Received 26 April 2018; | Revised 6 June 2018; | Accepted 6 July 2018

DOI: xxx/xxxx

RESEARCH ARTICLE

Applications of Event-based Image Sensors —Review and Analysis[†]

Juan A. Leñero-Bardallo* | Ricardo Carmona-Galán | Ángel Rodríguez-Vázquez

[1]Institute of Microelectronics of Seville (IMSE-CNM), Consejo Superior de Investigaciones Científicas y Universidad de Sevilla, C/ Américo Vespucio s/n, 41092, Seville, Spain

Correspondence
*Juan A. Leñero-Bardallo. Email: juanle@imse-cnm.csic.es

Summary

The spread of event-driven asynchronous vision sensors during the last years has increased significantly the industrial interest and the application scenarios for them. This article reviews the main fields of application that event-based image sensors have found during the last twenty years. We focus in the description of applications where such devices can outperform conventional frame-based sensors. The practical functions of the three main families of asynchronous event-based sensors are analyzed. The article also studies what are the factors that increase nowadays the demand of sensors that minimize the power and bandwidth consumption. Moreover, the technological factors that have facilitated the development of asynchronous sensors are discussed.

KEYWORDS:
Image sensors, event, spike, Address Event Representation, AER, asynchronous

1 | INTRODUCTION

Since Mahowald and her colleagues published in the early nineties the first integrated asynchronous bio-inspired vision sensor[1], the spread of event-driven devices has grown significantly[2]. Depending on the information from the visual scene that is processed, there are three main categories of event-based (also known as event-driven or just spiking) vision sensors. The first, and the oldest group, corresponds to sensors that detect spatio-temporal contrast within the visual scene. In their origins, they were bio-inspired sensors that try to emulate the early processing of the visual scene that is performed in the human retina. The second and the most successful category so far corresponds to Dynamic Vision Sensors (DVS). These are devices whose pixels spike whether they detect transient illumination changes. They can track very fast moving objects with very low power and bandwidth data consumption. Finally, the third category corresponds to spiking luminance sensors, also known in the literature as *octopus retinas*. Although they just perform a light-to-frequency conversion, they preserve the asynchronous communication

[†]This work was supported in part by the Spanish Ministry of Economy and Competitiveness under Grant TEC2015-66878-C3-1-R, Co-Funded by ERDF-FEDER; in part by Junta de Andalucía CEICE under Grant TIC 2012-2338 (SMARTCIS-3D); and in part by ONR under Grant N000141410355 (HCELLVIS).

based on spikes to encode the illumination levels. Only illuminated pixels can send information off-chip. Their dynamic range and their latency are quite competitive. Thus, they are attractive for applications where is necessary to detect illumination levels of certain regions with low resources consumption.

Although the commercial success of asynchronous vision sensors is still far away from frame-based sensors, there are several companies like Prophesee or Inilabs dedicated to the fabrication of event-based cameras. The rise of the demand of vision sensors with low power and bandwidth consumption, and the capability of processing the visual scene has put spiking sensors into the scene. Drone vision and autonomous vehicles prefer devices with such features instead ones that provide high image quality. For these reasons, asynchronous image sensors find accommodation and are highly competitive in situations where image quality is not a must.

There are several reviews and comparisons between the different event-based sensors published so far[2,3]. However, to the best of our knowledge, still there is not still published a review of the specific fields of application where they have been employed. In this article, we will focus on a description of a the fields of application where event-based vision sensors have been successfully accommodated. We will pay special attention to applications of the family of spiking luminance sensors in which the authors have made relevant contributions during the last years.

2 | SENSORS BASED ON SPATIAL CONTRAST DETECTION

Spatial contrast detection is the very first-processing of the visual scene performed by the human visual system[4]. It is implemented by the cells that compound the retina. Their outputs are transmitted as train of spikes through the optic nerve. The detection of spatio-temporal contrast allows to detect edges, and shapes. That is crucial for an early interpretation and classification of the elements within the scene. At first instance, spiking sensors capable of computing spatial contrast on the focal plane pursue to mimic a biological modelling of the human retina[4]. The first silicon retina chip was proposed by Mahowald,[1]. The original pixel topology is shown in Fig. 1 . Pixels detect light with a photoreceptor circuit which output is a voltage that is averaged within a neighborhood with a diffusive resistive network that interconnects all the pixels. If the average illumination value in the neighborhood, $\overline{V}_{p_{neighb.}}$, differs from the pixel's local illumination value, $V_{p_{i,j}}$, the pixel will spike, indicating the detection of spatial contrast. These circuits, in the literature, still they are known as silicon retinas due to their biological inspiration. There are improved versions of this preliminary approach[5,6,7]. This family of sensors opened a field of research that lead to ulterior development of event-based vision sensors. They also put emphases in the development of asynchronous circuitry to implement the well-known AER (Address Event Representation) communication protocol[8,1,9]. Modern arbitration schemes[10,11] are based on the early versions implemented on the first asynchronous bio-inspired vision sensors.

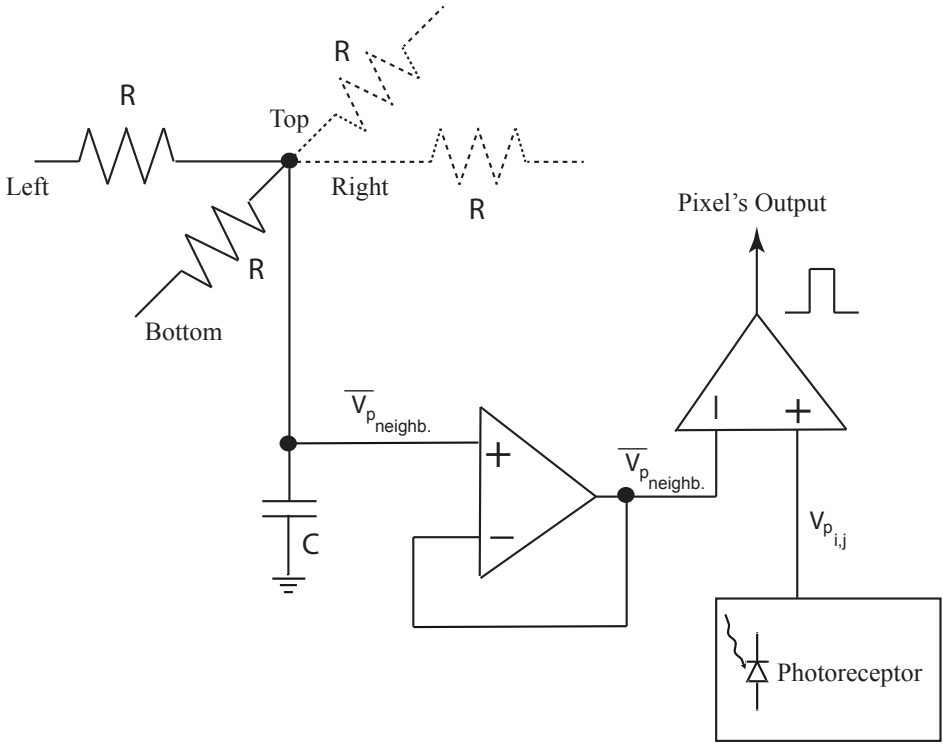

FIGURE 1 Retina contrast detection pixel according to[1]. Average illumination values within a pixel neighborhood ($\overline{V}_{p_{neighb.}}$) are computed with a diffusive resistive network. The average illumination value in the neighborhood ($\overline{V}_{p_{neighb.}}$) is compared to the local pixel illumination ($V_{p_{i,j}}$). If both values differ, the pixel spikes, with a frequency proportional to the spatial contrast magnitude.

Once the possibility of implementing these sensors was demonstrated, the community drew its attention about how to improve the quality of their images. Particularly, mismatch was an important challenge associated to the first generation of bio-inspired sensors, operating with transistors in subthreshold region. Several authors pointed out the need of reducing pixel Fixed-Pattern-Noise (FPN) associated to transistor mismatch. In that sense, calibration procedures and circuitry robust to mismatch were developed[12,7]. Nowadays, the event-driven spatial contrast computation is not highly competitive in terms of spatial noise against the traditional method: read-out an entire pixel matrix illumination values and process them on the digital domain to detect edges and shapes. The main reason is that asynchronous computation of spatial contrast on the focal plane requires non-linear operations that are difficult to implement precisely in the analog domain. However, they deserve to be mentioned in this survey because they led to the development of other families of asynchronous sensors whose applications will be described in the next sections.

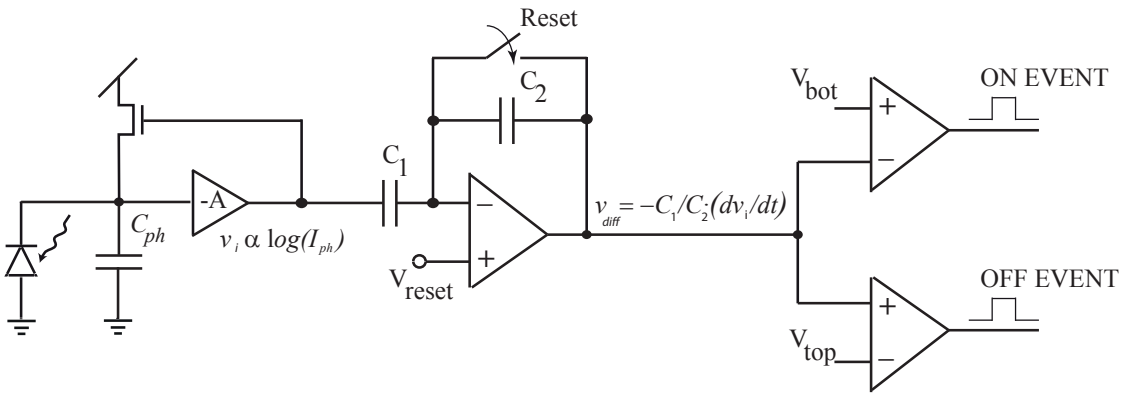

FIGURE 2 Architecture of a DVS sensor. Light is detected by a photologarithmic receptor with negative feedback and gain $-A$. Its voltage outputs feed a differencing circuits that amplifies transient voltage variations at its input, v_i. If the differencing circuit output, v_{diff}, exceeds a negative or a positive voltage threshold, events are transmitted indicating temporal contrast. The sign of the temporal contrast (positive or negative) is detected with two comparators and transmitted through two independent pathways.

3 | DVS SENSORS

The pixel's architecture of a Dynamic Vision Sensor (DVS) sensors is depicted in Fig. 2 . Light is processed by a logarithmic photoreceptor. Its output voltage, v_i, feds a differencing circuit that computes and amplifies its temporal derivative, v_{diff}. If the transient variation of differencing voltage exceeds a programmable threshold (either positive or negative), and event is sent off-chip, indicating a pixel's illumination change, its coordinates, and the sign of the temporal contrast. DVS sensors can detect fast changes of illumination, with low power and data bandwidth consumption. Movement detection is outperformed over transitional procedures, based on frame processing. Proof of that is the fact that several companies like Inilabs or Prophesee have put on their scope the development of these family of sensors.

The first functional DVS sensor was presented by Lichtsteiner et al. in 2008 [13]. In this first publication, the authors already devised possible applications for the sensor, i.e. traffic monitoring, gaming, sleep monitoring, etc. In more recent publications, several authors presented improved versions of the original DVS sensor [14,15,16,17].

Regarding the application fields of these sensors, we can highlight the same examples:

1. Datasets for implementation and training of spiking neural networks for object tracking, action recognition and/or object recognition. Spiking neural networks needs to be fed with spiking inputs. Images generated with asynchronous image sensors are ideal for this purpose. The spread and popularization of DVS sensors among the community has driven the creation of event-based image datasets recorded with DVS sensors [18,19].

2. Traffic monitoring. The control of moving objects is an ideal task for DVS sensors. Only pixels that detect illumination changes send data off-chip, minimizing the bandwidth and the power consumption. There are specific papers addressing this topic, for instance,[20]. Also DVS sensors have been used to implement proximity detectors,[21].

3. Fall detectors. Processing the output data flow of a DVS sensor, the authors propose a fall detector[22]. It was intended for home-care of elder people. In can be integrated among the sensors of an intelligent house. Gesture recognition is also possible,[23].

4. Robotics integration, control and navigation[24]. The idea is to provide images of visual scenes with cameras that are mounted on systems that are continuously moving or observing a non-static visual scene.

5. Tracking and particles detection on fluids. The excellent temporal resolution of DVS sensors has made possible to track fast movements of particles in fluids reducing the data throughout[25].

6. Color change detection[26]. Combining the pixel architecture of a DVS sensor with stacked photodiodes at different depths in standard silicon, it was possible to create an image sensor whose pixels only reacts to color changes (transient shifts within the spectrum).

7. Stereo vision[27,28]. The combination of two or more DVS sensors recording images from different localizations allows determine distances and depth, after processing their outputs. The DVS sensor reduced output data flow allows to alleviate the bandwidth and power processing consumption with regards to classic frame-based processing algorithms.

4 | SPIKING LUMINANCE SENSORS

In this paper, we will focus on summarizing practical applications where asynchronous spiking sensors have been employed. These devices, also known as *octopus retinas*, perform a light-to-frequency conversion to encode luminance levels within the visual scene[29]. Their operation principle is depicted in Fig. 3 . Pixel operation is initiated by reseting the voltage at the integration capacitance, C. Then, its voltage decreases linearly with a slope that is proportional to the pixel's photocurrent until it reaches a programmable analog voltage, V_{ref}. Thereafter, the pixel self-resets and starts integrating charge again. The pixel's oscillation period is approximately given by this expression:

$$f_{osc} \approx \frac{I_{ph}}{C \cdot \left(V_{DD} - V_{ref}\right)} = \frac{I_{ph}}{C \cdot \Delta V} \tag{1}$$

Where I_{ph} is the pixel photocurrent and C is the integration capacitance. V_{ref} is a voltage threshold to control the pixel sensitivity to light. Although these sensors do not implement any kind of processing, just measuring luminance levels, they have inherent advantages over conventional frame-based image sensors. The first one is speed; pixel spiking frequency can

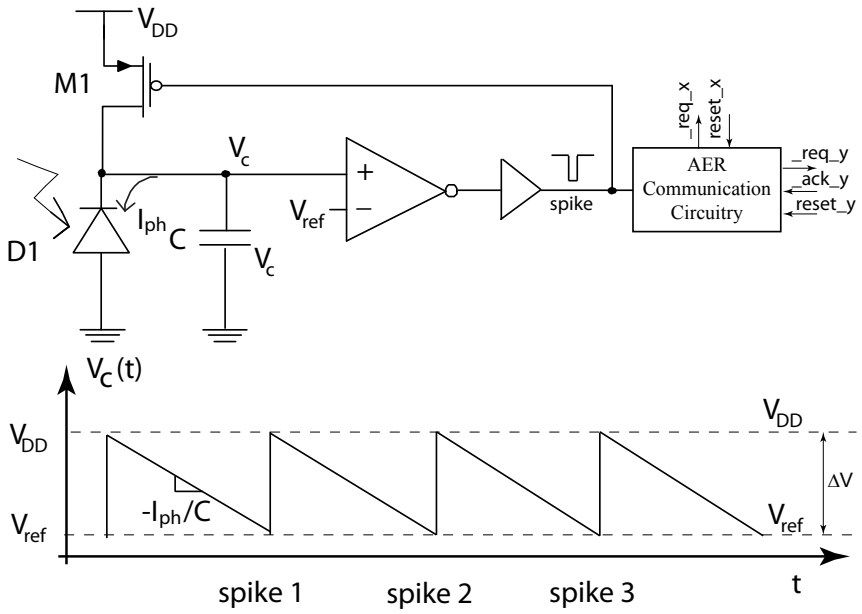

FIGURE 3 Integrate-and-fire (I&F) pixel block diagram. Pixels spike with a frequency proportional to light intensity. Every time that this occurs, pixels' addresses are transmitted asynchronously off-chip. Asynchronous digital signals involved in the data transmission are displayed on the right.

easily reach KHz rates in indoor environments[30,29]. This implies an equivalent temporal resolution in the order of Kilo-frames per second. The second advantage is their autonomous operation. Pixels have not to be scanned periodically to gauge their illumination values. They pulse continuously whenever they are illuminated. If they are dark, they do not send any data. That is also an advantage for application scenarios where dark pixels are meaningless. Another strength of them is their High Dynamic Range (HDR) operation. Pixel spiking frequencies can spam more than five decades[31,29]. This implies the capacity of detecting intra-scene illumination levels with a dynamic range higher than 100dB. Also, their power consumption is reduced with respect frame-based sensors.

Their main penalty is image quality. FPN noise is typically higher with these sensors[30,31]. However, there many applications that can sacrifice image quality in benefit of speed, HDR operation, and low power consumption. We summarize and explain some of them in the next subsections.

4.1 | Sun Sensors

Sun sensors detect the sun position (azimuth and latitude) referred to their centroid. They are demanded for green energy production in solar plants and to develop navigation systems for spacecrafts and sounding rockets, that take the sun as a reference for orientation. There are two main classic families of sun sensors: analog[33] and digital[34,35]. Analog ones are more simple and

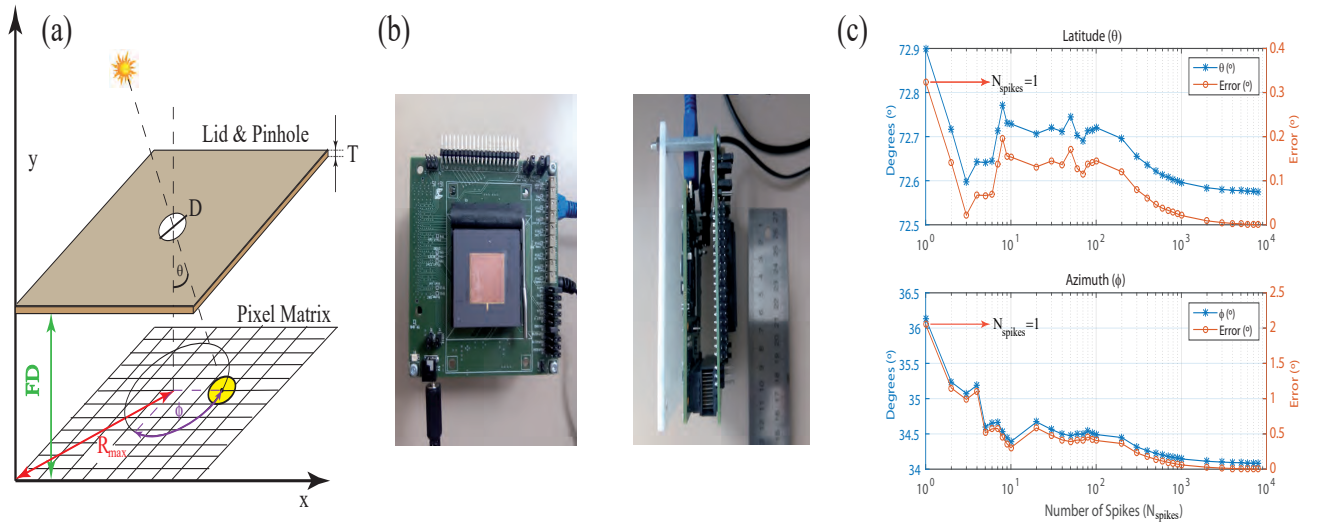

FIGURE 4 (a) Operation of a sun sensor[32] built with a pinhole lens. (b) Snapshot of an asynchronous sun sensor implemented with a pinhole lens. (c) Sensor's output data. The sun position is calculated computing the centroid of a certain number of incoming events. With only one event, it is possible to gauge the sun position with an error lower than 2.2%.

easy to design, but they are more prone to noise and sensitive to scene distractors. Digital ones are the most extended ones nowadays. They are reliable and easy to build. Deploying a dedicated optics on top of a frame-based sensor, it is possible to convert it into a sun sensor. Depending on the sun position, a group of pixels will be illuminated. Its centroid will depend on the sun localization. By processing the output data, the centroid is calculated. Then, establishing basic trigonometric relations, the sun position can be determined.

The main limitations of this approach are the data redundancy and the limited operation speed. Data redundancy is an important drawback because dark pixels needs to be readout and are useless to compute the sun position. This also drives to high bandwidth and power consumption derived of the data readout and the data processing. The second limitation is the operation speed. Since the entire pixel matrix has to be readout, every time that the sun position has to be updated, the operation speed is conditioned by the frame rate. Several authors have proposed strategies to avoid reading dark pixels[34]. However, still some of them have to be readout periodically.

A recent alternative to build digital sun sensors is the usage of octopus retinas[36,32]. By placing a dedicated optic over them, only pixels exposed to sunlight will send data. Therefore, data redundancy is avoided, simplifying the data processing, the sensor operation and reducing power consumption. Moreover, the excellent temporal resolution of octopus retinas make these devices suitable for space navigation. In Fig. 4 .a, we display the architecture of a spiking sun sensor with a pinhole lens on top of it. In Fig. 4 .b, the fabricated spiking sun sensor is displayed. Data recorded with the sensor is provided in Fig. 4 .c. The centroid computation of the illuminated region is fast and simple. Just computing the centroid of the pixels that fires, it is possible to do that. The operator has freedom to decide how many spikes are employed to compute it. Usually the centroid is closer to the highest

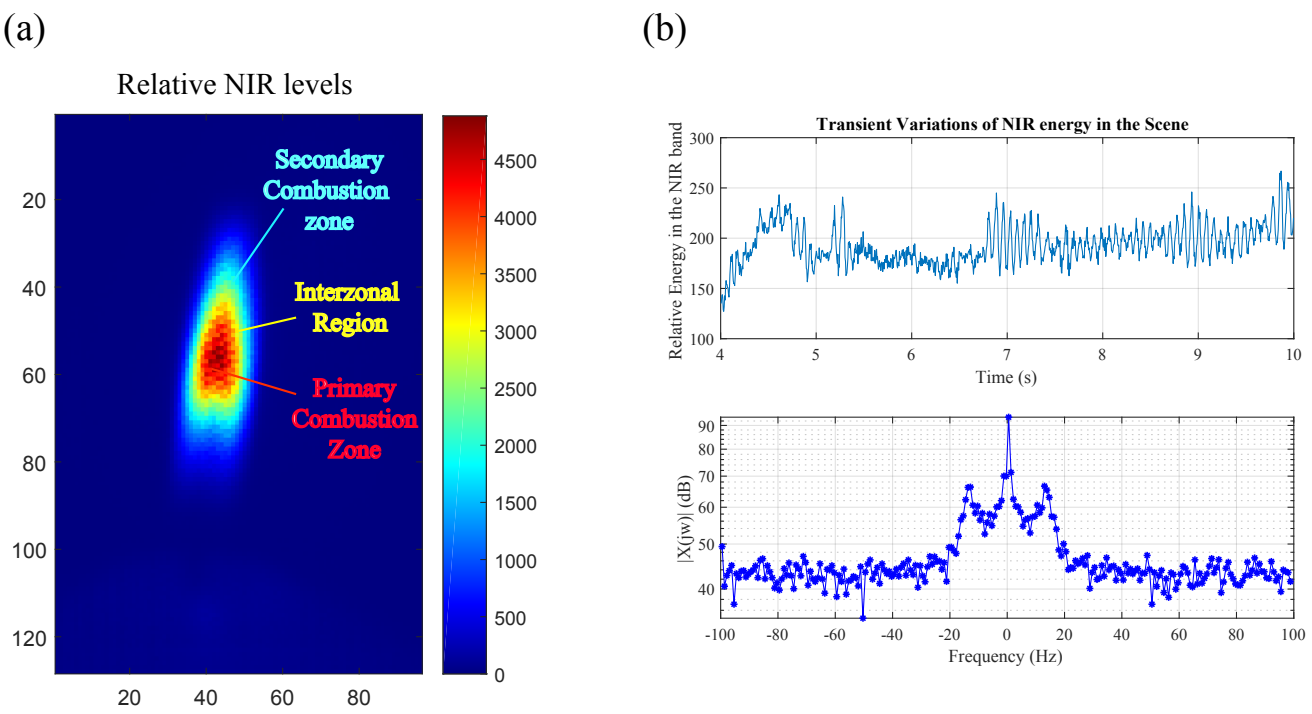

FIGURE 5 Flames monitoring with an asynchronous NIR spiking sensor[38,39]. (a) Snapshot of a flame taken with a luminance asynchronous sensor with an optical NIR filter. (b) Top: Transient NIR levels variations detected with the sensor in a visual scene with a flame. Bottom: FFT of the previous data to study the flame frequency components.

illuminated pixel that spikes first than the others. Thus, just taking into account its coordinates, it is possible to gauge the sun position with acceptable precision. In terms of dynamic range, spiking sensors allow to encode the illumination levels without saturating. Usually, frame-based sensors dynamic range is limited to 70dB[37], that is not enough to detect illumination levels inside very bright light sources like the sun that require sensors with more than 100dB dynamic range to be characterized[37].

4.2 | Flame monitoring

Asynchronous spiking sensors can be used in the detection and the analysis of flames activity,[38,39]. Flame flickering and flame emissions have to be monitored in some industrial processes[40,41]. These tasks were implemented, at first instance, with infrared (IR) cameras or microbolometers. Their main limitations are the high cost, the limited frame rate, the need of calibration in some cases, and their fragility[42].

As alternative to them, CMOS or CCD cameras working in the Near Infrared Band (NIR) have been employed. By placing a NIR filter over them, the become NIR imagers. Radiation inside this band is emitted by flames. Hence, with adequate image processing, flames activity can be monitored[43,44,45]. Flames flicker with frequency components that range between DC and 100Hz[40]. Thus, the analysis of their transient activity requires sensors with good temporal resolution. This requirement also

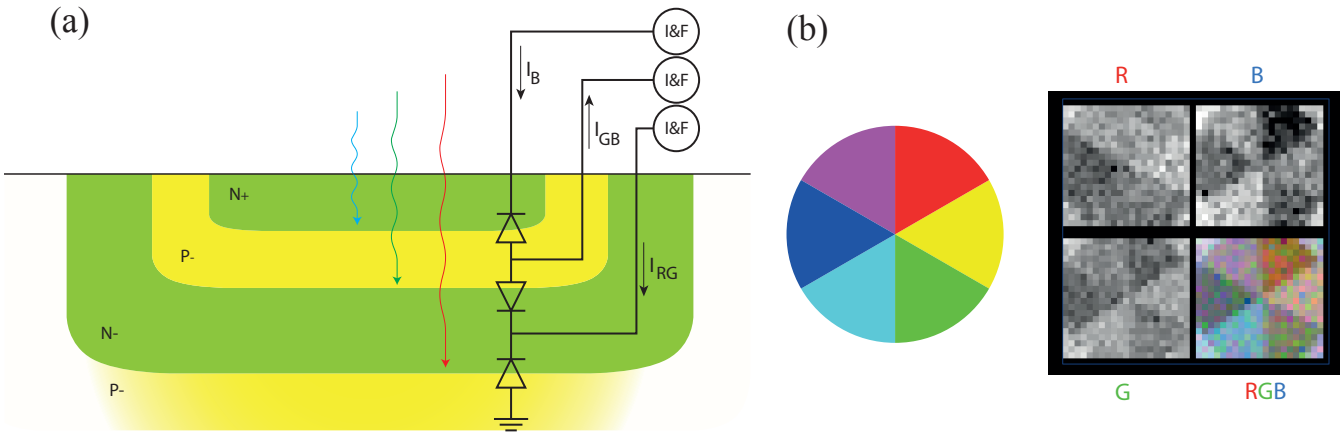

FIGURE 6 Asynchronous color sensor implementation,[46]. (a) Stacked photodiodes to achieve spectral selectivity within the visible spectrum. Each one is connected to integrate-and-fire (I&F) block as the one depicted in Fig. 4 . (b) Input stimulus and color images rendered with an asynchronous sensor implemented with the proposed pixel architecture.

implies that, operating with a frame-based camera, the amount of data to be processed is be very high, limiting the data analysis to off-line operation[40].

Recently, the authors advanced how octopus retinas can used to monitor flame activity,[39]. The excellent temporal resolution of these sensors can be exploited to track their transient activity efficiently. Filtering the radiation below the NIR band, the output data is limited to the regions exposed to NIR radiation. Hence, their output data flow can be processed real-time[39], outperforming the operation obtained with classic NIR imagers[40]. In Fig. 5 .a, a snapshot of a flame taken with an octopus asynchronous sensor fabricated in standard CMOS technology is displayed. An optical filter was placed over the sensor to remove the radiation outside the NIR band. Results are alike the ones that can be obtained with a commercial infrared camera. It is possible to identify the main regions inside the flame. On the top of the right plot in Fig. 5 .b, we have plotted the transient NIR levels detected by the sensor in a visual scene with flames. Flames provoke fast variations of the NIR levels. On the bottom of the same plot, the Fast Fourier Transform (FFT) of the previous data is displayed. Main frequency components range from DC to 30Hz, for this flame.

4.3 | High speed color detection with stacked photodiodes

Currently, standard CMOS technologies offer the possibility of incorporating a deep n-well or a deep p-well to the chip design. These layers allow to stack photodiodes at different depths, placing a deep n-well beneath an n-well, as it is depicted in Fig. 6 .a. Stacked photodiodes have different sensitivity to light wavelength, depending on their depth[47,48]. Hence, similar spectral discrimination can be achieved than employing traditional methods with discrete optical filters. In this previous work[46], there diodes were stacked to encode color at high speed. In Fig. 6 .a, there is a sketch of the color pixels. Each one is connected to an integrate-and-fire (I&F) module with the same architecture than an octopus sensor as the one shown in Fig. 3 . Processing the

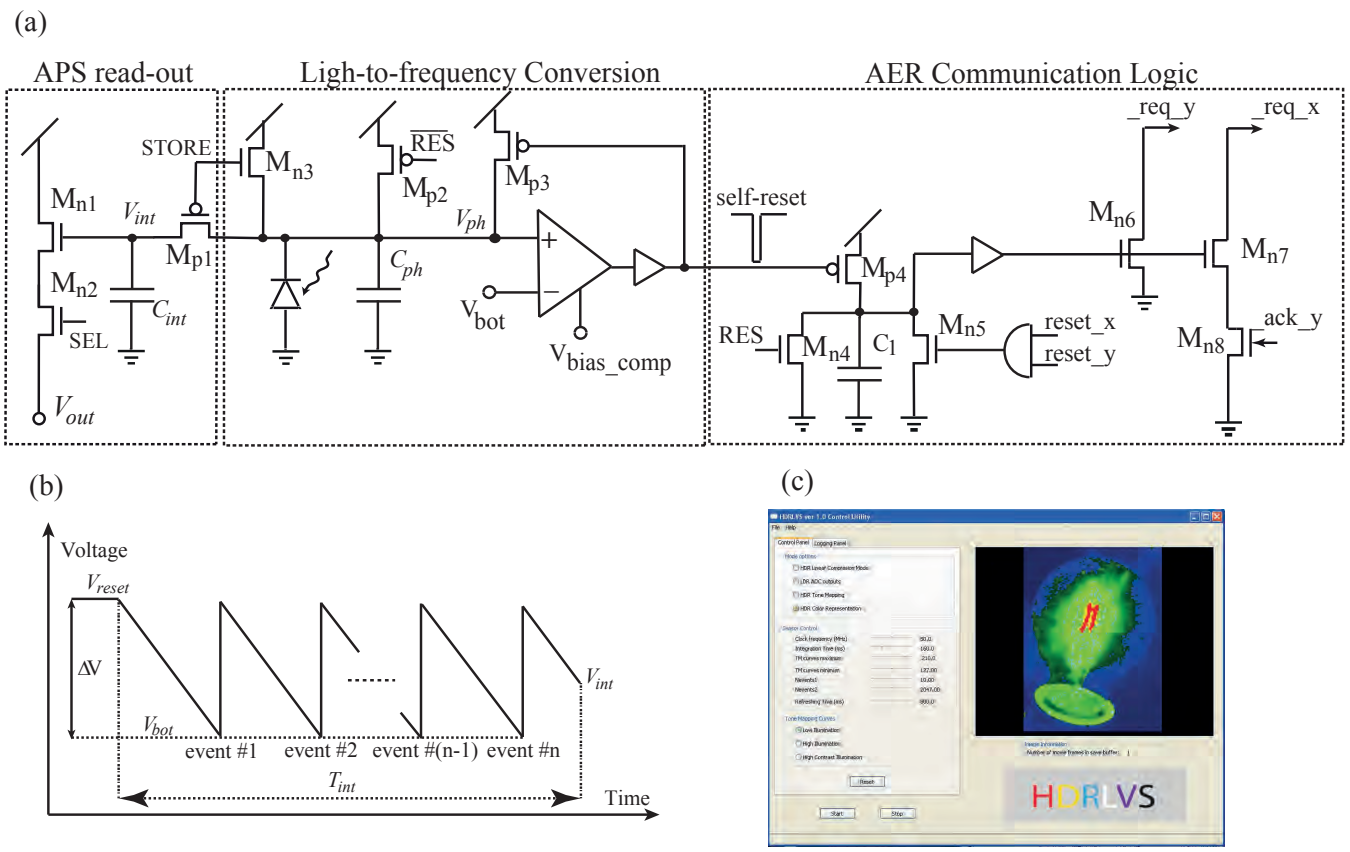

FIGURE 7 (a) Pixel schematics of the HDRLVS sensor[30]. (b) Operation principle. Pixels generate spikes every time that the voltage at the integration capacitance reaches a voltage threshold. At the end of an integration period, the remaining voltage at the integration capacitance is readout and digitized with an APS architecture. This information is combined with the number of spikes (if any) associated to each pixel to encode illumination levels. (c) Custom interface programmed to test and represent the sensor outputs. In the picture, a snapshot of a visual scene with large intra-dynamic range is shown. Intensity levels were rendered with a color scale of 24-bit depth.

outputs of each I&F block, it is possible to render real time images,[46]. On the right of Fig 6 .b, there is an input stimulus that was recorded with the sensor and the image captured with it. Although it was not possible to tune the depth of the photodiodes to adjust the peak of sensitivity of each one, tri-chromatic images could be rendered, demonstrating the feasibility of this approach.

5 | HYBRID SENSORS

Another tendency on the design of asynchronous sensors is the implementation of hybrid devices. These combine frame-based operation and event-based asynchronous operation. The approach aims to merge the advantages of both families of sensors in one unique design. The development of specific technologies for image sensor design that permit the layout of pixels with reasonable pitch, and the demand of sensors that can implement image processing on the focal plane, justifies this trend. Posch et al. implemented the ATIS sensor[16]. The device has pixels that detect motion, operating as a DVS sensor, and also measures

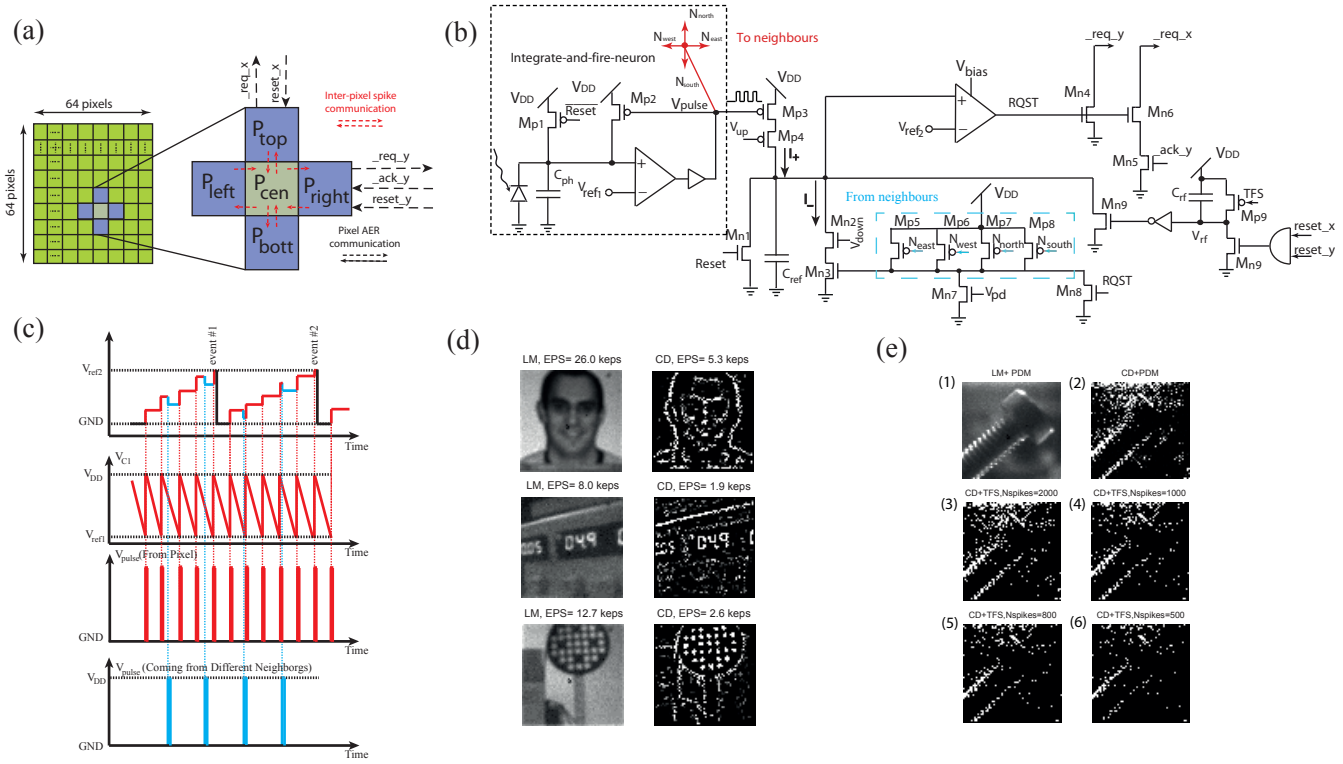

FIGURE 8 Sensor with dual operation modes and dual readout modes[31]. (a) Pixel's digital signals exchanged with the neighbors. (b) Pixel's schematics. (c) Illustration of how the pixel and the neighbors spikes increase/decrease the voltage at the reference capacitance, C_{ref}. Events are transmitted off-chip weather the pixel activity is higher than the neighborhood activity. (d) Snapshots taken with the sensor with its two operation modes: contrast detection (CD) and luminance mode (LM), combined with AER readout. (f) Snapshots taken with the sensor in contrast detection mode combined with TFS readout. The output image is rendered with different number of events.

their luminance levels whether they detect transient illumination variations. The design permits to measure all the illumination levels of the visual scene and the time stamps where they changed, minimizing the output data flow. Every time that a pixel illumination value changes, its new luminance value is gauged.

The DAVIS sensor[14] reported by Delbruck et al. combines DVS and/or APS operation. The user can toggle between the two operation modes, depending on the information of the visual scene to be captured. Pixel pitch is competitive, 18.5µm×18.5µm, for many applications, as drone vision or autonomous vehicles control, that do not require high resolution images.

The authors of this article, contributed in 2016 with a HDR image sensor (known as HDRLVS sensor) with linear operation[30]. The device has an output data format identical to any APS sensor. However, internally, its pixels combine classic APS readout with event operation. In Fig. 7 .a, pixel schematics of the HDRLVS sensor are displayed. In Fig. 7 .b, the sensor operation principle is depicted. The pixel has a light-to-frequency conversion module that generates spikes whenever the voltage at the integration capacitance, C_{ph}, reaches a voltage threshold. The coordinates of the pixel that spikes are transmitted off-chip every time that this occurs. At the end of an integration time, the remaining voltage is digitized and readout with APS circuitry.

Combining the number of events associated to each pixel (more significant bits) and the digital word that represents the remaining voltage at the integration capacitance (less significant bits), pixel illumination levels are encoded linearly with a 24-bit word. Therefore, an intra-scene dynamic range of 130dB can be achieved. In Fig. 7 .c, there is a custom interface that was programmed to test and demonstrate the sensor operation. The measured illumination levels in a visual scene with large intra-scene dynamic range are displayed. A color scale was used to represent the illumination levels encoded with 24-bit depth. A video showing the device capabilities is available on the net, [49].

There are situations where it is advantageous to trade between image quality and power and data consumption. For instance, surveillance cameras would not need to send detailed images of the visual scene as long as an intruder or a transient illumination change are detected. The authors reported an asynchronous sensor [31,50] that has two operation modes, namely luminance and spatial contrast detection. These two operation modes can be combined with two readout schemes. The first one corresponds to a classic asynchronous readout mode based on the AER communication protocol [10,11] and the second one implements (Time-to-First-Spike) TFS operation [51,52,53]. In this operation mode, the sensor's pixels are released after a global reset signal. The operator can decide how many events (spikes) are used to render one image.

The sensor allows to toggle instantaneously between the operation and the readout modes, depending on the requirements of image quality, latency, power, and bandwidth consumption. In Fig. 8 .a, pixel digital signals exchanged with its neighbors are illustrated. Pixel spikes are sent to the neighbouring pixels. Analogy, spikes coming from the neighbors are received. There is a competition between the pixel and its neighbors to charge/discharge a reference integration capacitance (C_{ref} in Fig. 8 .c). Every time that the pixel fires, a packet of charge is transferred to a capacitance. When neighbors spike, a packet of charge is subtracted from the same capacitance. If the pixel activity is higher than the neighborhood, the voltage at the capacitance will reach a voltage threshold and an spike will be sent, indicating the detection of spatial contrast. The strength of the neighbors influence can be adjusted by the user. Inhibiting the influence, the sensor behaves as an octopus retina whose pixels spike with a frequency proportional to illumination. In Fig. 8 .b and 8 .c, the pixel's schematics and its operation are depicted. In Fig. 8 .d, we display snapshots in the contrast detection and in the luminance mode, combined with classic AER readout. In Fig. 8 .e, the TFS readout mode is illustrated. A different number of events was used in each snapshot to reconstruct the resultant image in the Contrast Detection (CD) mode. It is demonstrated, that with a reduced number of events, it is possible to interpret the visual scene.

6 | COMPARISON BETWEEN DIFFERENT SENSORS FAMILIES

We summarize on Table 1 main sensor parameters for each sensor family. The key initial contribution for each sensor category and its first author are mentioned. Furthermore, we highlight the main aforementioned applications associated to each sensor type.

7 | ANALYSIS AND EXPECTATIVES

There are two main factors that facilitate the spread of asynchronous vision sensors. The first one is technology development. Transistor scalability and the development of CMOS fabrication technology makes possible to reduce significantly the pixel pitch when designing asynchronous sensors. To illustrate this, we can take as reference the first functional DVS sensor reported by Lichtsteiner et at.[13]. Its pixel pitch was 40μm×40μm. Recent improved versions, i.e. the DAVIS[14], has a pixel pitch of 18.5μm×18.5μm, including extra functionalities as APS luminance sensing. The recent publication of Son et al.[54] from Samsung reflects the growing industry interest in DVS sensors and the real possibility of fabricating DVS pixels with a competitive pitch. Modern fabrication technologies have larger number of routing metals and allow to place MiM capacitors on top of the transistor layout without requiring extra area. In connection with the technology development, the spread of 3D fabrication technologies opens a new horizon of pixel design possibilities[55]. Functional processing units could be placed in different layers than the one dedicated to light sensing. In consequence, pixel focal plane processing capabilities can be increased without reducing the pixel pitch.

Another factor that increases the development of asynchronous vision sensors is the demand of vision sensor for autonomous vehicles and drone vision. The specifications for theses systems are driven to interpret the visual scene with low consumption of power and bandwidth. Hence, asynchronous image sensors find a perfect field of application in these scenarios. Hybrid operation sensors facilitate the incorporation of spiking sensors to the market. The main reason is that their output data format can be compatible with classic frame-based image processing algorithms and displays. Thus, there is no required a deep knowledge of the sensor's inner operation from the operator's side.

8 | CONCLUSION

Applications fields for the different kinds of asynchronous vision sensors have been presented and explained. We have reviewed the main families of asynchronous sensors: vision sensors for spatial contrast detection, DVS sensors, spiking luminance sensors, and devices with hybrid operation. For each one, the most relevant application scenarios have been mentioned and analyzed. In the manuscript, we have paid special attention to the applications of spiking luminance sensors, explaining in more detail some

of the contributions made by the authors. The expectations for the further development of asynchronous sensors have been also discussed.

References

1. Mahowald M.. *An Analog VLSI System for Stereoscopic Vision*. Kluwer; 1994.

2. Posch C., Serrano-Gotarredona T., Linares-Barranco B., Delbruck T.. Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output. *Proceedings of the IEEE.* 2014;102(10):1470-1484.

3. Delbrück T., Linares-Barranco B., Culurciello E., Posch C.. Activity-driven, event-based vision sensors. In: :2426-2429; 2010.

4. Dowling John E.. *The Retina: An Approachable Part of the Brain*. The Belknap Press of Harvard University Press; 2 ed.2012.

5. Zaghloul K.A., Boahen K.. Optic nerve signals in a neuromorphic chip I: Outer and inner retina models. *Biomedical Engineering, IEEE Transactions on.* 2004;51(4):657-666.

6. Zaghloul K.A., Boahen K.. Optic nerve signals in a neuromorphic chip II: testing and results. *Biomedical Engineering, IEEE Transactions on.* 2004;51(4):667-675.

7. Leñero-Bardallo Juan A., Serrano-Gotarredona T., Linares-Barranco B.. A Five-Decade Dynamic-Range Ambient-Light-Independent Calibrated Signed-Spatial-Contrast AER Retina With 0.1ms Latency and Optional Time-to-First-Spike Mode. *Circuits and Systems I: Regular Papers, IEEE Transactions on.* 2010;57(10):2632-2643.

8. Silvilotti M.. Wiring Considerations in Analog VLSI Systems with Application to Field-programmable Networks. PhD thesisCal. Inst. of Tech.Pasadena, California1991.

9. Mortara A., Vittoz E. A.. A communication architecture tailored for analog VLSI artificial neural networks: intrinsic performance and limitations. *IEEE Transactions on Neural Networks.* 1994;5(3):459-466.

10. Boahen Kwabena A.. Point-to-point connectivity between neuromorphic chips using address events. *IEEE Trans. Circuits Syst. II.* 2000;47(5):416-434.

11. A.Leñero-Bardallo J., Pérez-Peña F., Carmona-Galán R., Rodríguez-Vázquez A.. Pipeline AER arbitration with event aging. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS):1-4; 2017.

12. Leñero-Bardallo Juan A., Serrano-Gotarredona T., Linares-Barranco Bernabe. A Calibration Technique for Very Low Current and Compact Tunable Neuromorphic Cells. Application to 5-bit 20nA DACs. *Transactions on Circuits and Systems, Part-II.* 2008;55(6):522-525.

13. Lichtsteiner Patrick, Posch Christoph, Delbruck Tobi. A 128×128 120dB $15\mu s$ Latency Asynchronous Temporal Contrast Vision Sensor. *IEEE Journal of Solid-State Circuits.* 2008;43(2):566-576.

14. Brandli C., Berner R., Yang Minhao, Liu Shih-Chii, Delbruck T.. A 240×180 130dB 3µs Latency Global Shutter Spatiotemporal Vision Sensor. *Solid-State Circuits, IEEE Journal of.* 2014;49(10):2333-2341.

15. Yang M., Liu S. C., Delbruck T.. A Dynamic Vision Sensor With 1% Temporal Contrast Sensitivity and In-Pixel Asynchronous Delta Modulator for Event Encoding. *IEEE Journal of Solid-State Circuits.* 2015;50(9):2149-2160.

16. Posch C., Matolin D., Wohlgenannt R.. A QVGA 143dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression. *IEEE Journal of Solid State Circuits.* 2010;46(1):259-275.

17. Leñero-Bardallo J. A., Serrano-Gotarredona T., Linares-Barranco Bernabe. A 3.6µs Latency Asynchronous Frame-Free Event-Driven Dynamic-Vision-Sensor. *IEEE Journal of Solid-State Circuits.* 2011;46(6):1443-1455.

18. Li Hongmin, Liu Hanchao, Ji Xiangyang, Li Guoqi, Shi Luping. CIFAR10-DVS: An Event-Stream Dataset for Object Classification. *Frontiers in Neuroscience.* 2017;11:309.

19. Hu Yuhuang, Liu Hongjie, Pfeiffer Michael, Delbruck Tobi. DVS Benchmark Datasets for Object Tracking, Action Recognition, and Object Recognition. *Frontiers in Neuroscience.* 2016;10:405.

20. Litzenberger M., Kohn B., G.Gritsch , et al. Vehicle Counting with an Embedded Traffic Data System using an Optical Transient Sensor. In: Proceedings of the 2007 IEEE Intelligent Transportation System Conference, Seattle, WA, USA:36-40; 2007.

21. Won J. Y., Ryu H., Delbruck T., Lee J. H., Hu J.. Proximity Sensing Based on a Dynamic Vision Sensor for Mobile Devices. *IEEE Transactions on Industrial Electronics.* 2015;62(1):536-544.

22. Fu Z., Delbruck T., Lichtsteiner P., Culurciello E.. An Address-Event Fall Detector for Assisted Living Applications. *IEEE Transactions on Biomedical Circuits and Systems.* 2008;2(2):88-96.

23. Lee J. H., Delbruck T., Pfeiffer M., et al. Real-Time Gesture Interface Based on Event-Driven Processing From Stereo Silicon Retinas. *IEEE Transactions on Neural Networks and Learning Systems.* 2014;25(12):2250-2263.

24. Delbruck T., Villanueva V., Longinotti L.. Integration of dynamic vision sensor with inertial measurement unit for electronically stabilized event-based vision. In: 2014 IEEE International Symposium on Circuits and Systems (ISCAS):2636-2639; 2014.

25. Drazen David, Lichtsteiner Patrick, Häfliger Philipp, Delbrück Tobi, Jensen Atle. Toward real-time particle tracking using an event-based dynamic vision sensor. *Experiments in Fluids.* 2011;51(5):1465.

26. Farian L., Leñero-Bardallo J.A., Häfliger P.. A Bio-Inspired AER Temporal Tri-Color Differentiator Pixel Array. *Biomedical Circuits and Systems, IEEE Transactions on.* 2015;9(5):686-698.

27. Domínguez-Morales Manuel Jesus, Cerezuela-Escudero Elena, Perez-Peña Fernando, Jimenez-Fernandez Angel, Linares-Barranco Alejandro, Jimenez-Moreno Gabriel. On the AER Stereo-Vision Processing: A Spike Approach to Epipolar Matching. In: Neural Information Processing:267–275Springer Berlin Heidelberg; 2013; Berlin, Heidelberg.

28. Domínguez-Morales M., Jimenez-Fernandez A., Paz R., et al. An Approach to Distance Estimation with Stereo Vision Using Address-Event-Representation. In: Neural Information Processing:190–198Springer Berlin Heidelberg; 2011; Berlin, Heidelberg.

29. Culurciello E., Etienne-Cummings R., Boahen K.A.. A biomorphic digital image sensor. *Solid-State Circuits, IEEE Journal of.* 2003;38(2):281-294.

30. Leñero-Bardallo Juan A., Carmona-Galán R., Rodríguez-Vázquez A.. A wide linear dynamic range image sensor based on asynchronous self-reset and tagging of saturation events. *IEEE Journal of Solid-State Circuits.* 2017;52(6):1605-1617.

31. Leñero-Bardallo Juan A., Carmona-Galán R., Rodríguez-Vázquez. A.. A bio-inspired vision sensor with dual operation and readout modes. *Sensors Journal, IEEE.* 2016;16(2):1-14.

32. Leñero-Bardallo J. A., Farian L., Guerrero-Rodríguez J. M., Carmona-Galán R., Rodríguez-Vázquez Á.. Sun Sensor Based on a Luminance Spiking Pixel Array. *IEEE Sensors Journal.* 2017;17(20):6578-6588.

33. Ortega P., Lopez-Rodriguez G., Ricart J., et al. A Miniaturized Two Axis Sun Sensor for Attitude Control of Nano-Satellites. *IEEE Sensors Journal.* 2010;10(10):1623-1632.

34. Xie N., Theuwissen A. J. P.. A Miniaturized Micro-Digital Sun Sensor by Means of Low-Power Low-Noise CMOS Imager. *IEEE Sensors Journal.* 2014;14(1):96-103.

35. Boldrini F., Monnini E., Procopio D., et al. Attitude sensors on a chip: Feasibility study and breadboarding activities. In: Proceedings 32nd Annual AAS Guided Control Conference:1197âĂŞ1216; 2009.

36. Farian L., Häfliger P., Leñero-Bardallo J. A.. A Miniaturized Two-Axis Ultra Low Latency and Low-Power Sun Sensor for Attitude Determination of Micro Space Probes. *IEEE Transactions on Circuits and Systems I: Regular Papers.* 2018;65(5):1543-1554.

37. Darmont Arnaud. Methods to extend the dynamic range of snapshot active pixel sensors. In: Proceedings of SPIE Vol. 6816, 681603 (2008); 2008.

38. Leñero-Bardallo J. A., Guerrero-Rodríguez J. M., Carmona-Galán R., Rodríguez-Vázquez Á.. On the analysis and detection of flames with an asynchronous spiking image sensor. *IEEE Sensors Journal.* 2018;:1-9.

39. Leñero-Bardallo J.A., Bryn D.H., Häfliger P.. Flame monitoring with an AER color vision sensor. In: Circuits and Systems (ISCAS), 2013 IEEE International Symposium on:2404-2407; 2013.

40. Yan Yong, M. Colechin, R. Hill. Monitoring of oscillatory characteristics of pulverized coal flames through image processing and spectral analysis. *IEEE Transactions on Instrumentation and Measurement.* 2006;55:226-231.

41. Jones A R. Flame failure detection and modern boilers. *Journal of Physics E: Scientific Instruments.* 1988;21(10):921.

42. Briz S., Castro A.J., Aranda J.M., Melendez J., Lopez F.. Reduction of false alarm rate in automatic forest fire infrared surveillance systems. *Remote Sensing of Environment.* 2003;:19-29.

43. Cheon Jimin, Lee Jeonghwan, Lee Inhee, Chae Youngcheol, Yoo Youngsin, Han Gunhee. A Single-Chip CMOS Smoke and Temperature Sensor for an Intelligent Fire Detector. *IEEE Sensors Journal.* 2009;9:625-628.

44. Maoult Y. Le, Sentenac T., Orteu J.J., Arcens J. P.. A New Approach Based on a Low Cost CCD Camera in the Near Infrared. *Fire Detection.* 2007;:193-206.

45. Bendiscio P., J F. Francescon, Malcovat P., Malobert F., Polett M., Valacca R.. A CMOS integrated infrared radiation detector for flame monitoring. *ISCAS.* 1998;6:625-628.

46. Leñero-Bardallo Juan A., Bryn D.H., Häfliger P.. Bio-Inspired Asynchronous Pixel Event Tricolor Vision Sensor. *Biomedical Circuits and Systems, IEEE Transactions on.* 2014;8(3):345-357.

47. Lange R.. 3D Time-of-Flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology. PhD thesisDepartment of Electrical Engineering and Computer Science at University of Siegen2000.

48. Sze S.M.. *Semiconductor Devices Physics and Technology*. AT&T Bell Laboratories: Wiley; 1985.

49. Leñero-Bardallo Juan A., Carmona-Galán R., Rodríguez-Vázquez Á.. Live Demonstration HDRLVS. In: https://www.youtube.com/watch?v=KrdpUpBRD60.

50. Leñero-Bardallo Juan A., Häfliger P.. A Dual-Operation-Mode Bio-Inspired Pixel. *Circuits and Systems II: Express Briefs, IEEE Transactions on.* 2014;61(11):855-859.

51. Ruedi P.-F., Heim P., Kaess F., et al. A 128 × 128 pixel 120dB dynamic-range vision-sensor chip for image contrast and orientation extraction. *Solid-State Circuits, IEEE Journal of.* 2003;38(12):2325-2333.

52. Barbaro M., Burgi P. Y., Mortara A., Nussbaum P., Heitger F.. A 100×100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding. *IEEE Journal of Solid-State Circuits.* 2002;37(2):160-172.

53. Shoushun Chen, Bermak A.. Arbitrated Time-to-First Spike CMOS Image Sensor With On-Chip Histogram Equalization. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on.* 2007;15(3):346-357.

54. Son B., Suh Y., Kim S., et al. 4.1 A 640×480 dynamic vision sensor with a 9μm pixel and 300Meps address-event representation. In: 2017 IEEE International Solid-State Circuits Conference (ISSCC):66-67; 2017.

55. Leñero-Bardallo J. A., Rodríguez-Vázquez Á.. Analog Electronics for Radiation Detection. In: Analog Electronics for Radiation Detection. 2016 (pp. 47-67).

TABLE 1 Comparative between different spiking sensor families

Type of sensor/ Functionality	Author / Year / Key first implementation	Key parameters	Key applications
Spatial contrast detection	Mahowald 1991, [1]	Spatial contrast sensitivity	Bio-inspired modelling, applications with reduced output data flow, surveillance, early scene interpretation
DVS sensors	Lichtsteiner 2008, [13]	Temporal contrast sensitivity	Datasets for spiking neural networks, traffic monitoring, robotic integration, domotic systems, stereo vision
Octopus sensors	Culurciello 2003, [29]	Dynamic range and latency	Light-to-frequency conversion, sun sensors, flame monitoring, fast color encoding
Hybrid sensors	Several authors 2010, [16,14,30]	Dynamic range, latency, power consumption, spatio-temporal contrast detection	Combinations of features from different sensors families, HDR imagers, motion detection