 1, p. e84614, Jan. 2014.
[27] Y. Nozaki and T. Delbruck, "Temperature and parasitic photocurrent effects in dynamic vision sensors," *IEEE Trans. Electron Devices*, vol. 64, no. 8, pp. 3239–3245, Aug. 2017.
[28] B. Son *et al.*, "4.1 A 640×480 dynamic vision sensor with a 9 μm pixel and 300Meps address-event representation," presented at the 2017 Int. Solid State Circuits Conf. Tech. Dig., San Francisco, CA, USA, Paper 4.1.
[29] Y. Gong *et al.*, "High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor," *Science*, vol. 350, no. 6266, pp. 1361–1366, Dec. 2015.



Diederik Paul Moeys (S'12) received the M.Eng. degree from University College London, London, U.K., in 2013, and the Ph.D. degree in electronic engineering from ETH Zürich, Zürich, Switzerland, in 2016. His research interests include chip design to event-based data processing and tracking using machine learning. He received the Goldsmid Faculty Medal in 2013, the IET Prize, and the IET Postgraduate Scholarship in 2016.

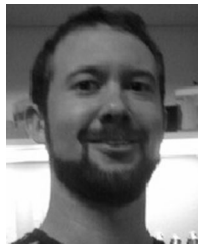
Federico Corradi (M'13) received the B.Sc. degree in physics from Università degli studi di Parma, Parma, Italy, the M.Sc. degree *(cum laude)* in physics from La Sapienza University, Rome, Italy, and the Ph.D. degrees in natural sciences from the University of Zürich, Zürich, Switzerland, and in neuroscience from Neuroscience Center Zürich, Zürich. He is currently an R&D Engineer with iniLabs GmbH, Zürich.

Chenghan Li received the B.Eng. degree in electronic engineering from Nanyang Technological University, Singapore, in 2010, and the M.Sc. degree in neural systems and computation from ETH Zürich, Zürich, Switzerland, in 2012. He is currently working toward the Ph.D. degree at the Institute of Neuroinformatics, University of Zurich and ETH Zürich, Zürich. His research includes neuroinspired two-stream vision sensors and visual processing.

Simeon A. Bamford received the Ph.D. degree in neuromorphic engineering from the University of Edinburgh, Edinburgh, U.K. He was a Postdoctoral Fellow with Istituto Superiore di Sanità, Rome, Italy, working on neuromimetic prosthesis and neuromorphic network dynamics simulation. Since 2013, he has been with iniLabs GmbH, Zürich, Switzerland, and he is currently the CTO of iniVation, Zürich.

Luca Longinotti received the B.Sc. degree in computer science from the University of Zürich, Zürich, Switzerland. He is currently with iniLabs GmbH, Zürich, as an R&D Software Engineer, where he is responsible for FPGA logic, firmware, and low-level software development. His main research interests include low-level and embedded software systems.

Fabian F. Voigt received the B.Sc. and M.Sc. degrees in interdisciplinary science from ETH Zürich, Zürich, Switzerland. He is currently working toward the Ph.D. degree at the Group of Fritjof Helmchen, Brain Research Institute, University of Zürich, Zürich. His research interests include the design of novel microscopes for neuroscience.



Stewart Berry received the B.Sc. degree in biology from Colorado State University, Fort Collins, CO, USA, and the M.Sc. degree in neuroscience from the University of Zürich, Zürich, Switzerland, where he is currently working toward the Ph.D. degree at the Brain Research Institute.

Gemma Taverni received the B.Sc. degree and the M.Sc. degree *(cum laude)* in biomedical engineering from the Università di Pisa, Pisa, Italy. She is currently working toward the Ph.D. degree at the University of Zürich, Zürich, Switzerland. She received the GNB 2014 Award for the best Bioengineering Master Thesis.

Fritjof Helmchen received the Ph.D. degree in neuroscience in 1996. He was with the Max-Planck Institute for Medical Research, Heidelberg, Germany, for many years. He was a Postdoctoral Fellow with Bell Labs, Lucent Technologies, Murray Hill, NJ, USA, for three years. Since 2005, he has been a Professor of neuroscience, the Codirector of the Brain Research Institute, and currently the Director of the Neuroscience Center, University of Zurich, Zürich, Switzerland.

Tobi Delbruck (M'99–SM'06–F'13) received the B.Sc. degree in physics from the University of California, Berkeley, CA, USA, in 1986, and the Ph.D. degree from California Institute of Technology, Pasadena, CA, in 1993. He is currently a Professor of physics with the Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zürich, Switzerland, where he has been since 1998. His group focuses on neuromorphic sensory processing.