



HAL
open science

A hybrid event-based pixel for low-power image Sensing

Mohamed Akrarai, Nils Margotat, Gilles Sicard, Laurent Fesquet

► **To cite this version:**

Mohamed Akrarai, Nils Margotat, Gilles Sicard, Laurent Fesquet. A hybrid event-based pixel for low-power image Sensing. 28th IEEE International Conference on Electronics Circuits and Systems (ICECS 2021), Nov 2021, Dubai, United Arab Emirates. 10.1109/ICECS53924.2021.9665509 . hal-03662342

HAL Id: hal-03662342

<https://hal.science/hal-03662342v1>

Submitted on 9 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

A hybrid event-based pixel for low-power image sensing

Mohamed Akrrarai[†], Nils Margotat[†], Gilles Sicard^{*} and Laurent Fesquet[†]

[†] *Univ. Grenoble Alpes, CNRS, Grenoble INP*, TIMA, 38000 Grenoble, France*

firstname.name@univ-grenoble-alpes.fr

^{*} *Institute of Engineering Univ. Grenoble Alpes*

^{*} *CEA-LETI, L3S, 38000 Grenoble, France*

firstname.name@cea.fr

Abstract—This article presents a hybrid pixel dedicated to event-based image sensors, which suppresses spatial and temporal redundancies. This pixel is a combination between two kinds of pixels, known as TFS and DVS. The first one is able to reduce - associated to a dedicated readout - spatial redundancies, while the second manages the temporal redundancies. Each pixel is classically built with a photodiode but this hybrid structure shares the photodiode between the DVS and the TFS pixels. In our design, 4 TFS pixel photodiodes are shared with 1 DVS pixel, which exploits the average current of the latter. In order to evaluate such a pixel, a comparison is made between a hybrid pixel matrix and a matrix composed of TFS and DVS pixels. The evaluation is made in terms of activity and power but also by analyzing their fill-factor, circuitry and layout. A testchip including the 2 event-based image sensors has been designed in FDSOI 28 nm technology from STMicroelectronics.

Index Terms—Event-based, trigger, Dynamic Vision Sensor pixel, Time To First Spike pixel.

I. INTRODUCTION

Asynchronous image sensors introduce a new approach for image acquisition. The approach does not rely on reading the full image at a fixed framerate, like standard CMOS image sensor, but rely on sparsely reading pixels when the pixels fire themselves. Based on an asynchronous logic protocol, the pixels autonomously generate an event when detecting luminance variations or measuring it. Furthermore, such pixels own intrinsic redundancy suppression properties.

One of the pixels is the Time-to-First Spike (TFS), first presented by Shoushun & Bermak [1], and introduced an implementation of the time-based encoding. A detailed characterization of the pixel was also given in [2], and various pixel matrix implementations followed, targeting low power image sensor design, for power critical application such as wireless endoscopy [3]. Another unique asynchronous pixel is the Dynamic Vision Sensor (DVS) pixel, first presented by Delbruck [4]. This pixel detects and reacts to luminance slope changes, and suppresses temporal redundancy internally. Furthermore, in [5], the same pixel was improved with pre-readout pixel-parallel noise and local spatial redundancy suppression, enabling more data compression. The most utility that can be obtained from the DVS pixel, is by using it as a trigger pixel. In [6], the DVS pixel is used to trigger a TFS pixel, when there is actual activity (luminance change) in the scene.

Thereby, the image sensor only updates at the end the pixels in the image memory that have a relevant information. The image sensor has a dynamic range of 143 dB, which is a high dynamic range compared to standard CMOS image sensors.

Our idea is built on the work in [6]. Instead of having 1 DVS triggering one TFS pixel, the average photo-current of 4 TFS pixels is monitored by a unique DVS pixel, placed in the middle of the four TFS. When this pixel detects a significant change compared to a preset threshold, it triggers the four TFS pixels. This architecture results in less events, since averaging 4 photo-currents, will filter single photo-current noise spikes, that might generate a false event. The proposed pixel architecture is compared to a non averaging architecture to highlight its advantages.

The paper is organized as follows. In Section II we present both the DVS and TFS pixels. In section III, we report the different sub-circuits comprising our hybrid pixel architecture. In Section VI, we compare the hybrid pixel architecture, which averages photocurrents to a non averaging triggered approaches through high-level simulations, and power consumption estimations. And finally, Section V concludes on the advantages of such an approach.

II. THE DVS AND TFS PIXELS

The two pixels have different circuits and operating principles, and are the backbone of asynchronous image capture. They both are autonomous. In this section, we highlight the structure of each pixel and how they operate.

A. DVS

This pixel only reacts to a luminance intensity change. Its block schematic is presented in Figure 1a. This pixel uses a switched capacitor differentiator circuit computing the signal derivative (in time), i.e. the voltage difference between two successive samples of the photo-detector voltage, V_p . This difference is compared afterwards to two thresholds to generate ON or OFF events according to the polarity of the slope of V_p . A positive change yields to an ON Event and a negative change to an OFF event. An event is represented by the triplet (Xe; Ye; Pe) where X and Y are the pixel coordinates in the matrix and P is the event polarity (ON or OFF event). Figure 1b presents the signals operating in the pixel. In red,

a photodiode current generated with a sinus current source is taken as a simple stimulus emulating a variation of the pixel luminance. The luminance slope changes are detected in the pixel thanks to a differentiator circuit, which translates to V_{diff} spikes (Green). The latter go into two comparators to generate ON (blue) or OFF (Orange) events. The reset signal is used after the detection of an event to clear the request, and drive V_{diff} back to its reset level, before the next detection. The full circuit schematic of the pixel will be presented in the sequel. In our hybrid pixel architecture, the DVS pixel events are only used to trigger the TFS pixels when there is a sufficient luminance change. Then, the TFS pixels start measuring the luminance.

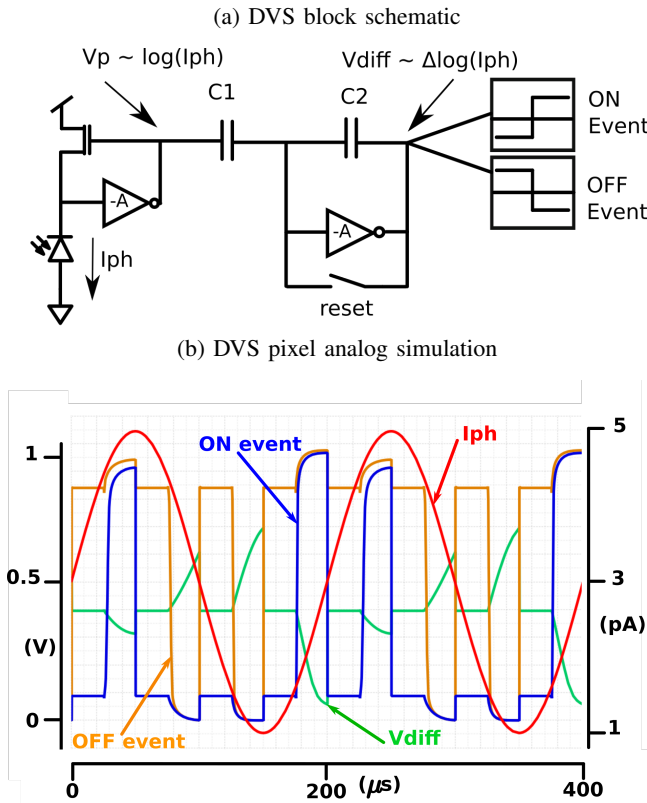


Fig. 1: DVS pixel

B. TFS

This pixel operates as a one level-crossing sampling circuit. When the photodiode voltage V_{ph} crosses an externally defined threshold, an event is generated. The pixel event is the logical address coordinates of the pixel in the pixel matrix. This is known as the Address Event Representation (AER). In order to translate this event into a gray level information, an external Time-to-Digital Converter (TDC) is used. The TDC generates the instant at which it received the event coded in a bit vector depending on the image resolution. In order to activate and communicate an event, a pixel exchanges the following signals with an external readout system (see Fig. 3 and 2):

- the pixel starts with the signals $localReset$, $Global_reset$ and ON_TFS inactive.
- Once the DVS pixel trigger is active using the ON_TFS (going to 0), the TFS pixel notifies the external readout of the start of photo-integration, by driving the two signals $RowStart$ and $ColStart$.
- Once the photodiode voltage level crosses an externally defined threshold, the comparator drives the $TFSComp$ signal to go down (active), and the latter drives the $RowReq$ and $ColReq$ signals to go up while at the same time, it saves a local request through the signal $ReqLoc$.
- The readout system (RS) detects the prior request signals on the row and column buses, and sends the acknowledge signals $RowAck$ and $ColAck$ acknowledging the pixel requests.
- After the pixel receives the acknowledgement signals, it deactivates the request signals $RowReq$ and $ColReq$.
- Afterwards, the readout system proceeds to the verification process, by setting the $RowVerif$ and $ColVerif$ signals. The pixel receives the latter signals and checks its local request memory. If the $ReqLoc$ signal is up, the response signal $PxPos$ will be pulled down to signal the presence of a request. If instead there is no local request, and the pixel is still integrating, signal $PxOn$ will go up, to signal the ongoing integration.

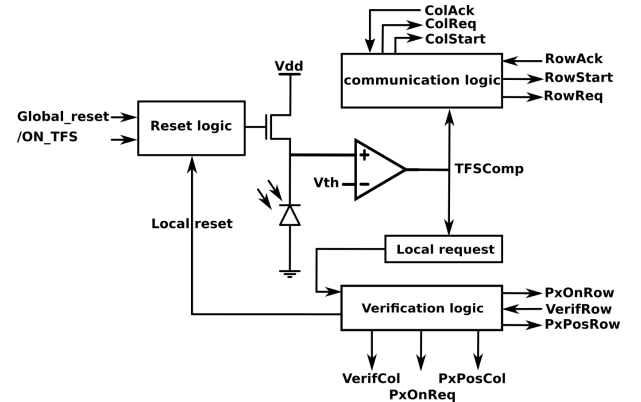


Fig. 2: TFS pixel block schematic

Figure 4 illustrates an example of address of AER address recovery. In the figure, both pixel in red (2,3) and (3,2) generate AER events through the signal buses $RowReq$ and $ColReq$, which writes at the output of the AER interfaces $RowReq(0, \dots, 1, 1, 0)$ and $ColReq(0, 1, 1, \dots, 0)$. However, when the readout system will receive the AER address, it interprets the address, as if pixels (3,2), (3,3), (2,2) and (2,3) have all fired. To mitigate this, the readout system checks the local request of each pixel amongst the four, to verify if they truly fired or not. The verification process is mandatory since, our AER request communication is used without any arbitration, which is the strength of our approach. Using arbiters introduces many issues, such as a limited event bandwidth, a reduced scalability and inaccurate luminance measurements. The verification process solves the AER address conflicts. Even if the

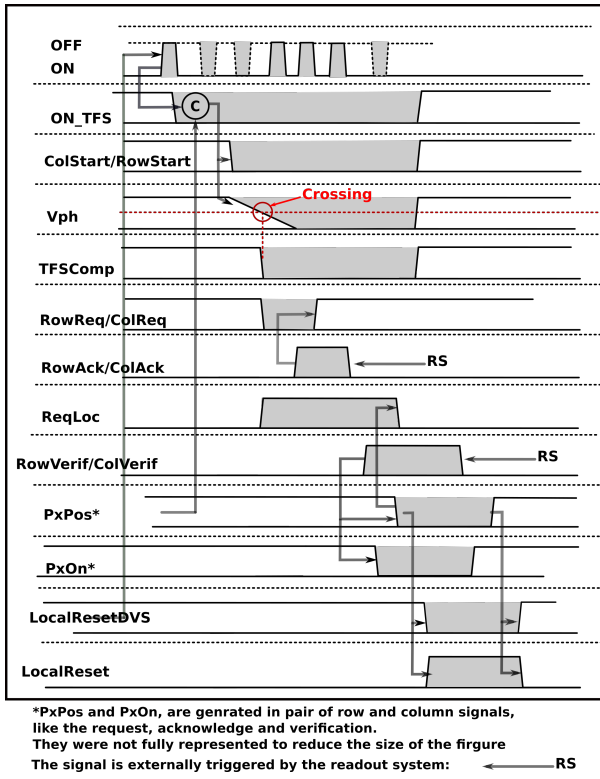


Fig. 3: TFS pixel communication signals

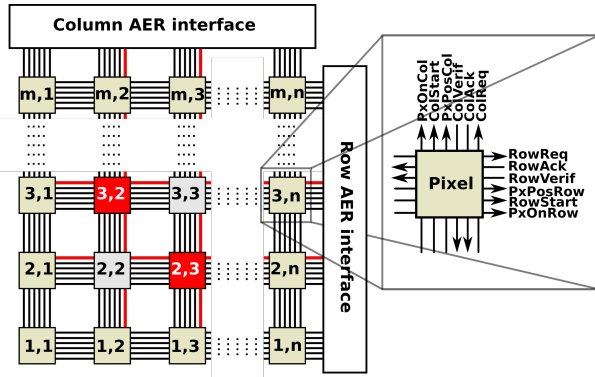


Fig. 4: Pixel matrix view illustrating the need for verification

verification is a sequential process, this approach is particularly interesting because it suppresses the risk of time measurement errors due to the arbitration tree. Indeed, this latter delays the requests inducing an inaccurate time stamping. Notice that the verification sequence is really fast compared to the photodiode integration time, which is several order of magnitudes slower. Thus, the verification speed is not an issue for such image sensors.

III. THE HYBRID PIXELS

In this Section, we present the hybrid pixels and the averaging circuit schematics and layouts. Using one DVS pixel - without photodiode - to trigger the surrounding 4 TFS pixels,

has many advantages over a non averaging trigger approach. This is carefully examined in the sequel.

A. The averaging pixel architecture

As we already mentioned in the paper, the circuit is a DVS pixel carefully placed and routed in the middle of the four TFS pixels. This DVS pixel has no photodiode. This latter has been advantageously replaced by the average photo-current of the 4 surrounding TFS photodiodes. Figure 5 below displays the block schematic of the global circuit. The circuit contains the 4 TFS pixels with their photodiodes, a DVS pixel without a photodiode but with the averaging circuit, and some asynchronous logic. Each component is detailed in separate subsections.

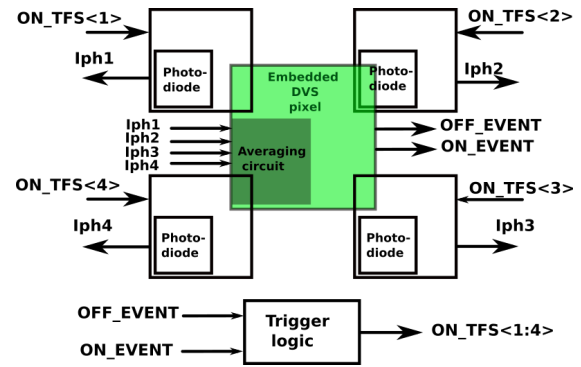


Fig. 5: Abstracted full circuit schematic

B. the averaging circuit

The averaged photo-currents are copied using 4 current-mirrors on each TFS pixel photodiode. The sum of the currents is then fed to the DVS pixel, thanks to a final current-mirror circuit. All current-mirrors are balanced. Thus, the averaged photo-current sum is multiplied by a gain G_{cm} of 4.

$$I_{dvs} = G_{cm} \frac{\sum_{n=1}^4 I_{ph_n}}{4} \quad (1)$$

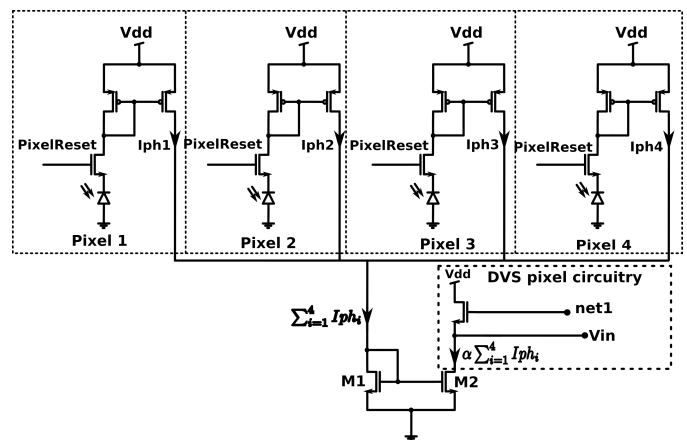


Fig. 6: Averaging circuit (net1 and Vin link to the DVS pixel)

Figure 6 shows how the current mirrors are linked to the photodiode of the pixels. In each TFS pixel, the current-mirrors are PMOS transistors, in order to be mounted on top of

the reset transistors. On the other hand, the DVS pixel receives the sum of the photo-currents using an NMOS current-mirror replacing its "own" photodiode.

C. The DVS pixel part

The circuitry of the pixel is completely standard to [4]. Figure 7 below presents the full schematic of the pixel.

- Transistors M1 and M2 constitute the last stage of the current mirror, form the past circuit.
- M4, M5 and M6 form an inverting amplifier with external polarization for testing purposes, which feeds back to the logarithmic compressor M3.
- M7 and M8 form a source follower isolating the previous stages of the circuit from the switching of the next stage (differentiator circuit).
- Transistors M9, M10, M11 and capacitors C1 and C2 constitute the switch capacitor differentiator circuit performing the luminance change detection.
- The ON and OFF comparators composed of M12, M13 and M14, M15 output the ON and OFF events, by comparing V_{diff} against global thresholds V_{off} and V_{on} .

In the pixel, the amplification gain is given by the capacitor ratio C_1/C_2 . At the differentiator output, the luminance is expressed as follows:

$$V_{diff} \sim \frac{C_1}{C_2} \cdot \Delta \log(I_{ph}) \quad (2)$$

Note that gain introduced prior to the logarithmic compressor stage is cancelled by the differentiator circuit.

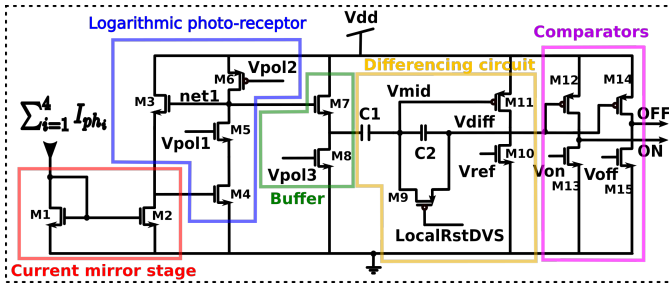


Fig. 7: DVS pixel schematic

D. The 4 TFS pixels

The schematic of figure 8 is divided into three sections. The first one is the analog part of the pixel, the second is the digital part, and the last represents the bus connections.

In the first part:

- The first transistor M1 resets the photodiode whenever it receives the *PixelReset* signal from the digital part of the pixel.
- Transistors M2 and M3 constitute the current-mirror copying the photo-current from the photodiode to the averaging circuit.
- M4, M5, M6 and M7 constitute the comparator circuit. It detects the threshold crossing $V_{Complvl}$ after the photo-integration starts (once started, the photodiode voltage decreases).

- M8, M9, M10, M11, M12 control the reading and writing of the local request memory, through the acknowledgement and verification signals.
- M13, M14, M15 and M16 compose the local request memory, which is no more than two feedback inverters.
- Mp is used to bias the comparator through an external voltage.
- Mr resets the memory when a reset signal is emitted by the digital pixel part.

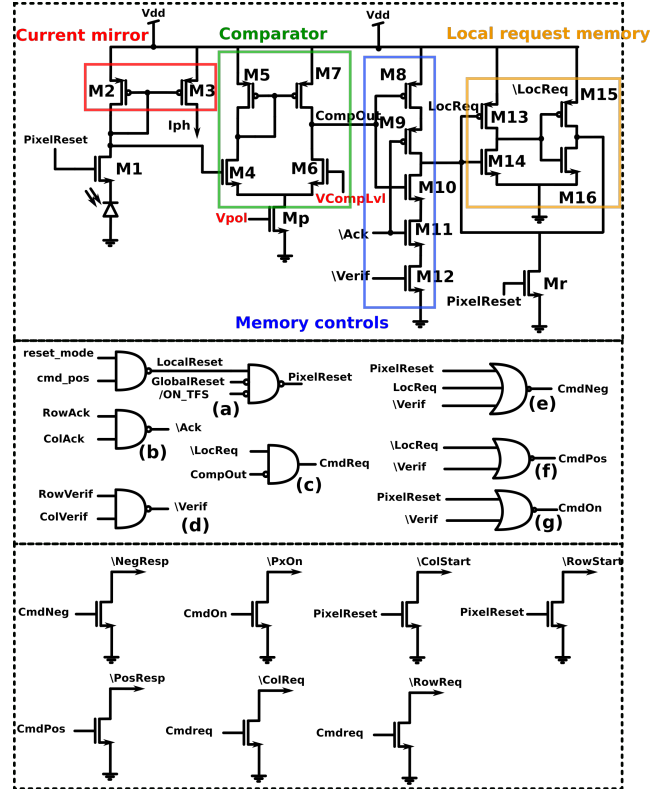


Fig. 8: TFS pixel schematic

The second part of figure 8 is the digital circuitry of the pixel. The NAND gates (b) and (d) synchronize both column and row signals, for verification and acknowledge, by selecting the pixel. The logic gates (c), (e), (f) and (g) all send the command signals (*CmdReq*, *CmdNeg*, *CmdPos* and *CmdOn*) to the bus transistors, in order to pull-down the adequate buses. The last 2-gate circuit (a) combines all the signals triggering the *PixelReset* signal. *ResetMode* selects the operating mode of the pixel, either a local reset mode where the pixel resets itself using the *LocalReset* signal when it fires, or a global reset mode, where all the pixels are externally reset through the *GlobalReset* signal. The *ON_TFS* signal removes the *PixelReset* when a request is coming from the DVS pixel to start the photo-integration.

Finally, the third part of Figure 8 shows all the TFS pixel NMOS transistors pulling down the buses when a pixel needs to send a request to the readout system. At the bottom of every pixel matrix row and column, there is a PMOS transistor for

each bus. This latter pulls up the the buses when there is no pixel requests in the rows and columns.

E. The trigger

This circuit (see Fig.9) connecting the DVS pixel and the 4 TFS pixels, is comprised of 4 C-elements, one for each TFS pixel (a), it is an asynchronous logic circuit memorizing the DVS trigger signal. The trigger also comprise a circuit that controls of the DVS pixel reset (b). Sub-figure (c) details the signals sequence of a trigger:

- The circuit starts functioning when both $\backslash DVSLocalReset$ and $\backslash ResetDVS$ are inactive (high). Once the DVS detects an event, ON or OFF signals start to output events, then the signal $\backslash ON_TFSi$ becomes low through the C element, and all four TFS pixel starts photo-integration.
- The first TFS pixel to successfully integrate and save an event, will be verified by the readout system. And when verified, the signal $CmdPos$ will go up to confirm the presence of a local request. Following this, the signal $\backslash DVSLocalReset$ is pulled down and $LocalReset$ goes up, so that both the DVS and the firing pixel are reset.
- The 3 other TFS pixels continue integrating until they fire an event. Then they follow the same process described above.

By switching the $Reset_mode$, the pixels work in global reset mode. They follow the same trigger and integration process, except that the TFS pixels are prevented from auto-resetting, and wait for the global synchronization signal $GlobalReset$ to start another cycle.

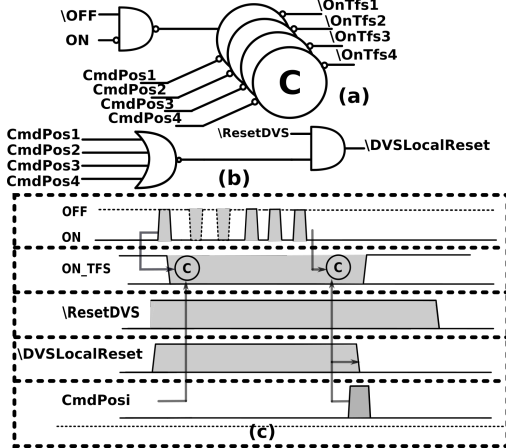


Fig. 9: Trigger logic circuit and signals diagram

IV. CIRCUIT DESIGN AND POWER ESTIMATION

The design of the pixel architecture described in the past sections, was done in 28 nm FDSOI technology from STMicroelectronics. Here, we present the physical implementation of our pixels, and the circuits.

A. Pixels and circuits design

In order to validate the advantages of an average trigger pixel architecture, we designed another pixel matrix with a

trigger architecture that does not use an averaging circuit, but solely relies on a full DVS pixel with it's photodiode. In this non averaging architecture only 1 DVS pixel and 3 TFS pixels are routable in the same area as the averaging architecture. Figure 10 displays the layouts of both averaging and non-averaging architectures. We can see that on the left, the averaging architecture enables the insertion of one more TFS pixel (we presented the 1 DVS / 4 TFS architecture in the previous sections), compared to the non-averaging architecture. This is mainly due to the need of maintaining a symmetrical photo sensitive area implementation. Which increases the fill factor (table I) for the averaging architecture.

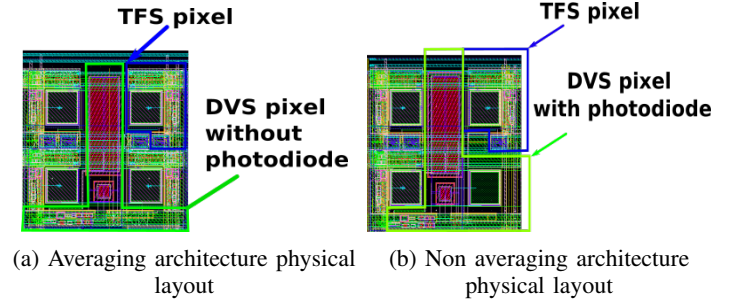


Fig. 10: The two designed pixel matrices

	Averaging	Non averaging
Photo-measurement area (μm^2)	3x3x4	3x3x3
Area(μm^2)	16.5x15	16.5x15
Fill factor (%)	14.5	10.4

TABLE I: Averaging and non averaging physical implementation comparison

throughout the rest of the paper, we will refer to the averaging pixel architecture with 1 DVS and 4 TFS pixels as the kernel (Av), the non averaging architecture will be referred to as kernel the (NAv).

B. Determination of the Event Rates

The power consumption of each pixel architecture, was estimated using post layout simulation data of the TFS and DVS pixels, as well as the event rate estimation of there high-level pixel matrices simulations for different kinds of stimuli. In *matlab*, we designed behavioral models of the two pixels to serve as building blocks of the higher level model of the pixel matrices. The models of the pixel matrices were fed by stimuli extracted from two scenarios: the first one is a car driving on the road (High activity scene 11a) and the second is a parked car with almost no activity, only a person walking in front of the car for a few seconds (Low activity scene 11b). The duration of stimuli is 10 seconds in each scenario.

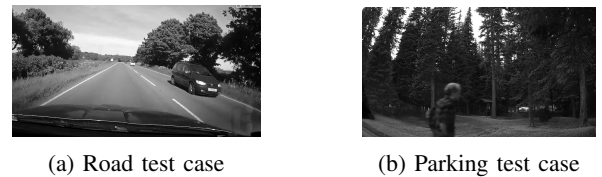


Fig. 11: The simulation test cases

The *matlab* simulations output the following: kernel (NAv), generated 168k events (road) and 4,4K events (parking), while kernel (Av) generated 78K events (road) and 2,1K events (parking) (see figure 12). The data collected from the simulations were used to calculate the total power consumption of the pixel matrices (table II). The collected data from the matlab simulation also yields an approximation of the Signal to Noise Ratio (SNR) of our pixel architectures: for the non-averaging (NAv), the SNR is approximately 11 dB, while for the averaging pixel architecture the SNR is 20 dB. This can be mostly attributed to the area difference of the photo-sensitive part and to the averaging.

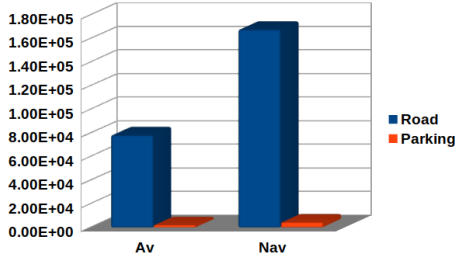


Fig. 12: Total numbers of events generated during 10 seconds of stimulus

C. Image Sensor Power Estimation

The estimated power consumption of the two pixels was obtained by dividing the total number of events of the original image (with a larger resolution) by the resolution of each pixel matrix. Indeed, our pixel matrix in Figure 10a has an effective resolution of $4 \times 32 \times 32$ pixels (32×32 kernels with 4 TFS pixels each), while pixel matrix in Figure 10b has an effective resolution of $3 \times 32 \times 32$ (32×32 kernels with 3 TFS pixel each). This yields to the average number of events generated for a single TFS pixel. The pixels are simulated post-layout to measure their power consumption and the result is multiplied by the number of events to approximate the power consumed by each DVS/TFS pixel. This average power consumption per pixel is then multiplied again by the resolution of the pixel matrices. Table II summarizes the power consumption of the pixels and matrices. Even though the averaging architecture is a few more complex (1 DVS pixel + 4 TFS pixels + an averaging block and a trigger) compared to the non-averaging architecture (that only contains 1 DVS pixel + 3 TFS pixels and the trigger), the table II highlights that the averaging pixel architecture reduces the event throughput, consumes less with high activity scenarios and a little bit more during low activity scenes. This is due to the photodiode integration of the TFS pixels, which becomes during the quiet scenes.

	Non-averaging		Averaging	
	Road	Parking	Road	Parking
DVS (μ W)	13.2	13	13	12.9
TFS (μ W)	14.13	4.15	8.54	4.62
Total pixel matrix power consumption (mW)	53.5	22.7	48.3	32.2

TABLE II: Power consumption results

V. CONCLUSION

The presented hybrid event-based architecture, is based on the combination of two event-based pixels (TFS and DVS). This architecture has proven limiting the firing rate of the pixels, since it only measures luminance when there is relevant information in the scene. This architecture also enhances the fill factor of the pixel matrix, because the full photo-sensitive area is used for measuring luminance compared to the non-averaging pixel architecture, which is only uses 3 of the 4 photodiodes for luminance acquisition. The hybrid pixel also shows better performances in terms power consumption when there is activity in the scene. This is due to the reduced number of events produced by this new pixel. The image quality is clearly increased because there is 33% more photodiodes in the averaging matrix for the same area with no geometrical distortion, thanks to a regular padding. Notice that the SNR is also improved by the averaging kernel.

Two image sensors, one with the hybrid averaging kernel and the other one with the non-averaging kernel, have been designed in FDSOI 28 nm from STMicroelectronics. The testchip is currently under fabrication and will be integrated on a camera testboard for measurements, which could confirm our simulations and analyzes more easily and more accurately the power consumption with several scenarios.

REFERENCES

- [1] Chen Shoushun and A. Bermak, "A low power CMOS imager, based on time-to-first-spike encoding and fair AER," 2005 IEEE International Symposium on Circuits and Systems, Kobe, 2005, pp. 5306-5309 Vol. 5.
- [2] X. Guo, X. Qi and J. G. Harris, "A Time-to-First-Spike CMOS Image Sensor," in IEEE Sensors Journal, vol. 7, no. 8, pp. 1165-1175, Aug. 2007, doi: 10.1109/JSEN.2007.900937.
- [3] M. Zhang, A. Bermak, X. Li and Z. Wang, "A low power CMOS image sensor design for wireless endoscopy capsule," 2008 IEEE Biomedical Circuits and Systems Conference, Baltimore, MD, 2008, pp. 397-400, doi: 10.1109/BIOCAS.2008.4696958.
- [4] P. Lichtsteiner, C. Posch and T. Delbruck, "A 128×128 120 dB 15 μ s Latency Asynchronous Temporal Contrast Vision Sensor," in IEEE Journal of Solid-State Circuits, vol. 43, no. 2, pp. 566-576, Feb. 2008, doi: 10.1109/JSSC.2007.914337.
- [5] C. Li, L. Longinotti, F. Corradi and T. Delbruck, "A 132 by 104 10μ m-Pixel 250μ W 1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression," 2019 Symposium on VLSI Circuits, 2019, pp. C216-C217, doi: 10.23919/VLSIC.2019.8778050.
- [6] C. Posch, D. Matolin and R. Wohlgenannt, "A QVGA 143dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression," 2010 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, 2010, pp. 400-401, doi: 10.1109/ISSCC.2010.5433973.
- [7] A. Darwish, L. M. G. Rocha, L. Fesquet and G. Sicard, "Design of a fully asynchronous image sensor reading system," 2015 Conference on Design of Circuits and Integrated Systems (DCIS), Estoril, 2015, pp. 1-5, doi: 10.1109/DCIS.2015.7388583.