A Bio-Inspired Vision Sensor With Dual Operation and Readout Modes

Juan A. Leñero-Bardallo, Philipp Häfliger, *Senior Member, IEEE*, R. Carmona-Galán, *Member, IEEE*, and Ángel Rodríguez-Vázquez, *Fellow, IEEE*

Abstract—This paper presents a novel event-based sensor with two operation modes: 1) intensity mode and spatial contrast detection. They can be combined with two different readout approaches: 1) pulse density modulation and time-to-first spike. The sensor is conceived to be a node of an smart camera network made up of several independent an autonomous nodes that send information to a central one. The user can toggle the operation and the readout modes with two control bits. The sensor has low latency (below 1 ms under average illumination conditions), low power consumption (19 mA), and reduced data flow, when detecting spatial contrast. A new approach to compute the spatial contrast based on inter-pixel event communication less prone to mismatch effects than diffusive networks is proposed. The sensor was fabricated in the standard AMS4M2P 0.35-μm process. A detailed system-level description and experimental results are provided.

Index Terms—Bio-inspired vision, AER (address event representation), spatial contrast, asynchronous vision sensors.

I. INTRODUCTION

E VENT-BASED vision sensors differ from conventional frame-based image sensors in that their outputs are not encoded representations of image intensity maps. Instead, these sensors deliver reduced sets of data signaling the presence of spatio-temporal contrast in incoming scenes. These data can be encoded in different ways and are typically conveyed by spikes –similarly to biological vision systems [1]. Encoding techniques include, among others, Pulse Density Modulation (PDM) [2], [3] and Time-to-First Spike (TFS) [4], [5], which are just the two employed in the chip reported in this paper.

The rationale for event-based vision sensors is that they preclude transmitting and manipulating huge amounts of irrelevant data which are typically encountered in images. Event-based sensors are information-centric entities while conventional image sensors are data-centric ones. Event-based sensors extract relevant information from visual scenes right at the point where these scenes are captured (at the focal plane), discard irrelevant data, such as for instance pixel values corresponding to uniform backgrounds, and transmit only a reduced amount of data with the relevant information needed to identify people, objects, shapes, etc. [6], [7]. When embedded into neuromorphic systems, their usage report significant advantages in terms of system size, speed, and energy budget [8]–[10]. These features are extremely important for applications like vision-enabled wireless vision sensors [11], [12]. In these applications, a large number of interacting sensor nodes is deployed for distributed monitoring of large areas under the control of a central processing node. Each node monitors a small region and is responsible of detecting events happening within its corresponding region. Information is shared locally among nodes themselves and with a central processing unit. Sensor ability to extract information in autonomous manner, with low-latency (i.e., at high-speed) and with reduced power consumption are cornerstone in these application scenarios. Furthermore, event outputs can be easily conveyed into radio signals using an UWB modulator [12], thus facilitating inter-node communications.

Following the detection of suspicious events, nodes may be required to upload complete images for further analysis. For instance, the Viola-Jones algorithm [13] can be used to detect people faces [14]. To that purpose, event sensors should also be capable of uploading complete images upon demand. Regarding this, the question arises why not using conventional image sensors, which naturally produce complete images, and employ embedded digital processors for extracting events. Although there is not a closed answer to this question, empirical evidences tell us that using event-based sensors are more efficient in terms of power consumption and speed [15]–[17]. This is not surprising since scenes acquired by nodes of wireless sensors do not contain relevant for information during most of the time. Hence, downloading, transmitting and processing irrelevant data, as it would happen whether conventional image sensors were employed, represents a waste of precious resources. Particularly data transmission takes most of the power; for instance, for the wireless sensor reported by Chen et al. [12], power consumption is 0.9mW for the sensor and 15mW for radio transmission. There are specific network communication standards like ZigBee (www.zigbee.org), aimed to reduce the power consumption in low data rate communication networks. Dedicated studies to

J. A. Leñero-Bardallo, R. Carmona-Galán, and Á. Rodríguez-Vázquez are with the Institute of Microelectronics of Seville, CSIC-Universidad de Sevilla, Sevilla 41004, Spain (e-mail: juanle@imse-cnm.csic.es; rcarmona@imse-cnm.csic.es; angel@imse-cnm.csic.es).

P. Häfliger is with the Nanoelectronics Group, Department of Informatics, University of Oslo, Blindern 1072, Norway (e-mail: hafliger@ifi.uio.no).
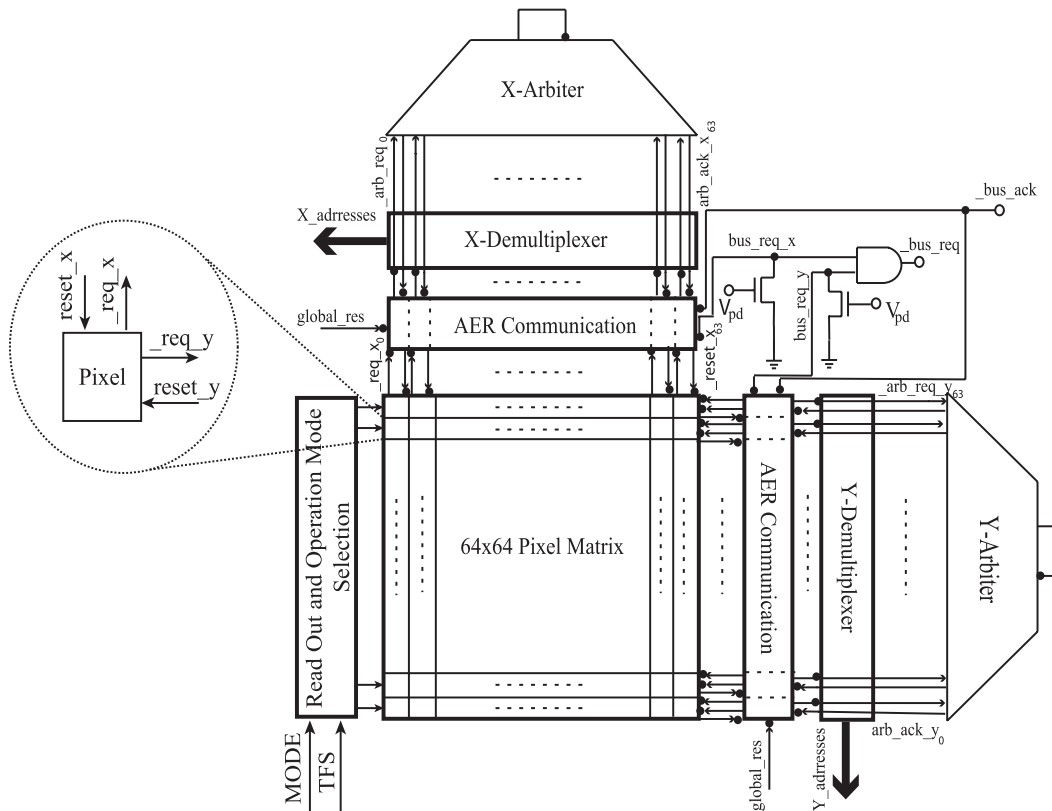
Fig. 1. System block diagram. Center: Pixel matrix. Left: Circuitry for selection of the sensor's readout and operation modes. Top and right: Asynchronous arbitration circuitry. Row and column petitions are arbitrated with the same circuitry: asynchronous communication logic, demultiplexers, and arbiters.

minimize the power consumption when transmitting images wirelessly have also been published [18]. Some authors propose on-chip image processing to avoid raw data over wireless channels at the expense of area consumption and system complexity, [19]–[21].

This paper reports an event-based sensor aimed for surveillance and inspection applications that compute spatial contrast and also provide intensity images. Two operation modes, namely:

- Intensity Mode (IM)
- Contrast Detection (CD)

and two readout modes:

- Pulse Density Modulation (PDM)
- Time-to-First Spike (TFS)

are available. Two bits are used to reconfigure device operation and readout modes without switching it off.

The different operation/readout modes involve tradeoffs between image quality, speed, bandwidth, and power consumption. Thus, the intensity mode provides better image quality and lower latency at the expense of a larger power and bandwidth consumption. On the other hand, the contrast mode delivers compressed images, with significant energy and bandwidth consumption saving. Regarding the two readout modes, the first one (PDM) provides pixels spikes with an output frequency proportional to light intensity or spatial contrast, depending on the operation mode.

On the other hand, with the second readout mode (TFS), pixels spike, at the most, one time after a global reset.

In this case, light intensity or spatial contrast are inversely proportional the pixel spiking latency after the global reset. In these readout modes, image quality (PDM) can be traded for a reduced output data flow (TFS).

The sensor chip reported in this paper uses an improved approach to compute the spatial contrast based on inter-pixel communication with spikes. The pixel itself has been proposed at the outcome of a previous research work conducted in the design of pixels capable of detecting contrast without being calibrated [22], [23]. These previous papers reported only an isolated pixel, while this one reports the design and characterization of a 64×64 focal plane array with all supporting circuits to make it a complete vision sensor system. The contrast computation method employed is less prone to mismatch and crosstalk than prior ones based on diffusive networks. The reasons are that pixel transistors do not operate in subthreshold region and that neighbor pixels exchange digital signals to compute spatial contrast. Diffusive networks are quite dependent on photocurrents diffusive network components to compute the spatial contrast and employ transistors operating in subthreshold region [5], [24]–[27]. The sensor does not need to be calibrated [5], [27]. There are prior event-based vision sensors that can provide intensity images [3], [28], [29] (also known as octopus retinas) and devices that can compute spatio-temporal contrast [5], [24]–[26]. However, to the best of our knowledge, devices that can be reconfigured in-line to switch between these operation modes have not been reported yet.
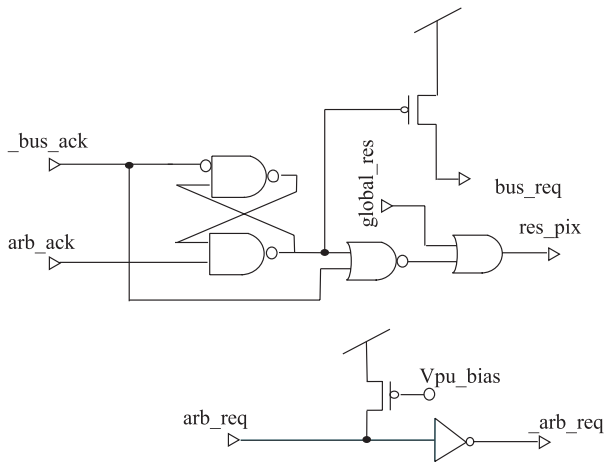
Fig. 2. AER communication circuitry. Each row and each column is connected to one of this elements as is depicted in Fig. 1.

The sensor was fabricated in the AMS4M2P 0.35 μm standard process. This article provides a insightful description of the main sensor blocks and detailed experimental results, analyzing the sensor performance in different environments. The article is organized as follows: Section II explains the main sensor constitutive blocks and the interconnections between them. The arbitration system is described there. Section III describes the contrast computation approach proposed in this paper. The pixel implementation and its operation are described there. Section IV depicts how the different sensor readout and operation modes can be selected. Section V presents the experimental results conducted to characterize the sensor and benchmarks it against the state-of-the-art. Finally Section VI draws some conclusions and describes further work.

II. SYSTEM DESCRIPTION

Fig. 1 displays the system block diagram. The main system blocks and their interconnections are indicated. The sensor is made of a pixel matrix, a block dedicated to control the sensor operation and readout modes, and the Address Event Representation (AER) arbitration system. The AER logic and the arbitration system is reported elsewhere [30]. Each pixel can be considered as an asynchronous block that generates events. Such events are handled by the arbitration system that arbiters the pixel petitions at the row and column levels (*_req_y* and *_req_x* signals in Fig. 1). Firstly, row pixel petitions are arbitered. When a pixel receives the acknowledge signal *reset_y*, initiates a column petition enabling the signal *_req_x*. Finally, when the pixel receives the signal *reset_x*, its address has already been sent through the AER bus and the request signals are released.

Fig. 2 displays the schematics of one AER communication logic cell. This circuitry is dedicated to receive and send the request and acknowledge signals of each row/column. It also generates the bus request signal (*_bus_req*) to send a new event out the chip. Whenever it receives the off-chip signal *_bus_ack*, it releases the pixels petitions.

Fig. 3 displays the arbitration logic and how it is arranged. We use two arbiters to arbitrate the row and columns petitions
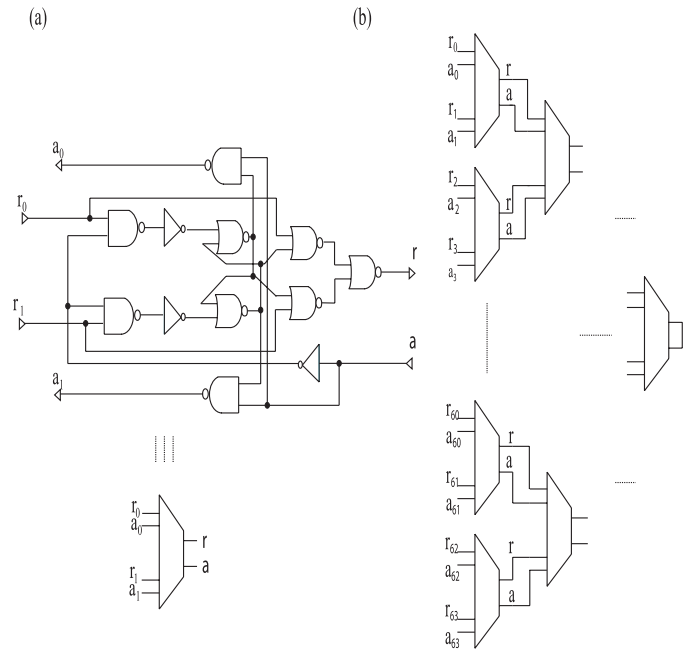

Fig. 3. (a) Greedy arbiter schematics and its symbol. (b) 64-input arbitration tree.

as is shown in Fig. 3. An arbiter is a circuit that avoids collisions between simultaneously requests. If it receives simultaneously two request signals from two pixels located in different rows/columns, only one will be attended. The circuitry of one arbiter is displayed in Fig. 3.(a). Arbiters are arranged as is depicted in Fig. 3.(b) with a tree shape.

Finally, we display in Fig. 4 the detailed connectivity of the AER blocks depicted in Fig. 1. Note that the AER blocks are replicated twice. One of them arbiters the rows requests and another one is dedicated to the columns requests. In between the AER communication logic and the arbitration tree, there is a demultiplexer. Its purpose is to encode the X- and Y-addresses of the pixels that have fired.

The arbitration system can handle event rates up to 10Meps (where *eps* means events per second) for pixels of the same row, and up to 2Meps for pixels of different rows, [30]. Such event rates are fast enough for low-medium resolution arrays, like the one reported in this article (64×64 pixels). This arbitration scheme adds also some pipeline operation. When the *_bus_ack* signal is received, the pixel who fired is reset immediately, allowing pixels to elicit row petitions before the *_bus_ack* signal is disabled. The external signal *_bus_req* triggers an off-chip datalogger to save the AER bus address. Once the pixel which fired coordinates are stored, the datalogger sends back to the chip the signal *_bus_ack*.

As it will be shown in Section V-E, the proposed arbitration system is fast enough to operate under all the illumination conditions where the sensor was tested. However, if we scaled arbitrarily the number of pixels and we exceeded such limit, pixels requests would not be attended instantaneously and the event timing would be corrupted, i.e. pixels would need to wait an arbitrary amount of time until their petitions are attended. To amend this limitation, there are faster arbitration schemes like the one proposed by Boahen et al. that can
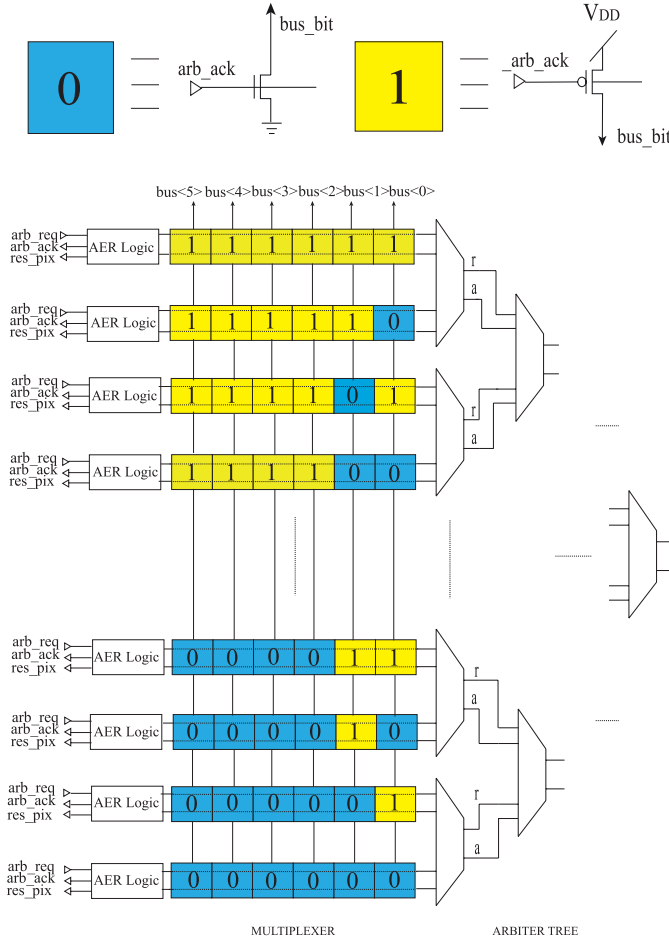
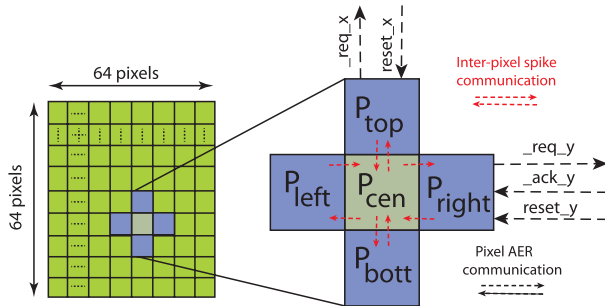Fig. 4. Detailed connectivity of the AER blocks depicted in Fig. 1.

Fig. 5. Diagram displaying the inter-pixel communication to compute the spatial contrast and pixel communication signals to handle the AER protocol.

reach 20Meps [31]. Another solution, would be to divide the pixel array into two independent blocks and use two arbitration systems operating in parallel, i.e. the maximum event rate throughout would be doubled.

III. CONTRAST COMPUTATION AND PIXEL OPERATION

The approach to compute the spatial contrast is new. It is based on inter-pixel communication (see Fig. 5). Pixels send pulses to their neighbors (North, South, West, and East) with a frequency that is proportional to the light intensity they receive. There is a competition between each pixel and its neighbors. If one pixel has higher illumination than

its neighbors, that pixel will be exposed to spatial contrast, and enabled to send out events with average frequency ($f_{i,j}$), proportional to the difference between its local photocurrent ($I_{i,j}$) and the average photocurrent within its neighborhood (I_{avg}), i.e.:

$$f_{i,j} \propto max\left(0, I_{i,j} - \left(I_{i+1,j} + I_{i-1,j} + I_{i,j+1} + I_{i,j-1}\right) \cdot k\right)$$
$$= max\left(0, I_{i,j} - I_{avg}\right) \quad (1)$$

Where k is a constant that can be tuned, depending on the desired neighbor influence in the contrast computation. The lower k is the higher is the influence of the neighborhood in the contrast computation, i.e. the central pixel should be exposed to a higher gradient illumination to detect contrast. If $k = 4$, $f_{i,j} = max\left(0, I_{i,j} - I_{avg}\right)$. This computation is equivalent to first-order Laplacian operator. If we compare this approach to the traditional one with diffusive networks [5], [24]–[27], there are a few advantages: Pixel transistors operate in strong inversion. Signals exchanged between adjacent pixels are digital and less prone to crosstalk. The implementation of a diffusive network (resistive two-dimensional grids) requires area. Usually their transistors operate in subthreshold region being more prone to mismatch and crosstalk. Inter-pixel communication with spikes models more accurately some kinds of neuron communication in biological systems. The limitation of our approach is that we do not have the possibility of tuning the number of neighbors involved in the contrast computation. In our design, the pixel neighborhood could be extended, adding one transistor per extra neighbor. In this implementation, as a proof-of-concept 4-pixel neighborhood was considered to compute the spatial contrast. As it will be discussed in Section V, there is a tradeoff between latency and power and bandwidth consumption.

Fig. 6 displays pixel schematics. Each pixel has an integrated-and-fire neuron that behaves as an oscillator generating pulses with a frequency that is proportional to its local photocurrent, $I_{i,j} = I_{ph}$. The oscillation frequency is given by:

$$f_{i,j} = \left[\frac{(V_{DD} - V_{ref})}{I_{i,j}} \cdot C_1 + \Delta T_{active}\right]^{-1}$$
$$\approx \frac{I_{i,j}}{C_1 \cdot (V_{DD} - V_{ref_1})} \quad (2)$$

Where V_{DD} is the reset voltage and V_{ref_1} is the comparator threshold voltage, that can be adjusted, depending on the desired sensor speed response. ΔT_{active} is the amount of time that the integrate-and-fire neuron is active. ΔT_{active} is determined by the sum of the delays introduced by the integrate-and-fire comparator and output digital buffer. The integrate-and-fire neuron outputs are connected to the neighboring pixels (N_{north}, N_{south}, N_{west}, and N_{east}) as is depicted in Fig. 6. Every pulse generated by the integrate-and-fire neuron, provokes a voltage increment of the capacitor C_2 and every pulse received from the neighbors elicits a voltage decrement at C_2:

$$\Delta V_{C2} = \pm \frac{I_{\{+,-\}}}{C_2 \cdot \Delta T_{active}} \quad (3)$$
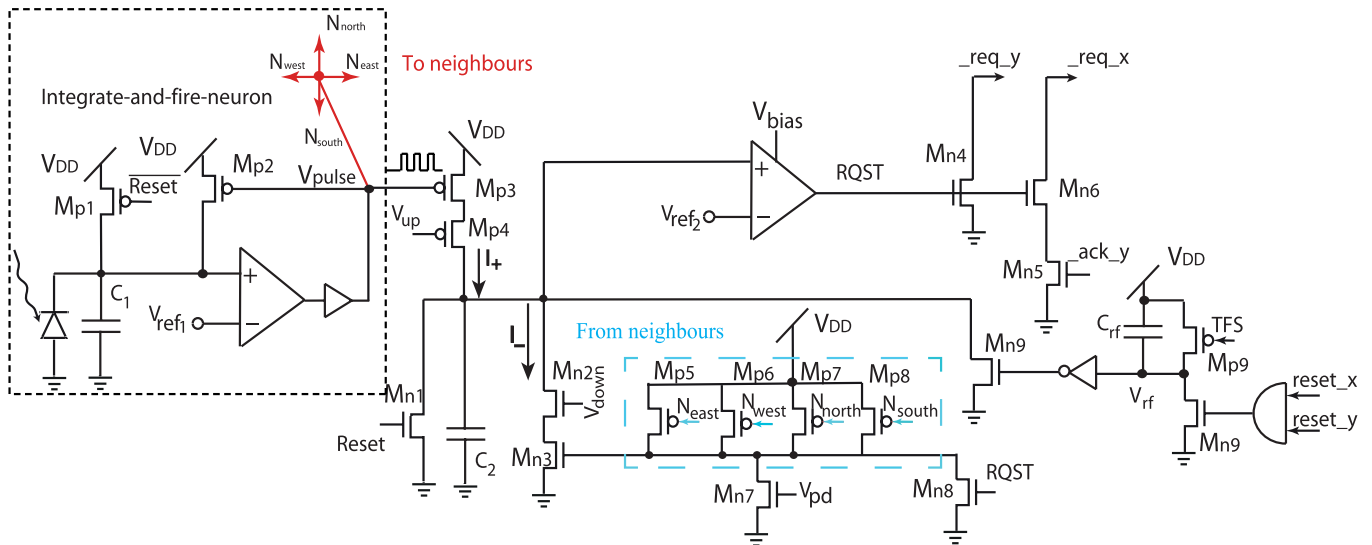
Fig. 6. Pixel schematics. Connections from/to neighbors have been highlighted in red. Transistor sizes are (W/L) $\mu m/\mu m$ $M_{p1} = 0.5/1$, $M_{p2} = M_{p3} = M_{p4} = 4/1$, $M_{p5} = M_{p6} = M_{p7} = M_{p8} = 1/2$, $M_{p9} = 0.5/1$, $M_{n1} = 1/0.35$, $M_{n2} = M_{n3} = 1/3$, $M_{n4} = 1/0.55$, $M_{n5} = M_{n6} = 0.6/0.55$, $M_{n7} = M_{n8} = 1/3$, $M_{n9} = 3/1$, and capacitance values are $C_1 = C_2 = 20fF$, $C_{rf} = 4fF$.

The currents $I_{\{+,-\}}$ can be adjusted tuning the voltages V_{up} and V_{down}, i.e.:

$$I_+ = \frac{V_{DD}}{R\left(V_{up}, V_{C_2}\right)} \qquad I_- = \frac{V_{C_2}}{R\left(V_{down}, V_{C_2}\right)} \qquad (4)$$

V_{up} controls the influence of the central pixel in the contrast computation. V_{down} controls the influence of neighbourhood in the contrast computation (k parameter of Eq. 1. Thus, there is a competition among the central pixel and its neighbors. To illustrate this, Fig. 7 depicts how the voltage at C_2 changes depending on the input pixel photocurrents of the pixel ($I_{i,j}$) and one of its neighborgs. If the central pixel is exposed to higher illumination than I_{avg}, the voltage at C_2 will rise up every time that the voltage V_{pulse} is active and an event will be sent out of the chip when the voltage threshold of the second comparator V_{ref_2} is reached. If there is no spatial contrast (pixel illumination is similar to the average illumination in the neighborhood), the voltage at C_2 will never reach V_{ref_2}. That is the principle of operation computing spatial contrast, equivalent to the Laplacian operator described by Eq. 1.

If we disable the influence of the neighborhood in the contrast computation and maximize the voltage increments at C_2 (by setting $V_{up} = V_{down} = 0V$), every pulse of the integrate-and-fire neuron can be conveyed into an output event. Hence, pixel outputs frequencies will be given by Eq. 2 and will be proportional to the illumination received by each pixel. As a result, the sensor will be operate as an octopus retina [3], [28], [29].

On the right of Fig. 6, we find the in-pixel digital circuitry to handle the AER communication with the peripheral blocks depicted in Fig. 4.

Finally, the two different readout modes can be set with the digital switch TFS. If $TFS = 0$, the voltage V_{rf} will rise up immediately after the pixel is reset by the periphery ($reset\_x = reset\_y = 1$). Transistor M_{p9} will operate as a pull-up transistor. Thus, pixels can spike as many times
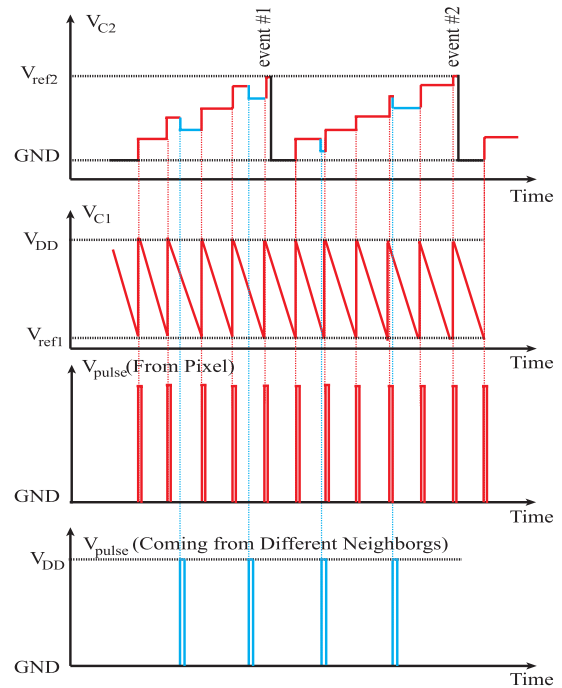

Fig. 7. Illustration of the contrast computation principle with inter-pixel communication. The transient voltages V_{C_1} and V_{pulse} inside one pixel (see Fig. 6) are displayed in red. We also display in blue the voltage V_{pulse} coming form its neighbouring pixels. The effect of the pixel and its neighbors on V_{C_2} is shown. The number of positive increments is proportional to the pixel photocurrent. The number of voltage decrements is proportional to the neighbors photocurrents. In this example, the neighbors' average photocurrent is lower than the pixel photocurrent. As a result, events are elicit when the voltage V_{C_2} reaches V_{ref_2}.

as necessary. This is the Pulse Density Mode (PDM) readout. If $TFS = 1$, after resetting the sensor with the signal *Reset*, the voltage V_{rf} will be high. The first time that the pixel is reset by the arbitration system V_{rf} will be pulled down by transistor M_{n_9}. After that, the voltage of V_{rf} will rise very
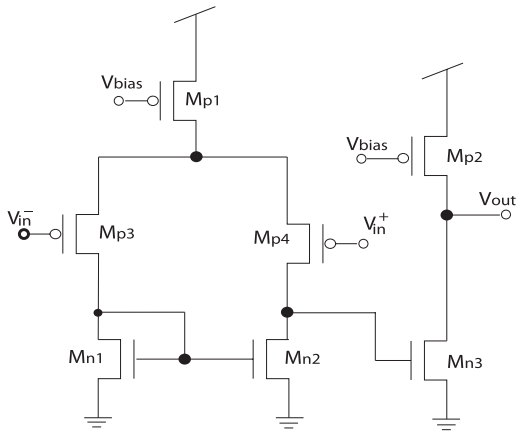
Fig. 8. Comparator schematics. Transistor sizes are: ($\mu m/\mu m$): $M_{n1} = M_{n2} = 2/0.5$, $M_{p1} = M_{p2} = 2/4$, $M_{p3} = M_{p4} = 2/0.35$, $M_{n3} = 6/0.35$.
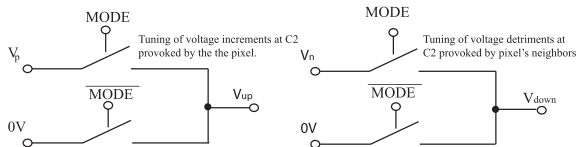

Fig. 9. Set of analog multiplexors to set the pixel voltages for the intensity and the contrast detection modes.

TABLE I
OPERATION MODES AND READOUT METHODS TRUTH TABLE

| Mode and readout method	MODE	TFS
Contrast Detection + PDM	1	0
Contrast Detection + TFS	1	1
Intensity+PDM	0	0
Intensity+TFS	0	1

Fig. 10. Experimental setup displaying the vision sensor, lens mount, and the USBAERmini data logger. The elements have been labeled in the picture.

slowly because transistors M_{p_9} and M_{n_9} are switched-off. The pixel will remain reset in the meantime. In practice, pixels will be allowed to spike at the most once. That is the Time-to-first-spike (TFS) readout.

Fig. 8 displays the comparator schematics. A two-stage comparator was selected for the design. For this chip implementation, we desired to control the threshold voltages V_{ref_1} and V_{ref_2}. To make the integrate-and-fire neuron circuit stable, a controlled delay must be introduced between the comparator output and the self-resetting signal that closes the loop. By adding a second-stage to the comparator, we make the delay bigger, stabilizing the circuit operation. Under nominal operation, the analog power consumption of each comparator was set to be lower than 150nA. The comparator has been previously implemented in prior designs offering good performance [5]. To mitigate mismatch, we have paid special care selecting the transistors sizes of the differential pair. The comparator has an estimated DC offset voltage of approximately $V_{os} = 15$mV. The input voltage has to exceed the threshold voltage (V_{ref_1} or V_{ref_2}) plus V_{os} to elicit one pulse. This offset voltage does not affect significantly the pixel operation because usually V_{ref_1} and V_{ref_2} values are much higher than V_{os}.

IV. OPERATION AND READOUT MODES

The sensor has been designed to allow the operator to toggle among operation and readout modes using two control bits. The binary input *MODE* sets the operation mode. The binary input *TFS* sets the readout mode. The sensor has not to be reset when toggling between operation modes. Analog multiplexors have been added to the chip (see Fig. 9) to set the values of the voltages V_{up} and V_{down}, depending on the circuit operation mode. In intensity mode, every pulse generated by
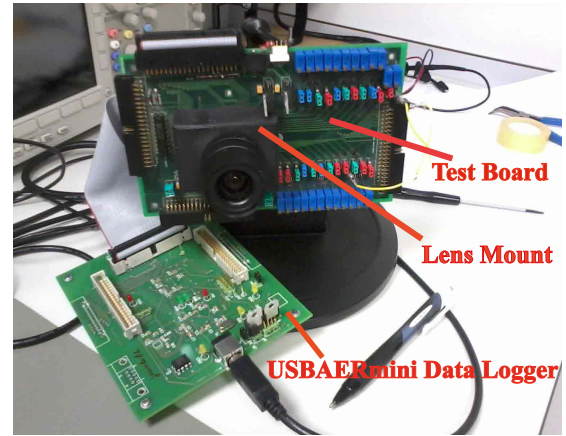
the integrate-and-fire neuron is directly conveyed into and output event. When *MODE = 0*, we maximize the voltage increments provoked by the pixel setting $V_{up} = 0V$. Thus, every voltage increment when the integrate-and-fire neuron fires will reach the voltage V_{ref_2} and an event will be sent. Neighbors influence is also disabled by setting $V_{down} = 0V$. In the contrast detection mode (*MODE = 1*), the voltages V_{up} and V_{down} can be tuned by the user, depending on the desired pixel/neighbors influence in the contrast computation. Under normal operating conditions, we set their values to $V_{up} = 2.4$V and $V_{down} = 0.53$V. Table I summarizes the truth table of the binary inputs *MODE* and *TFS* for the different operation and readout modes.

V. EXPERIMENTAL RESULTS

A. Experimental Setup

Fig. 10 displays the experimental setup. A custom PCB and a dedicated lens mount were fabricated to host the sensor. The board USBAERmini [32] was used as a data logger to send the sensor outputs to a PC and render real time images. The sensor was fabricated in the standard CMOS AMS4M2P 0.35 μm process. The sensor was tested with a 4mm $F = 1 : 2$ CS-mount lens. Fig. 11 shows a chip photo including pads and the pixel layout. In order to avoid crosstalk between neighbouring pixels, capacitor C_2 and the more sensitive pixel nodes have been placed in one specific region that is not crossed by digital communication signals. Chip dimensions are 3.6mm×3.8mm. Main sensor specifications are summarized on Table II.

B. Sensor Performance

A dedicated Java interface was programmed to display images on a PC, [33]. Fig. 12 shows the interface displaying
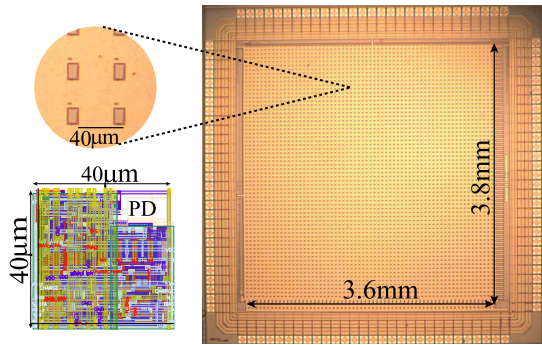
Fig. 11. Chip micro photograph and pixel layout. Chip dimensions, including pads, are 3.6 mm × 3.8 mm. Pixels size is 40 μm × 40 μm.

TABLE II
SENSOR SPECIFICATIONS

Technology	AMS4M2P 0.35μm
Power Supply	3.3V
Resolution	64×64
Pixel Size	40μm×40μm
Pixel Complexity	49 Transistors + 3 Capacitors
Fill Factor	8.5%
Power Consumption	1.85mA@0.1Meps
FPN	3.3% (CD) / 2.4% (IM)
Latency	400μs @ 1.2klux
Dynamic Range	>90dB
Contrast Sensitivity	7000Hz/WC

Fig. 12. Custom jAER interface rendering real-time images. We show snapshots of the same scene with the sensor operating in intensity (top) and contrast detection (bottom) modes.

snapshots of the same scene in intensity and contrast detection modes. Fig. 13 displays several snapshots of static images taken in both operation modes in indoor environments with
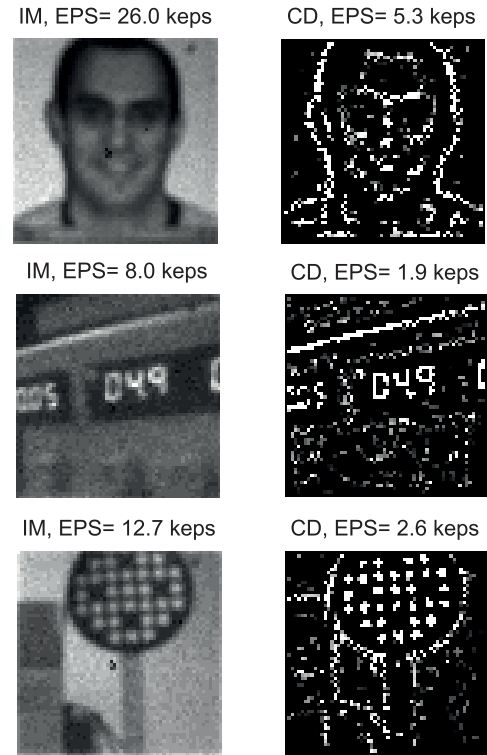

Fig. 13. Snapshots taken in the Intensity Mode and the Contrast Detection modes. Event rates in each operation mode are indicated.

scene illuminations ranging between 300-500lux and PDM readout. The output data flow is lower operating in contrast detection mode. Typically, the event rate is 4-6 times lower in contrast detection Mode and PDM readout, depending on the visual scene. The contrast detection mode allows to use the vision sensor to recognize objects or people in the visual scene. The intensity mode can be used to obtain more detailed and quality images.

C. Power Consumption

Power consumption mainly depends on the event rate. Fig. 14 displays the dependency between the output event rate and the chip power consumption, including pads. Operating in indoor environments with average illumination conditions between 300-600lux, the power consumption is always below 2mA in all the operation modes. It comes mainly from the analog circuitry and does not change significantly with the event rate. In average, consumption is below 2mA in indoor environments. For event rates above 300keps, the digital power consumption becomes dominant and there is a linear dependence between the event rate and the power consumption. Linear data fitting for event rates above 300keps has been added to the plot. There is approximately an increment of 0.13mA for every event even rate increment of 100keps. The analog power consumption mainly comes from the pixels comparators. The digital power consumption comes from the arbitration system and the periphery, that is very active for high event rates. Examining the graph of Fig. 14, we observe that the contrast detection mode will lead to save power by reducing the event rate. Note that in wireless sensor network the main power consumption comes from the data
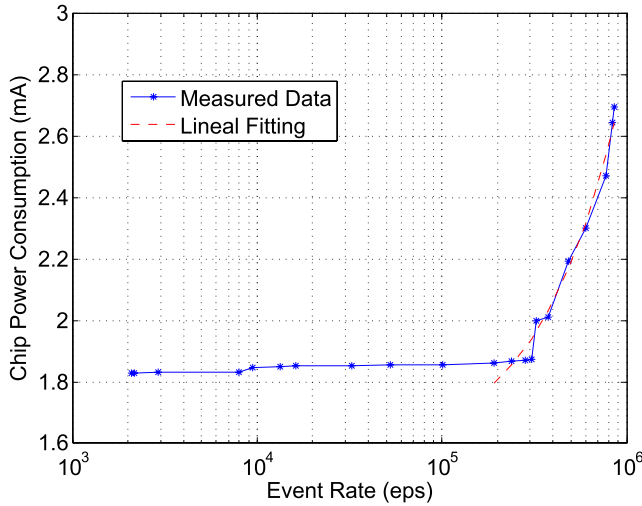
Fig. 14. Chip power consumption versus event rate.

transmission [12]. Hence, it is crucial to reduce the output data flow.

D. Fixed Pattern Noise

The sensor was conceived to compute spatial contrast without being calibrated, avoiding the traditional contrast computation approach with diffusive networks [5], [27]. With the new approach, we avoid calibration by using inter-pixel communication with spikes. Only pixels exposed to higher illumination than their neighbors can spike. Only one transistor per neighbor is necessary to take it into account in the contrast computation. In our implementation, a 4-pixel neighborhood is considered. Pixel neighbors placed in the North, South, East, and West have influence in the contrast computation. To compute the fixed pattern noise, we illuminated all the pixel array with uniform and constant light. Events send out by the pixels where recorded during a few seconds. We define the Fixed Pattern Noise (FPN) for event-based vision sensors as the ratio between the standard deviation of the output frequencies normalized by the average output frequency, i.e.:

$$FPN = \frac{\sum_{i=1}^{N} \sqrt{\left(f_i - \bar{f}\right)^2}}{N \cdot \bar{f}} = \frac{\sigma_f}{\bar{f}} \quad (5)$$

Operating in contrast computation mode, with an average chip illumination of 100lux, we gauged the FPN. The sensor optics was removed and a diffuser was placed over it to obtain uniform illumination. The measured FPN was 3.2%. FPN is mainly provoked by mismatch of the voltage increments/decrements at C_2. The value is higher than the predicted by simulation, [22]. The main reason is that our mismatch models were not completely accurate. Usually circuit simulators do not model mismatch of slope factor (or gamma). Since the transistors that inject or subtract charge from C_2 (M_{p_4}, M_{p_3}, M_{n_2} and M_{n_3}) are sensitive to body-effect, such mismatch affects performance, even though it is not detected by most simulators. Charge injection has also a random component not modelled. Another reasons are the arbitrary delays introduced by the periphery and the USBAERmini

board that handles the communication with a PC to display real-time images. Other non-idealities that we have observed and have an influence in the contrast computation is the charge leakage at C_2. This leakage provokes voltage drops of the stored voltage at C_2 that alter the contrast computation, in situations where the average event rate is low (due to low illumination and/or low spatial contrast). We also believe that for future system integrations the more feasible option to improve FPN would be to implement a more populated neighbourhood to compute the contrast computation. By doing this, mismatch between neighbors influence would be averaged. For instance, considering an 8-pixel neighbourhood would have an extra cost of four transistors per pixel, but could improve the FPN figure. The choice of the neighborhood has a tradeoff between layout and pixel complexity, FPN, and spatial resolution. By increasing neighbors that participate, FPN will be lower. However, the area consumption will be higher, and the spatial resolution detecting contrast will be lower (edges will be wider in the resulting image). Operating in Intensity Mode, the FPN is 2.4%. In that case, voltage increments/decrements have not any impact in the contrast computation, for this reason, the FPN is lower. Still arbitrary delays introduced by the AER arbitration system affect the contrast computation.

If we scaled down the technology, the transistor mismatch impact would be presumably higher and could be predicted according the Pelgrom's law [34]. However, reducing the transistors size, more transistors could be added in the same pixel area. Hence, we could add more connectivity among neighbouring pixels as previously discussed, making the system more robust against mismatch when computing spatial contrast.

The impact of dark current and transistor leakage at the integration capacitance C_1 was also quantified. Dark current and transistor leakage provoke charge losses at the integration capacitance. If the voltage at the integration capacitance reaches the voltage threshold V_{ref_1}, an undesired packet of charge will be injected in C_2. In intensity mode, this voltage increment will elicit one event. There is a tradeoff between speed and the impact of dark current. If we lower the value of V_{ref_1}, the residual event activity will be lower, but also the latency response of the sensor. Without any illumination, with $V_{ref_1} = 2.4$V, the average event rate is $\bar{f} = 1.2$Hz, and with $V_{ref_1} = 1$V the average event rate due to the dark current and the transistor leakage was $\bar{f} = 0.45$Hz.

E. Dynamic Range

Fig. 15 shows the dependence between the event rate and and chip illuminance in intensity mode. For illumination values higher than 100lux, there is approximately a linear dependence between event rate and illumination. For illumination values below 10lux, dark current and leakage effect start being noticeable. The sensor can operate satisfactorily in environments with uncontrolled illumination conditions. We define the dynamic range as the ratio between the highest and the lowest illumination values with illumination conditions where the sensor can operate in both operation modes. The lowest illumination value is limited by the dark current and the
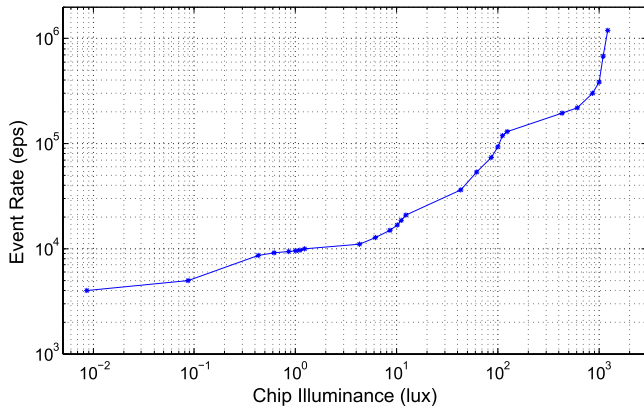
Fig. 15. Dependency between scene illumination and event rate in Intensity Mode.

transistor leakage. Both factors discharge slowly C_1 generating output events. Under very low illumination conditions, these events can be comparable to the average pixel event activity and limit the sensor performance. With high illumination values, the event activity growth is proportional to illumination. In such situations, the periphery can be a bottleneck if the event rate is higher than 10Meps, that is the maximum event rate that our arbitration system can handle [30]. In such situations, the arbitration system can introduce random delays that corrupt the event timing. Operating in intensity mode, the event rate depends on illumination, and the periphery saturates earlier than operating in contrast detection mode. In the last operation mode, the dynamic range is sensibly higher because the event rate is lower. The sensor dynamic range (operating in intensity mode) is higher than 90dB. The lowest chip illuminance where the sensor was able to render images was 0.02lux. The highest illumination value where the sensor was tested without corrupting the event timing was higher than 1klux. Note that operating in contrast detection mode, this dynamic range can be extended. In practical situations, with an array of 64 × 64, is difficult to exceed event rates higher than 10Meps. Hence, the practical dynamic range is higher than 90dB. To saturate the periphery, we placed the sensor very close to a very bright light source. Such situations are difficult to find in real application scenarios.

F. Latency

Fig. 16 displays the sensor response latency operating in contrast detection mode. The sensor was focusing a pattern with black and white transitions. The average pixels latency detecting spatial contrast was measured for different chip illuminance values. Error bars indicating the inter-pixel deviation for each measurement have been added. Under very bright illumination values, the pixel latency response is close to 400μs with an inter-pixel deviation of 50μs. Inter-pixel deviations are mainly due to pixel mismatch and arbitration delays introduced by the asynchronous arbitration system.

G. Contrast Sensitivity

Let us considerer a pixel neighborhood exposed to a illumination step. The half of the neighbors and the central pixel
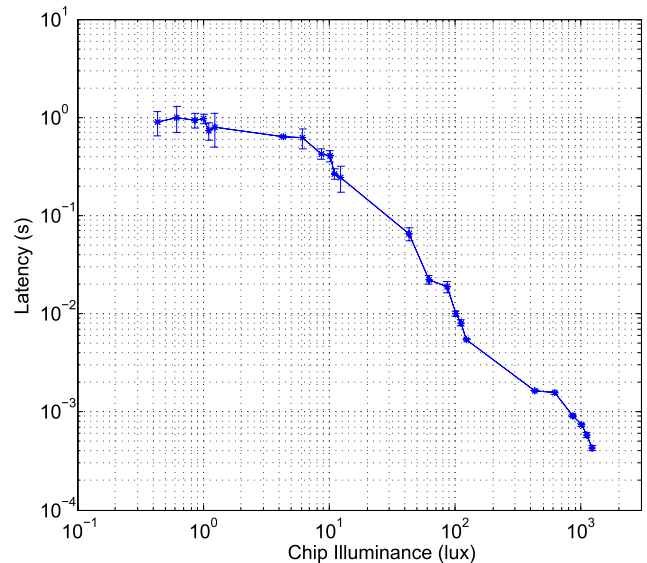

Fig. 16. Chip illuminance versus latency response operating in contrast mode. Error bars showing the inter-pixel deviation for each measurement have been added to the plot.
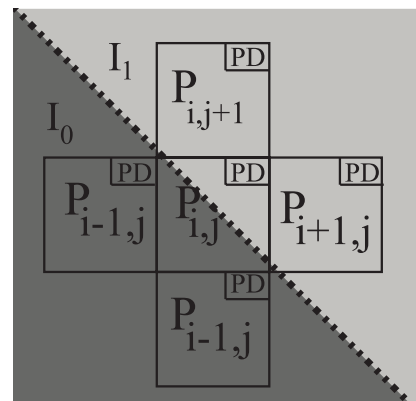

Fig. 17. Experiment to determine the pixel contrast sensitivity. We generated a illumination step between two adjacent regions. The central pixel and two of its neighbors were exposed to a illumination level I_1. The other two neighbors were exposed to an illumination level I_0, lower than I_1. I_1 was swept to vary the value of the WC, while I_0 was kept constant.

receive a certain level of illumination I_1. The rest of neighbors receive a level of illumination I_0, as it is depicted in Fig. 17. If $I_1 > I_0$, the central pixel will spike with a frequency that will depend on the difference between illumination levels $I_1 - I_0$, and the value of voltage variation at C_1 necessary to elicit one voltage increment at C_2, $\Delta V_1 = V_{DD} - V_{ref_1}$. ΔV_1 was kept constant during the experiment, $\Delta V_2 = V_{DD} - V_{ref_2} = 1V$.

The Weber Contrast (WC) between two adjacent regions with different illumination values is defined as:

$$WC = \frac{I_1 - I_0}{I_1 + I_0} \qquad (6)$$

Where I_1 and I_0 are the photocurrent values of pixels of the adjacent regions. Operating, we can express the difference between illumination levels $I_1 - I_0$ as:

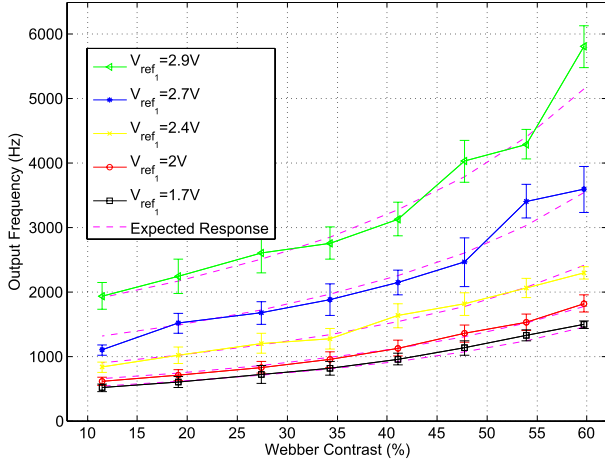$$I_0 - I_1 = \frac{2 \cdot WC \cdot I_0}{1 - WC} \qquad (7)$$

Fig. 18. Contrast sensitivity measurements. Pixels of one column were exposed to illumination steps with different Weber contrast values. Their output frequencies were measured and averaged. Error bars indicate responses standard deviation.

The spiking frequency of the second integrate-and-fire neuron in Fig. 6 is given by:

$$f = \frac{I_+ - I_-}{C \cdot \Delta V_2} \tag{8}$$

If we assume that two neighbors photodiodes generate a photocurrent value I_1, and the other two are exposed to a photocurrent value I_0, the spiking frequency would be proportional to:

$$f \propto \frac{I_{i,j} - k' \left(I_{i+1,j} + I_{i-1,j} + I_{i,j+1} + I_{i,j-1} \right)}{C \cdot \Delta V_{ref_2}} \tag{9}$$

The constant of proportionality, α, will depend of ΔV_{ref_1}. Taking into account that $I_1 = I_{i+1,j} = I_{i-1,j}$, $I_0 = I_{i,j+1} = I_{i,j-1}$, and introducing the constant $k = 2k'$ in Eq. 9, we can express the firing frequency as:

$$\begin{aligned} f &= \alpha \frac{I_1 (1-k) - k \cdot I_0}{C \cdot \Delta V_{ref_2}} \\ &= \alpha \frac{1-k}{C \cdot \Delta V_{ref_2}} \left(I_1 - I_0 - \frac{k \cdot I_0}{1-k} \right) \end{aligned} \tag{10}$$

The value of I_0 is kept constant. We vary the value of I_1 to change the value of WC and measure the output frequency. α is a constant that is proportional to V_{ref_1}, i.e. $\alpha = f(V_{ref_1})$. Hence, the pixel spiking frequency can be expressed as a function of the WC and I_0:

$$f = \alpha \frac{1-k}{C_2 \cdot \Delta V_{ref_2}} \left[\frac{2 \cdot WC \cdot I_0}{1 - WC} + \frac{1 - 2 \cdot k}{1-k} \cdot I_0 \right] \tag{11}$$

Fig. 18 shows the average response of the pixels of one column to different illumination steps with values specified in Table III. In our experiment, we measured the output frequency for different values of the WC. Measurements have been repeated for different values of ΔV_{ref_1}, keeping constant the rest of system parameters. We observed that we can fit the experiment results with the family of curves described by Eq. 11 for different values of parameter α. Error bars indicating the inter-pixel deviation of each measurement have

TABLE III
LUMINANCE STEPS AND WEBER'S CONTRAST RELATIONSHIP

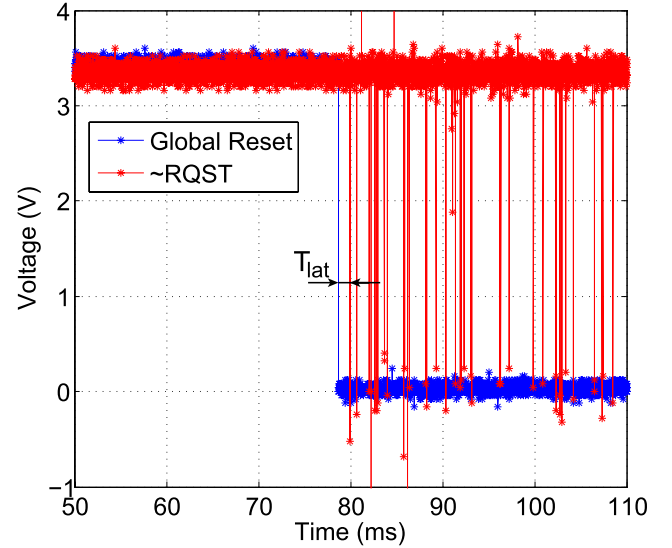| Luminance Step	100% to 0%	90% to 0%	80% to 0%	70% to 0%	60% to 0%	50% to 0%	40% to 0%	30% to 0%
I_{bright}/I_{dark}	3.9	3.3	2.8	2.4	2.0	1.75	1.5	1.25
Weber's Contrast	0.59	0.53	0.47	0.41	0.34	0.27	0.19	0.11

Fig. 19. Measured pixels responses when reseting the whole array. Pixels with higher illumination/contrast will spike first and one time.

been added. The value of the Weber contrast was determined for each illumination step computing the value of the ratio I_0/I_1. To do that, we measured the average event activity in each region operating in intensity mode. The higher V_{ref_1} is, the higher the event rate is. The sensor contrast sensitivity can be defined as the slope of the linear fitting of each plot. With $V_{ref_1} = 2.9V$, we can achieve a contrast sensitivity of 7khz/WC. We can interpret such feature in the following way: under normal values of WC (see Fig. 18), all pixels have a temporal resolution below one millisecond, i.e. they spike with frequencies above 1Khz when detecting spatial contrast.

H. Time-to-First Spike Readout Mode

The time-to-first spike (TFS) readout mode is an readout mode that starts will a chip global reset. Pixels can spike at the most once. The ones exposed to higher illumination/contrast (depending on the operation mode) will spike first. This readout mode, combined with the contrast detection mode, provides an extremed compressed data flow. A low number of events is necessary to identify elements of the visual scene [37]. If the pixel array is reset several times to track a moving scene, the operation is similar to a frame-based operation where time between consecutive frames can be adjusted dynamically, depending on the desired speed. Fig. 19 displays how pixels send out the chip events after a global reset. Initially, pixels are reset and silent. No events are sent out. After the reset is disabled, pixels start firing, sending

TABLE IV

STATE-OF-THE-ART COMPARISON

Work	Culurciello 2003 [3]	Chen 2007 [4]	Zaugh 2004 [25], [26]	Ruedi 2003 [35]	Costas 2007 [27]	Juanle 2010 [5]	Juanle 2013 [28]	Brandli 2014 [36]	This work
Functionality	Intensity Images	Intensity Images	Spatio-temporal Contrast Detection	Spatial Contrast and Direction	Spatial Contrast Detection	Spatial Contrast Detection	Intensity Images (color)	Temporal Contrast and APS	Intensity Images and Spatial Contrast Detection
Technology	0.6µm 3M	AMIS 0.35µm 5M 1P	0.35µm 4M	0.5µm 3M 2P	AMS4M2P 0.35µm	AMS4M2P 0.35µm	STM 90nm	0.18µm 1P6M MIM CIS	AMS4M2P 0.35µm
Number of operation/readout modes	1/1	1/1	1/1	2/1	1/1	1/2	1/2	2/2	**2/2**
Instantaneous Operation Mode Reconfiguration	NA	NA	NA	No	NA	NA	NA	No	**Yes**
Readout Type	PDM	TFS	PDM	TFS	PDM	PDM and TFS	PDM and TFS	PDM and frame-based	PDM and TFS
Spatial Contrast Computation	N/A	N/A	Yes (diffusive grid)	Yes (4 nearest pixels (North, South, West, East))	Yes (adjustable diffusive grid(up to 10 pixels))	Yes (adjustable diffusive grid)	N/A	No	**Yes (Inter-pixel spike communication)**
Calibration for spatial contrast detection	NA	NA	No	No	Yes	Yes	NA	NA	**No**
Pixel Pitch	32µm×30µm	17µm×17µm	34µm×40µm	69µm×69µm	58µm×56µm	81.5µm ×76.5µm	31µm×31µm	18.5µm ×18.5µm	40µm×40µm
Array resolution	80 × 60	128 × 128	96 × 60	128 × 128	32 × 32	32 × 32	22 × 22	240 × 180	64 × 64
Fill factor	14%	17%	7.3%	9%	3%	2%	28%	22%	8.5%
FPN	4%	4.6%	1-2dec	1.7%	6.6%	0.9%	3 − 5.8%	0.5% (APS)-3.5% DVS	3.3% (CD) / 2.4% (IM)
FPN (Spatial Contrast Detection Mode)	NA	NA	1-2dec	1.7%	6.6% (after calibration)	0.9% (after calibration)	NA	NA	3.3% **(CD)**
Dynamic Range	ND	>100dB	50dB	110dB	100dB	100dB	>60dB	130dB (DVS), 51dB (APS)	>90dB
Latency	120-125µs	10µs -1s	Not reported	2ms-150ms	Not reported	0.1-10ms	N/A	3µs	400µs -1s
Power consumption	3-71mW	N/A	63mW	300mW	33µW-10mW	0.66-6.6mW	5-14mW	0.5-5mW	6.1mW

RQST signals. There is a latency (T_{lat}) between the global reset and the events that is inversely proportional to the light intensity or spatial contrast detected by pixels. This latency is used to render the resultant images. The user can decide how many events are selected to reconstruct one image.

Fig. 20 depicts how this readout mode performs taken snapshots with the camera. We took an snapshot of a static screw combining the different operation and readout modes. Fig. (a) displays the image taken in intensity mode with PDM readout. Fig. 20.(b) displays the image taken in contrast detection mode with PDM readout. The TFS readout is useful in the contrast detection mode because only pixels that detects spatial contrast will spike at the most once. Fig. 20. (c, d, e, f, g, h) shows the same image taken in contrast detection mode with TFS readout, reconstructed after receiving N_{spikes}. We can observe that a number of events of lower than 1,000 and delay lower than 30ms, it is possible to identify the screw. The user can decide how many events, N_{spikes}, receive to render one image and reset the sensor again, after receiving them. This operation mode is particularly useful for surveillance applications, where bandwidth and power consumption have to be saved. If changes in the visual scene are slow, time between resets can be enlarged, giving the possibility of adjusting the sensor temporal resolution dynamically.

I. Benchmarking and Comparison With the State-of-the-Art

If we compare the performance of the proposed sensor with designs based on frame-based sensors, the power consumption saved in the data transmission within a wireless camera network could be reduced significantly. For comparative purposes,
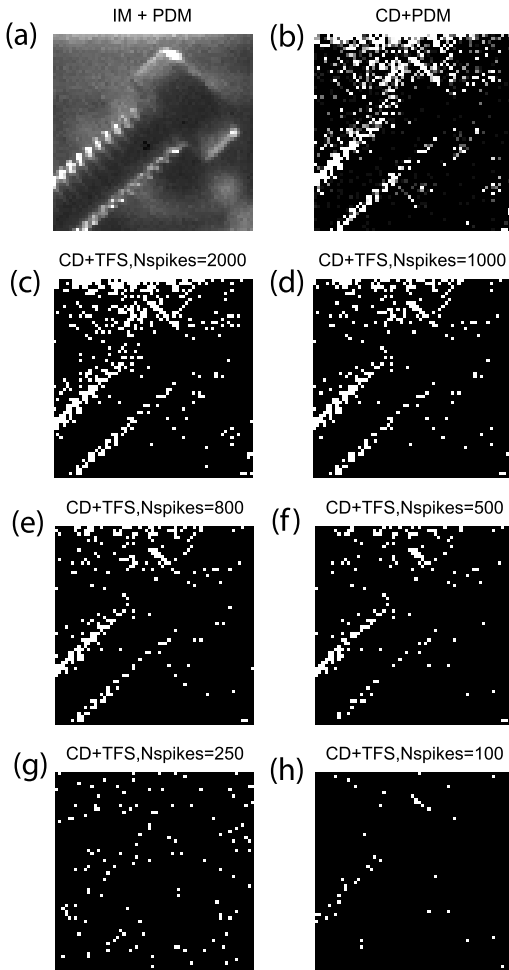
Fig. 20.   Snapshot taken combining different operation and readout modes. (a) Intensity mode + PDM, (b) Contrast mode + PDM, (c, d, e, f, g, and h) Contrast mode + TFS, reconstructing the image representing the first N_{spikes} received.

the sensor reported by Chen et al. [12] has a UWB transmitter with a power consumption of 15mW@1.3Mbps, i.e. 11.5nJ/bit. If we assume that we are operating with a frame-based sensor, with the same pixel array resolution, a frame rate of 25fps, and 8 bits to codify pixel illumination, its output data flow would be 800kbits/s. Under average illumination conditions (see snapshots in Fig.13 and event rates in Fig. 15), in all the operation modes, the event rate of the proposed sensor is always lower than 100kevents/s. Therefore, assuming a linear dependence between the event rate and the power consumption, the transmission power consumption would be 9.2mW with the frame-based sensor and 1.15mW with our sensor. The power saved in the data transmission could even be higher operating with TFS read out.

Comparing the new sensor to previously reported event-based vision sensors, the new device abandons the traditional approach to compute spatial contrast based on diffusive networks, [5], [27]. Instead, inter-pixel spike communication is proposed, improving the pixel fill factor over other devices that compute spatial contrast with diffusive networks [5], [25]–[27].

Table IV summarizes the most relevant and recently published event-based vision sensors that can provide intensity

and/or contrast detection on the focal plane. The new sensor is the only one who has two operation and two readout modes, offering a good deal between fill factor, dynamic range, power consumption, and pixel complexity. It is possible to toggle between operation modes instantaneously, without reconfiguring the sensor, with only one control bit. The FPN values are competitive, taking into account that the sensor does not need to be calibrated. In this work, a 4-pixel neighborhood was considered to compute spatial contrast. The neighborhood considered in the contrast computation could be extended adding one extra transistor per extra neighbor.

VI. Conclusion

A new event-based vision sensor that can provide intensity images and detect spatial contrast have been presented. Two readout modes are available: PDM and TFS. The operator can toggle among operation and readout modes with two control bits, trading between image quality and bandwidth and/or power consumption. The sensor is suitable to be a node of smart camera networks, reconfiguring its operation depending on the requirements given by a central node or an operator. The spatial contrast computation is achieved, without calibration, with a novel approach based on inter-pixel spike communication, more robust against mismatch than the traditional one with diffusive networks. Detailed experimental results and sample images are provided. The new sensor has been bench marked and offers a flexible and competitive alternative to previously reported sensors. As a future work, we plan to improve the contrast detection by adding a large neighborhood when computing spatial contrast. We also plan to run experiments to test the sensor in real Embedded Smart Camera Networks and quantify the benefit that we expect from its reconfigurability.

References

[1] J. E. Dowling, *The Retina: An Approachable Part of the Brain*, 2nd ed. Cambridge, MA, USA: Harvard Univ. Press, 2012.

[2] D. G. Chen, D. Matolin, A. Bermak, and C. Posch, "Pulse-modulation imaging—Review and performance analysis," *IEEE Trans. Biomed. Circuits Syst.*, vol. 5, no. 1, pp. 64–82, Feb. 2011.

[3] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, "A biomorphic digital image sensor," *IEEE J. Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, Feb. 2003.

[4] C. Shoushun and A. Bermak, "Arbitrated time-to-first spike CMOS image sensor with on-chip histogram equalization," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 15, no. 3, pp. 346–357, Mar. 2007.

[5] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, "A five-decade dynamic-range ambient-light-independent calibrated signed-spatial-contrast AER retina with 0.1-ms latency and optional time-to-first-spike mode," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 10, pp. 2632–2643, Oct. 2010.

[6] G. Orchard, C. Meyer, R. Etienne-Cummings, C. Posch, N. Thakor, and R. Benosman, "HFirst: A temporal approach to object recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 10, pp. 2028–2040, Oct. 2015.

[7] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014. [Online]. Available: http://www.sciencemag.org/content/345/6197/668.abstract

[8] G. Orchard, J. G. Martin, R. J. Vogelstein, and R. Etienne-Cummings, "Fast neuromimetic object recognition using FPGA outperforms GPU implementations," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 8, pp. 1239–1252, Aug. 2013.

[9] C. Posch, T. Serrano-Gotarredona, B. Linares-Barranco, and T. Delbruck, "Retinomorphic event-based vision sensors: Bioinspired cameras with spiking output," *Proc. IEEE*, vol. 102, no. 10, pp. 1470–1484, Oct. 2014.

[10] L. Camuñas-Mesa, C. Zamarreño-Ramos, A. Linares-Barranco, A. J. Acosta-Jimenez, T. Serrano-Gotarredona, and B. Linares-Barranco, "An event-driven multi-kernel convolution processor module for event-driven vision sensors," *IEEE J. Solid-State Circuits*, vol. 47, no. 2, pp. 504–517, Feb. 2012.

[11] M. Gottardi, N. Massari, and S. A. Jawed, "A 100 μw 128 $\times$ 64 pixels contrast-based asynchronous binary vision sensor for sensor networks applications," *IEEE J. Solid-State Circuits*, vol. 44, no. 5, pp. 1582–1592, May 2009.

[12] S. Chen, W. Tang, X. Zhang, and E. Culurciello, "A 64 $\times$ 64 pixels UWB wireless temporal-difference digital image sensor," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 12, pp. 2232–2240, Dec. 2012.

[13] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.

[14] J. Fernández-Berni, R. Carmona-Galán, R. del Río, R. Kleihorst, W. Philips, and Á. Rodríguez-Vázquez, "Focal-plane sensing-processing: A power-efficient approach for the implementation of privacy-aware networked visual sensors," *Sensors*, vol. 14, no. 8, pp. 15203–15226, Aug. 2014.

[15] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128 $\times$ 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.

[16] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB dynamic range frame-free PWM image sensor with lossless pixel-level video compression and time-domain CDS," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 259–275, Jan. 2010.

[17] J. A. Leñero-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.

[18] G. Pekhteryev, Z. Sahinoglu, P. Orlik, and G. Bhatti, "Image transmission over IEEE 802.15.4 and ZigBee networks," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 4, May 2005, pp. 3539–3542.

[19] A. Nilchi, J. Aziz, and R. Genov, "Focal-plane algorithmically-multiplying CMOS computational image sensor," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1829–1839, Jun. 2009.

[20] S. Chen, A. Bermak, and Y. Wang, "A CMOS image sensor with on-chip image compression based on predictive boundary adaptation and memoryless QTD algorithm," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 4, pp. 538–547, Apr. 2011.

[21] Y. Nishikawa, S. Kawahito, M. Furuta, and T. Tamura, "A high-speed CMOS image sensor with on-chip parallel image compression circuits," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Sep. 2007, pp. 833–836.

[22] J. A. Leñero-Bardallo and P. Häfliger, "A dual-operation-mode bio-inspired pixel," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 11, pp. 855–859, Nov. 2014.

[23] J. A. Leñero-Bardallo and P. Häfliger, "A dual operation mode bio-inspired vision sensor," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct./Nov. 2013, pp. 310–313.

[24] M. Mahowald, *An Analog VLSI System for Stereoscopic Vision*. Norwell, MA, USA: Kluwer, 1994.

[25] K. A. Zaghloul and K. Boahen, "Optic nerve signals in a neuromorphic chip I: Outer and inner retina models," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 657–666, Apr. 2004.

[26] K. A. Zaghloul and K. Boahen, "Optic nerve signals in a neuromorphic chip II: Testing and results," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 667–675, Apr. 2004.

[27] J. Costas-Santos, T. Serrano-Gotarredona, R. Serrano-Gotarredona, and B. Linares-Barranco, "A spatial contrast retina with on-chip calibration for neuromorphic spike-based AER vision systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 7, pp. 1444–1458, Jul. 2007.

[28] J. A. Leñero-Bardallo, D. H. Bryn, and P. Häfliger, "Bio-inspired asynchronous pixel event tricolor vision sensor," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 3, pp. 345–357, Jun. 2014.

[29] J. A. M. Olsson and P. Häfliger, "Two color asynchronous event photo pixel," in *Proc. ISCAS*, May 2008, pp. 2146–2149.

[30] P. D. Häfliger, "A spike based learning rule and its implementation in analog hardware," Ph.D. dissertation, Dept. Appl. Comput. Sci., ETH Zürich Univ., Zürich, Switzerland, 2000. [Online]. Available: http://www.ifi.uio.no/~hafliger

[31] K. A. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 47, no. 5, pp. 416–434, May 2000.

[32] R. Berner, T. Delbruck, A. Civit-Balcells, and A. Linares-Barranco, "A 5 Meps \$100 USB2.0 address-event monitor-sequencer interface," in *Proc. ISCAS*, New Orleans, LA, USA, May 2007, pp. 2451–2454.

[33] *jAER Open Source Project*. [Online]. Available: http://sourceforge.net/projects/jaer/, accessed Nov. 8, 2006.

[34] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct. 1989.

[35] P.-F. Ruedi *et al.*, "A 128 $\times$ 128 pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2325–2333, Dec. 2003.

[36] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A 240$\times$180 130 dB 3μs latency global shutter spatiotemporal vision sensor," *IEEE J. Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, Oct. 2014.

[37] R. Van Rullen, J. Gautrais, A. Delorme, and S. Thorpe, "Face processing using one spike per neurone," *Biosystems*, vol. 48, nos. 1–3, pp. 229–239, Nov. 1998.

Juan A. Leñero-Bardallo received the M.Sc. degree in telecommunications engineering and the Ph.D. degree in microelectronics from the University of Seville, Seville, Spain, in 2005 and 2010, respectively.
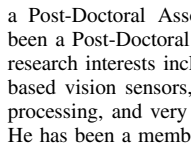
From 2006 to 2010, he was working towards the Ph.D. degree, sponsored by a national grant. In 2008, he was a Visiting Scholar with the University of Oslo, Oslo, Norway, for two months. In 2010, he was a Post-Doctoral Associate at Yale University, New Haven, CT, USA. From 2010 to 2013, he was a Post-Doctoral Associate at the University of Oslo. Since 2013, he has been a Post-Doctoral Associate at the University of Seville, Spain. His main research interests include address event representation vision systems, frame-based vision sensors, smart sensors, wireless vision sensor networks, signal processing, and very large scale integration emulators of biological systems. He has been a member of the IEEE Sensory Systems Committee since 2012.

Philipp Häfliger (M'03–SM'10) received the M.Sc. degree in computer science (with astronomy as a second subject) from the Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, in 1995, and the Ph.D. degree from the Institute of Neuroinformatics, ETH, in 2000. He then moved for a post-doctoral position to the Nanoelectronics Group at the Institute of Informatics, University of Oslo, Norway, where he is now an Associate Professor. In Oslo, his research focus shifted gradually from neuromorphic (inspired by the nervous system) electronics toward ultralow-power ASIC design for biomedical devices, in particular, wireless microimplants.

Dr. Häfliger has been the Chairman of the IEEE Circuits and Systems (CAS) Society's Biomedical and Life Science CAS (BioCAS) Technical Committee (2010–2012), and has served and still serves in various functions for the IEEE ISCAS and IEEE BioCAS Conferences. He is also an Associate Editor of the IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.

R. Carmona-Galán (M'04) received the Licenciado and Ph.D. degrees in physics, in the speciality of electronics, from the University of Seville, Spain, in 1993 and 2002, respectively. He was a Research Assistant with the University of California, Berkeley. From 1999 to 2005, he was an Assistant Professor with the Department of Electronics and Electromagnetism, School of Engineering, University of Seville. Since 2005, he holds a tenured position with the Institute of Microelectronics of Seville (CSIC). He also held a post-doctoral position at the University of Notre Dame, IN, USA (2006–2007). His main research areas are vision chips, in particular, smart CMOS imagers for low-power vision applications like robotics, vehicle navigation, and vision-enabled wireless sensor networks. He is also interested in CMOS-compatible sensing structures for LWIR and MWIR imaging, single-photon detection, and detector for X-ray and high-energy physics, and also in the implementation of high resolution smart imagers in 3-D integrated circuit technologies. He is a member of the IEEE-CASS Technical Committees on Cellular Nanoscale Networks and Array Computing and on Sensory Systems. He is member of the Steering Committee of the Workshop on Architecture of Smart Cameras, and has recently chaired the 9th International Conference on Distributed Smart Cameras in-cooperation with ACM-SIGBED. He has served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS from 2012–2013

Ángel Rodríguez-Vázquez (M'80–F'96) received the Ph.D. degree. He is currently a Full Professor of Electronics with the University of Seville.

His research is on the design of analog and mixed-signal front-ends for sensing and communication, including smart imagers, vision chips and low-power sensory-processing microsystems. He has authored 11 books, 34 additional book chapters, and some 150 journal articles in peer-review specialized publications. He has presented invited plenary lectures at different international conferences and has received a number of awards for his research (the IEEE Guillemin-Cauer Best Paper Award, two Wiley's IJCTA Best Paper Awards, two IEEE ECCTD Best Paper Awards, one SPIE-IST Electronic Imaging Best Paper Award, the IEEE ISCAS Best Demo-Paper Award, and the IEEE ICECS Best Demo-Paper Award). He was elected as a Fellow of the IEEE for his contributions to the design of chaos-based communication chips and neuro-fuzzy chips. His research work received some 6,600 citations; he has an h-index of 43 and an i10-index of 133.

He has always been looking for the balance between long-term research and innovative industrial developments. He founded AnaFocus Ltd., in 2001, on the basis of his patents on vision chips, and served as CEO, on leave from the University, until 2009, when the company reached maturity as a worldwide provider of smart CMOS imagers and vision systems-on-chip.

He has served as an Editor, an Associate Editor, and a Guest Editor of different IEEE and non-IEEE journals, is in the committee of several international journals and conferences, and has been the Chair of several international IEEE and SPIE Conferences. He served as VP Region 8 of the IEEE Circuits and Systems Society (2009-2012) and as the Chair of the IEEE CASS Fellow Evaluation Committee (2010, 2012, 2013, 2014, and 2015).