- APROVIS3D -

Analog PROcessing of bioinspired VIsion Sensors for 3D reconstruction

Document Reference:		
Title: **D2.1.1 Stereo Event Cameras**		
Contractor: **IMSE**		
Prepared by: **Teresa Serrano-Gotarredona**		
Document Type: **Deliverable**		
Version: **1.1**		Pages: 29
Classification: **External document**		

Document Track

Version	Date	Remarks and Authors
1.0	25/02/2022	First draft (T. Serrano - IMSE)
1.1	22/05/2022	Final version (T. Serrano - IMSE)

Authors

	Role / Function	Name	Organisation
Prepared by	WP2 Leader	T. Serrano-G	IMSE
Checked by	Quality check		
Released by	Project Coordinator/WP0L	J. Martinet	UCA
Approved by	Project Coordinator/WP0L	J. Martinet	UCA

TABLE OF CONTENTS

FIGURES

TABLES

1 Introduction

Among the collection of spiking vision sensors, the DVSs (dynamic vision sensors) due to their low pixel mismatch, high dynamic range, high pixel bandwidth, low latency, low power consumption and temporal redundancy data suppression properties have become very popular and have raised the industrial interest. Several spin-off companies (PROPHESEE [1], CELEPIXEL [2] or INIVATION [3,4]) have put in the market high-resolution prototypes and other established companies as Samsung [5] and Sony (https://www.sony.net/SonyInfo/News/Press/202002/20-0219E/), have also developed DVS prototypes. Megapixel DVSs allow performing high-resolution detection of small objects while keeping a broad field of view. However, the data rate produced by the sensor may saturate the sensor output bandwidth especially if the observed image has high dynamic contents as it is the case of fast car driving. Furthermore, as the sensor resolution increases, it also increases the size and energy consumption of the subsequent vision processing systems. Alternatives to improve the sensor bandwidth have been proposed as communicating in parallel the events of a group of pixels [5]. Other alternatives aim at reducing the amount of sensor output data by including a spatio-temporal correlation logic block shared by a group of pixels to filter out noisy pixels as well as reduce spatial redundancy [4] or by including a digital event signal processor at the sensor output that can be programmed with some spatio-temporal filtering capabilities [1]. These late alternatives [1,4] aim to increase the sparseness in the output data but the size complexity of the subsequent processing system remains the same. This dilemma is solved in some biological systems. Humans are equipped with *foveal vision* [6].

Our eyes have a foveal region equipped with a high density of color sensitive photodetectors where a high resolution (HR) vision is sensed. The photoreceptor density decreases logarithmically towards the peripheral region where color insensitive low resolution (LR) scene parts are acquired [7]. High-speed localization of objects of interest is performed using computation on peripheral low-resolution regions (also known as the 'where' computation) to control the eye saccade movements that center the foveal fixation point on the target region that will be processed with high resolution for further observation and processing by the recognition 'what' subsystem. Not only are less the retina resources dedicated to acquiring the peripheral regions, but also the cortical area of computational resources dedicated to process the peripheral regions are less. Whereas the foveal region corresponds to only 5º of the visual field, 25% of the cortical computational resources are dedicated to process it. Figure 1(a) illustrates the proposed electronically multi-foveated vision system. The top subfigure in Figure 1(a) illustrates a moderate resolution image with 259,2k pixels. The lower subfigure illustrates a x64 lower resolution representation with 4k pixels that would be used by the LR Deep-Net attention subsystem in Figure 1(b) to localize the ROIs (regions of interest) in the visual field and control electronically the foveation regions where a high resolution object recognition is done simultaneously in parallel over small regions by the HR recognition subsystems, reducing the computation demands that would be needed for processing the full image.

The use of foveated vision to achieve efficient object recognition of high-resolution target objects combined with a wide low resolution vision field allowing localizing the target in peripheral regions with low computational load has been previously proposed. In the robotic vision community, hardware systems combining several cameras with different vision fields have been proposed [8,9]. However, these systems require mechanical control to move the central point of the foveation sensor which increases the system delay and power consumption. Other systems are based on the use of high-resolution sensors and a foveal transformation is done in subsequent processing stages to save computation power and energy. However, these systems do not alleviate the data

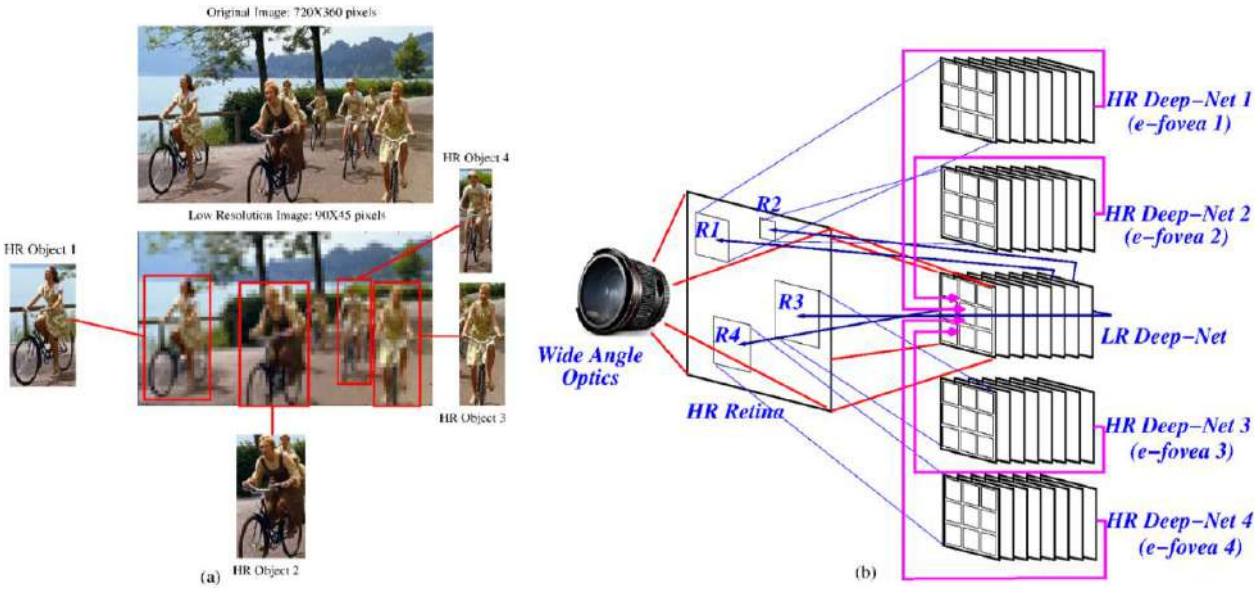


Figure 1. (a) Illustration of a multifoveal resolution sensor, and (b) proposed vision system with electronic control of multiple foveation points and parallel object recognition subsystems.

bandwidth communication at the sensor level [10,11]. There have been some proposals of sensors with foveal physical geometry; many of them suffer from non-optimum use of the silicon area [12–14]. A sensor with uniform high resolution foveal pixels in the center and low resolution motion detection peripheral pixels has been proposed [15], however it still needs a mechanical control of the center of foveation. A foveated sensor with electronic control of the foveation point where spatial resolution is traded off for temporal resolution has been published [16]. However, it does not reduce the output data rate of the sensor. In **APROVIS3D**, we have implemented an electronically foveated dynamic vision sensor EF-DVS. The proposed sensor is equipped with a foveation mechanism where nonfoveated regions save power and data bandwidth in favor of foveated high resolution regions. The foveation regions are configurable with external electronic signals. Multiple foveation points and region sizes can be electronically activated. This multi-foveation capability could allow recognition capabilities that can outperform human vision in tasks where several ROIs should be simultaneously attended as real-time surveillance systems, automatic driving, or autonomous flying drones.

2 Foveated Pixel

The pixel design is based on a DVS pixel previously reported[17]. In the next subsection we briefly describe the operation principle of a single DVS pixel. Afterwards, we describe how to combine or group neighboring pixels into a single low-resolution pixel that averages the information of the grouped pixels while keeping the bandwidth consumption and data rate of a single pixel.

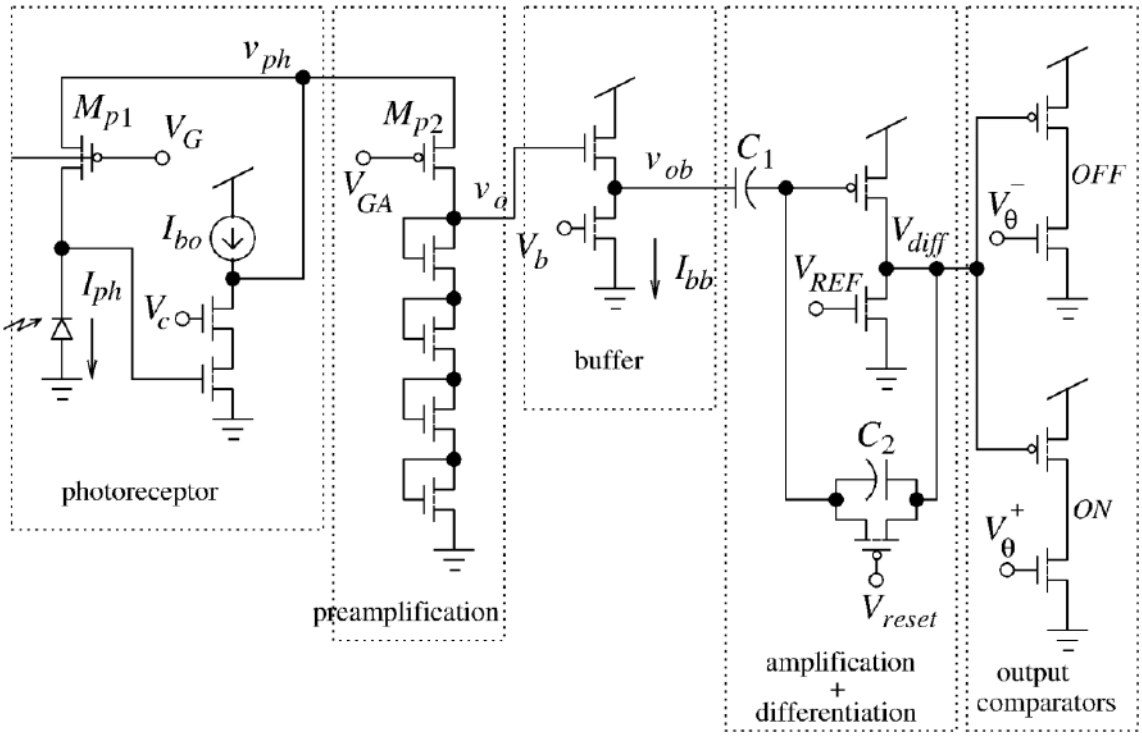


Figure 2. DVS pixel schematic

2.1 DVS Pixel

Figure 2 plots the block diagram of the DVS pixel which is based on a simplified version of a previously reported pixel[17]. The pixel consists of a photoreceptor stage that transduces the photocurrent generated in a photodiode to a voltage, a pre-amplification stage that introduces a voltage gain, a buffer to decouple the photo voltage from the capacitive input of a capacitive voltage amplification and differentiation stage. Then, the amplified voltage is compared to two voltage thresholds to generate the ON and OFF output events. Further circuitry (not shown in Figure 2) drives the row and column request and receives acknowledgement signals to communicate the events to peripheral arbitration and decoding circuits.

As shown in Figure 2, the photoreceptor block is composed of a photodiode connected to a source-driven active stage that generates a voltage V_{ph} which is logarithmic with the photocurrent, such that

$$V_{ph} = \frac{V_G}{n_p} + U_T log \frac{I_{ph}}{I_{sp}} \qquad (1)$$

Where n_p and I_{sp} are the sub threshold slope factor and current factor of transistor M_{p1}. Transistors M_{p1} and M_{p2} form a current mirror which, when operating in sub threshold region, can introduce a current gain A_I that can be adjusted through the voltage difference $\Delta V_P = V_G - V_{GA}$, such that the current flowing through transistor M_{p2} can be expressed as,

$$I_{phA} = I_{ph} e^{\frac{V_G - V_{GA}}{n_p U_T}} = A_I I_{ph} \qquad (2)$$

The pre-amplification stage is composed of the voltage to current conversion transistor M_{p2} connected to a stack of 4 NMOS diode connected transistors that generates an output voltage v_o which is approximately equal to

$$V_o = 4 n_n U_T log \frac{I_{phA}}{I_{sp}} \qquad (3)$$

So that voltage v_o is also logarithmic with the photocurrent but a voltage amplification factor approximately equal to the number of stacked diodes connected transistors can be achieved. This amplified voltage is buffered to v_{ob} to avoid slowing down time constant due to coupling with capacitor C_1.

Voltage v_{ob} is the input to an amplification and differentiation stage as proposed by Delbruck[18], so that

$$\Delta V_{diff} = -\frac{C_1}{C_2}\Delta V_{ob} \cong -\frac{C_1}{C_2}\Delta V_o = -\frac{C_1}{C_2} 4 n_n U_T log \frac{I_{ph2}}{I_{ph1}} = -A_T U_T log \frac{I_{ph2}}{I_{ph1}} \quad (4)$$

Where, A_T is the total voltage gain introduced by the pre-amplification and the amplification and differentiation stages.

Voltage V_{diff} is used as input to the ON and OFF output comparators. Whenever output ON/OFF current comparators are activated, the pixel generates a positive/negative output event, and voltage V_{diff} is reset to its resting reset value such that

$$I_{sp} e^{\frac{n_p V_{dd} - V_{rest}}{n_p U_T}} = I_{sn} e^{\frac{V_{REF}}{n_n U_T}} \quad (5)$$

Whenever there is a voltage excursion ΔV_{diff} *such* that

$$\Delta V_{diffON} = -\frac{n_p}{n_n} V_{\theta R}{}^{+}$$
$$\Delta V_{diffOFF} = +\frac{n_p}{n_n} V_{\theta R}{}^{-} \quad (6)$$

where $V_{\theta R} = \left| V_{REF} - V_{\theta}{}^{+/-} \right|$, an ON/OFF output event is generated and voltage V_{diff} returns to its reset value.

Turning to the input, let us call the minimum contrast stimulus that generates a single positive event through output channel ON "*ON contrast threshold*" (or "ON contrast sensitivity") $\theta_{ev}^{+} > 0$, and the minimum contrast stimulus that generates a single negative event through output channel OFF "*OFF contrast threshold*" (or "OFF contrast sensitivity") $\theta_{ev}^{-} > 0$.

$$\theta_{ev}^{+} = log \frac{I_{bright}}{I_{dark}} = \frac{\frac{n_p}{n_n} V_{\theta R}{}^{+}}{A_T U_T}$$
$$\theta_{ev}^{-} = log \frac{I_{dark}}{I_{bright}} = -\frac{\frac{n_p}{n_n} V_{\theta R}{}^{-}}{A_T U_T} \quad (7)$$

The minimum detectable stimulus contrast can be adjusted through contrast sensitivity control voltages $V_{\theta R}$.

2.2 Pixel Grouping

In this section we study how we can combine the output of different pixels in a way that the combined output represents the temporal variation of the spatial average of the individual pixel photocurrents in the neighborhood but reducing the output activity of the pixel group compared to the total output activity of the individual pixels.

We have considered several options to combine the local photocurrent information obtained by individual pixels into combined meaningful low-resolution information that averages the local information in the neighborhood while saving output bandwidth, reducing noise and keeping the dynamic range:

1. Interconnect the pixels at the output node of the voltage buffers v_{ob}

When interconnecting pixels at node v_{ob} we obtain (where G is the number of pixels per group)

$$\sum_{i=1}^{G} I_{sn} e^{\frac{v_{oi} - v_{ob}}{n U_T}} = G I_{bb} \quad (8)$$

This is the equation of a multi-input differential pair, where the pixel with the highest voltage v_{oi} is going to take most of the bias GI_{bb} current of the multiple input differential pair and is going to drive the voltage at the output node v_{ob}. Consequently, the output node v_{ob} is going to follow only the changes in the maximum photocurrent from the group of pixels.

2. Adding voltages V_{phi} (similarly adding voltages v_{oi} or v_{obi} using some voltage adding circuitry) to keep a logarithmic relation with the photocurrent

As we know that V_{phi}, V_{oi}, V_{obi} are proportional to log(I_{phi}), so that any of this voltages is proportional to the logarithm of the photocurrent. Let us define a generic voltage V_{logi} so that $V_{logi} \alpha \log(I_{ph\_final}/I_{ph\_initial})$,, and consider the added voltage $V_{log}=\sum_{i=1}^{K} V_{logi}$

Then,

$$\Delta V_{log}=$$

$$V_{\log\_final} - V_{\log_{initial}} = (\sum_{i=1}^{G} V_{logi_{final}} - \sum_{i=1}^{G} V_{logi_{initial}}) \propto (\sum_{i=1}^{G} \log(I_{phi_{final}}) - \sum_{i=1}^{G} \log(I_{phi_{initial}})) =$$

$$log \frac{\prod_{i=1}^{G} I_{phi_{final}}}{\prod_{i=1}^{G} I_{phi_{initial}}}$$

(9)

The effect of adding the voltages of the individual pixels is equivalent to multiplication of their photocurrents. The pixel group will be sensitive to the relative variation of the product of all the photocurrent in the pixel group. The sensitivity of the group is proportional to the total number of pixels in the group; thus the event frequency generation would be multiplied by the number of pixels for the same photocurrent relative variation and the output bandwidth consumed by the group in low resolution mode will be the same that the one of all the individual pixels in high-resolution mode.

A practical simple circuitry for adding the pixel voltages would be interconnecting the floating node of the C_1/C_2 capacitors of the group pixels. In this case,

$$\Delta V_{diff} = -\frac{C_1}{C_2} \sum_i \Delta V_{obi} \cong -\frac{C_1}{C_2} 4 n_n U_T log \frac{\prod_{i=1}^{G} I_{phi_{final}}}{\prod_{i=1}^{G} I_{phi_{inicial}}} = -A_T U_T log \frac{\prod_{i=1}^{G} I_{phi_{final}}}{\prod_{i=1}^{G} I_{phi_{inicial}}} \quad (10)$$

This pixel grouping results in an increased contrast sensitivity for the pixel group.

3. Interconnecting nodes v_{oi}. In this case, the resulting v_o voltage of the interconnected pixel group can be expressed as

$$V_o = 4 n_n U_T log \frac{\sum_{i=1}^{G} A_I I_{phi}}{G I_{sp}} \quad (11)$$

In this case, it can be noted that an averaging of the total current variation over the pixels in the group is obtained, as desired. Figure 3.(a) depicts the modified schematic of a DVS pixel that allows the interconnection of pixels at nodes v_{oi}. Digital signals m_{up} and m_{right} control pixel reconfiguration. When both of them are deactivated, the pixel operates individually in high resolution mode. However, by activating digital signals m_{up} and m_{right}, node v_o can be interconnected with neighboring pixels implementing the averaging of the amplified photocurrents.

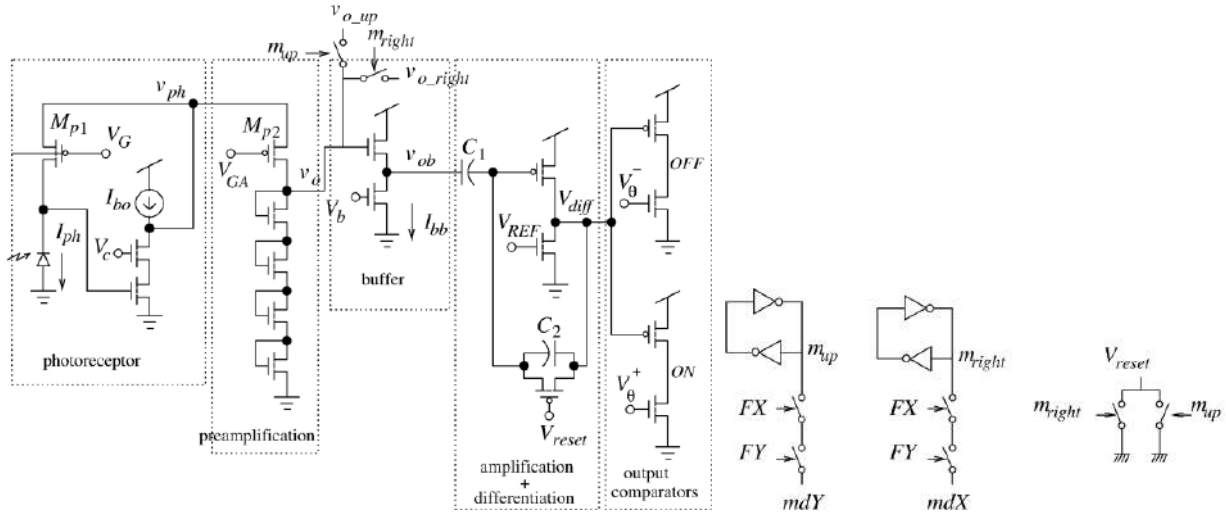


Figure 3. (a) Photodetector design for pixel interconnection (b) Pixel circuitry for foveation control

Voltage v_o is buffered and further amplified and differentiated, so that

$$\Delta V_{diff} = -\frac{C_1}{C_2}\Delta V_o \cong -\frac{C_1}{C_2} 4n_n U_T log\frac{\sum_{i=1}^{G} I_{phi_{final}}}{\sum_{i=1}^{G} I_{phi_{initial}}} = -A_T U_T log\frac{\sum_{i=1}^{G} I_{phi_{final}}}{\sum_{i=1}^{G} I_{phi_{initial}}} \quad (12)$$

The behavior is thus mathematically equivalent to joining the photodiodes of the neighboring pixels and adding their photocurrents.

Based on the above considerations, we have selected the third option of pixel grouping so that the output nodes of the pre-amplified photocurrent v_o are shared by neighboring pixels to build low resolution regions.

3 Digitally Reconfigurable Foveated Architectures

At the system level, different implementations and corresponding control circuitry for the pixel grouping mechanisms have been considered. We have considered different system architectures for the electronically foveated sensors. In the following, we explain three possible architectures for the EF-DVS.

3.1 Digitally Reconfigurable Foveated Architecture 1

3.1.1. Pixel Design

Figure 3(a) illustrates a particular pixel realization where the output nodes of the pre-amplified photocurrent v_o are shared by neighboring pixels. Digital control signals m_{up} and m_{right} are digitally stored and are individually reconfigurable for every pixel in the array as shown in Figure 3(b).

Whenever, digital signal m_{up} is active the pixel will be configured in low resolution mode and its v_o signal will be connected to the upper neighboring pixel. In a similar way, whenever, signal m_{right} is active, the pixel is in low resolution connected to its right neighboring pixel. Only when both signals m_{up} and m_{right} are simultaneously low, the pixel is operating individually in full HR mode. Using this scheme, low resolution regions with arbitrary shapes and sizes can be externally programmed in the sensor as will be explained next.

As illustrated in Figure 3(b), signal V_{reset} is shorted to ground so that the pixel is reset and unable to generate output events, whenever one of the signals m_{up} or m_{right} is active. That way, only the upper-right pixel of each pixel group is allowed to generate the output events corresponding to the averaged signal of the group. Figure 4 plots pixel signals during a simulation in which the pixel photocurrent is going through periodic variations (upper signal), and the pixel configuration

signals m_{up} and m_{right} are varied changing the pixel resolution configuration. As can be observed, only when m_{up} and m_{right} are simultaneously low, the pixel generates positive and negative output events. Consequently, only the upper-left pixel in each low-resolution group would generate output events representing the temporal changes of averaged photocurrents in the group.

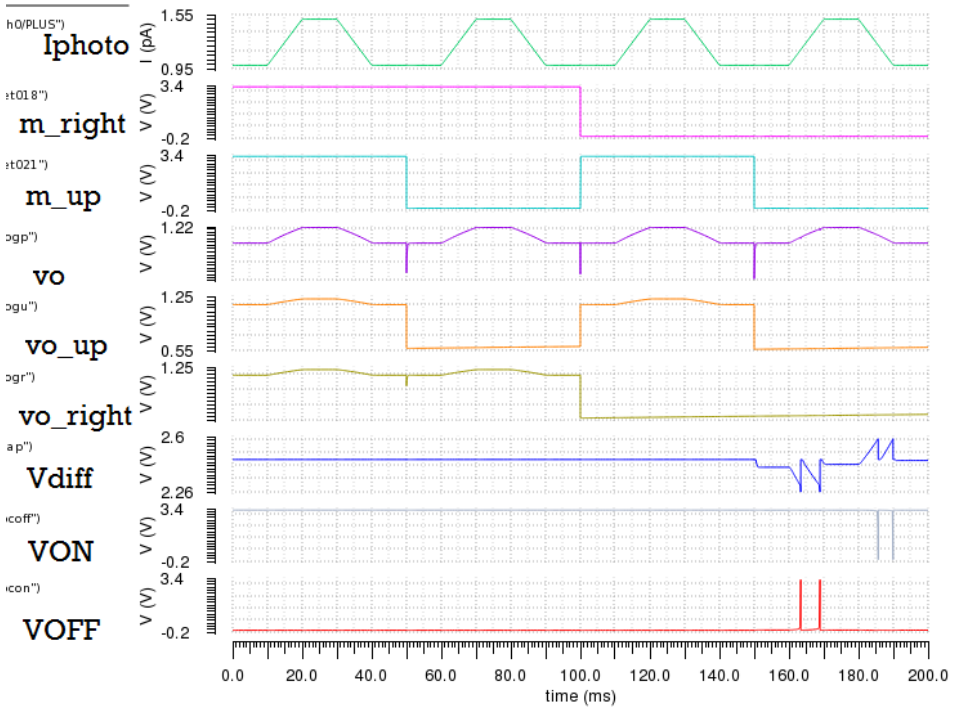


Figure 4. Transient simulation of pixel response under different foveation configuration signals

3.1.2. Array Design

The architecture of the DVS array for fully digital control of foveal reconfiguration is shown in Figure 5. The sensor contains an array of NxN pixels and the corresponding peripheral circuitry to generate the address-event-output. Additional X-FOV and Y-FOV foveation control blocks are added to the sensor for digital configuration of the foveal regions. Figure 5(b) shows the schematic of X-FOV foveation block.

The sensor receives as input the address [***fax, fay***] of the pixel that should be reconfigured and a two bit digital configuration mode signal [*mdX,mdY*]. The X-FOV and Y-FOV foveation control blocks contain decoders that activate the corresponding column and row signals [FX_i, FY_j]. As can be observed in Figure 3(b), the corresponding pixel where the [FX_i, FY_j] signals are simultaneously selected, updates it foveation configuration signals m_{up} and m_{right} to the global digital inputs [*mdX,mdY*]. Additional control signals *AEX*, *AEY* and *AD* are added to the X-FOV and Y-FOV foveation control blocks to allow simultaneous configuration of all the array, all pixels in a row, or all the pixels in a column. During single pixel configuration signals *AEX*, *AEY* and *AD* should be set to logic 1. Setting *AEX* to 0, all the column wise selection signals FX_i are simultaneously active, so that all pixels in the selected FY_j row are parallely configured with the same [*mdX,mdY*] input. In a similar way setting *AEY* to zero, all row wise selection signals FY_j are simultaneously active, so that all pixels in the selected FX_i column are parallely configured with the same [*mdX,mdY*] input. Setting *AEX* and *AEY* simultaneously to 0, allows parallel configuration of all the array pixels to the same [*mdX,mdY*] state. Setting signal *AD* to 0 deactivates all row and column selection [FX_i, FY_j] signals disabling foveal reconfiguration.

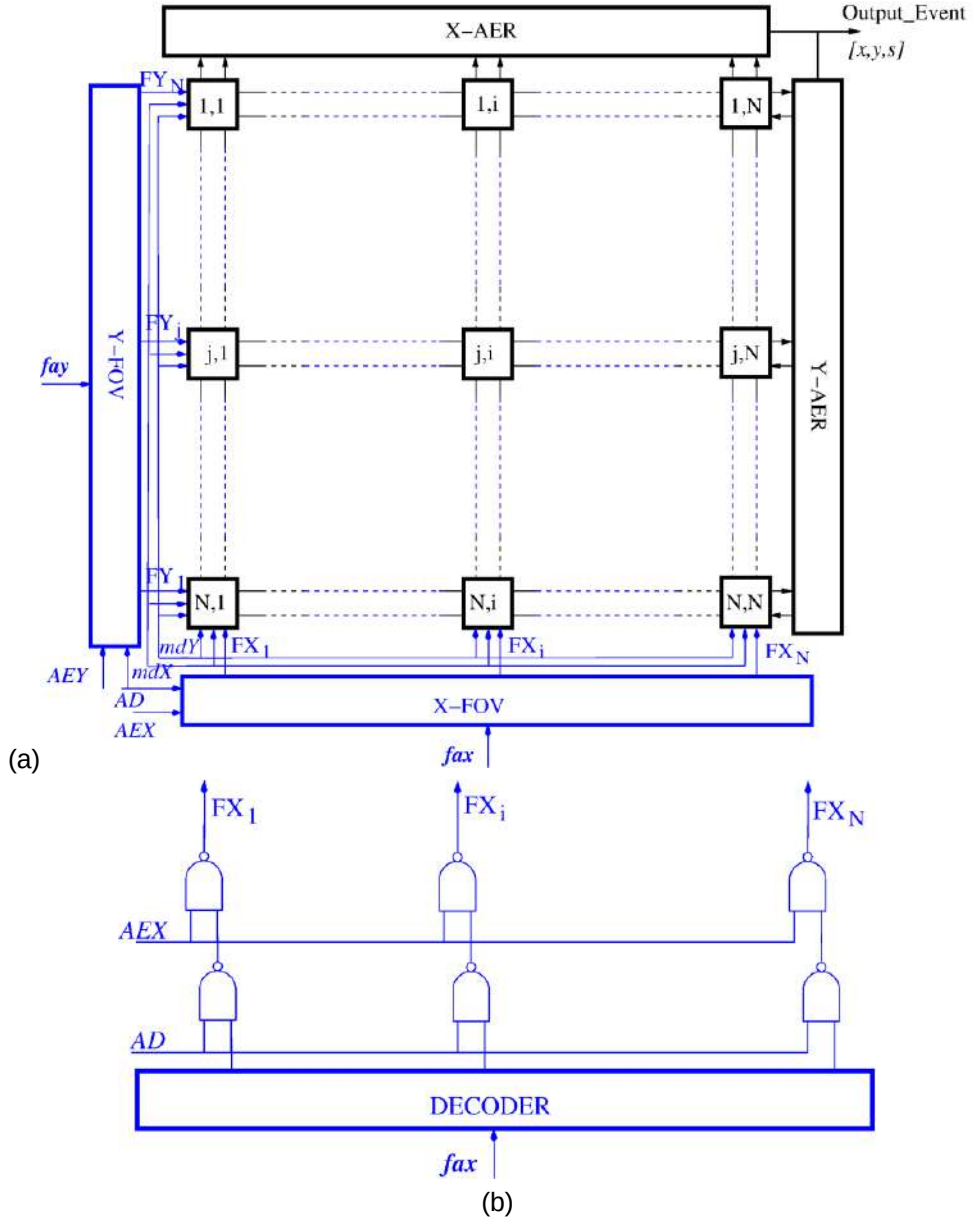


Figure 5. EF-DVS architecture 1. (a) Foveated DVS Array Architecture, (b) X-Foveation control block

3.2 Digitally Reconfigurable Foveated Architecture 2

An alternative architecture to define the foveation regions is depicted in Figure 6(a). The NxN pixels in the sensor are organized in groups of MxM pixels each. Each group can be configured using peripheral X-FOV and Y-FOV foveation control blocks in high-resolution (HR) mode or low resolution (LR) mode using just a single bit stored in the group as shown in Figure 6(b). When the stored control bit HR is high, each pixel operates individually. On the contrary, when HR is deactivated all the pixels in the group are interconnected. When HR is deactivated, a switch controlled by *not(HR)* is activated in every pixel group except for the central pixel that generates the group output events. Using this architecture, the number of control lines for configuring the foveal regions is reduced as well as the number of extra circuitries needed for storing the foveal configuration which is now shared by all the pixels in a group. However, the unfoveated regions have now a fixed size with no possible adaptation.

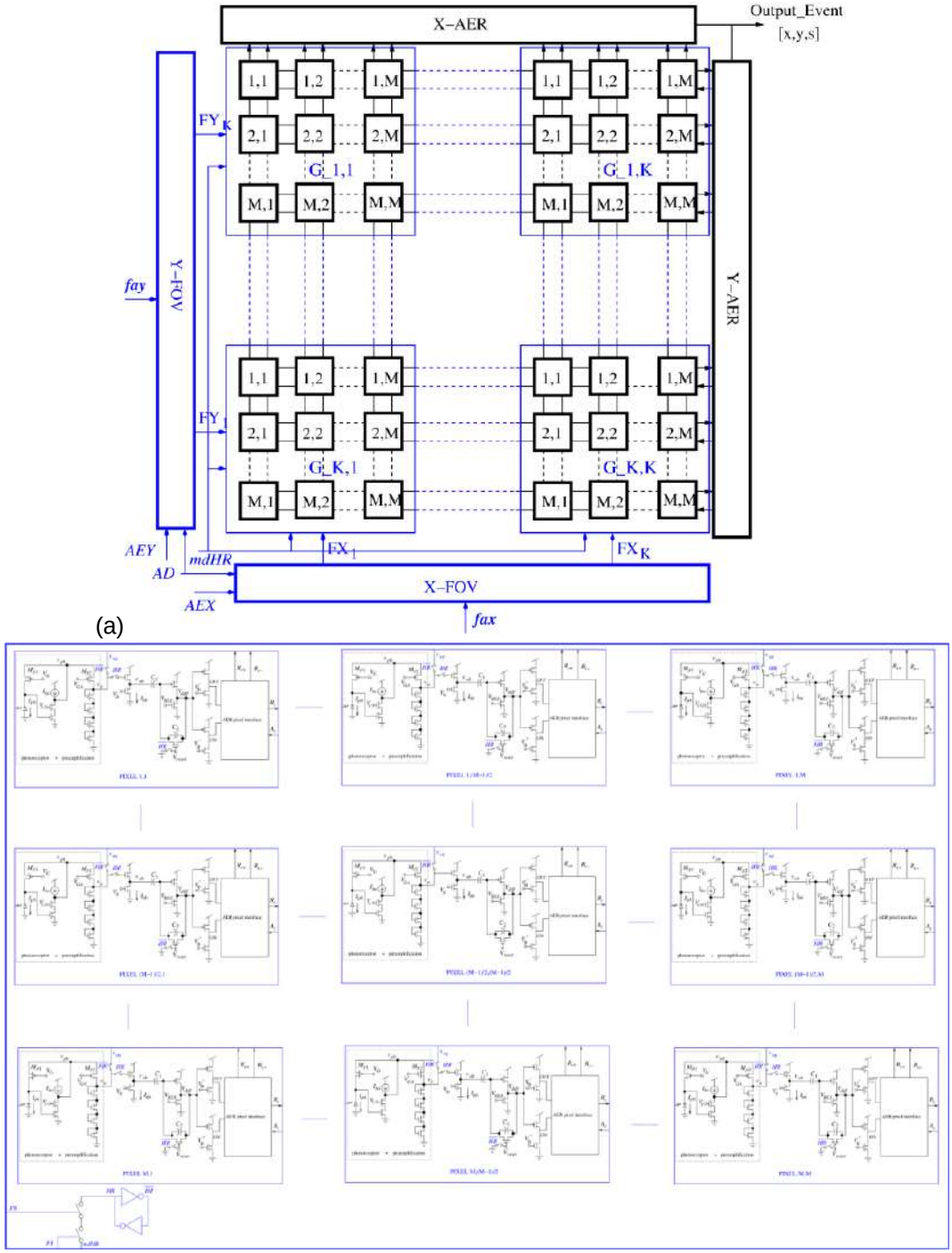
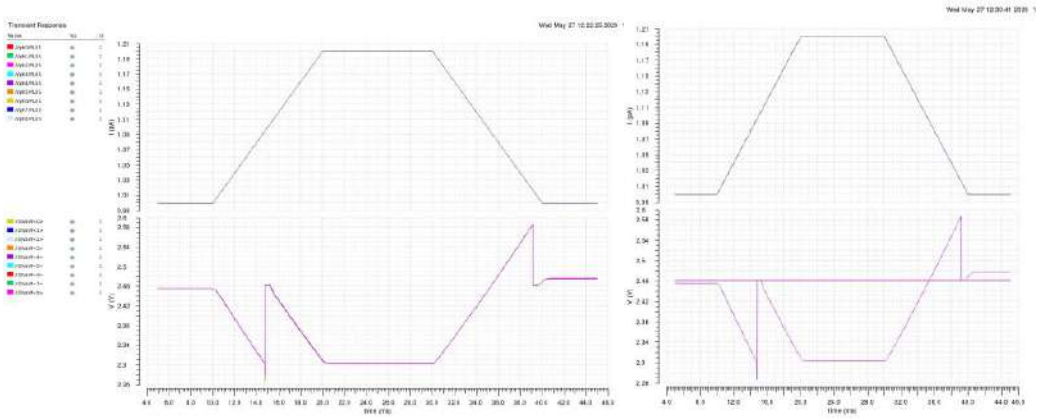


Figure 6. EF-DVS architecture 2. (a) Foveated DVS Array Architecture, (b) X-Foveation control block

A group of 3x3 pixels using the foveal configuration circuitry shown in Figure 6 has been simulated in the AMS0.35micron technology. Figure 7(a) corresponds to a simulation where the group pixels are configured in HR mode, while Figure 7(b) shows simulation results when the group is configured in LR mode. In the simulations, each pixel is stimulated with identical stimulus of a photo current that goes up and down following a ramp as shown in the upper subfigures. The lower subfigure of Figure 7(a) shows the voltages at node V_{diff} of all the pixels in the group in HR configuration, while in Figure 7(b) the group is in low resolution configuration. As can be observed, when operating in LR mode the central pixel (pixel 4

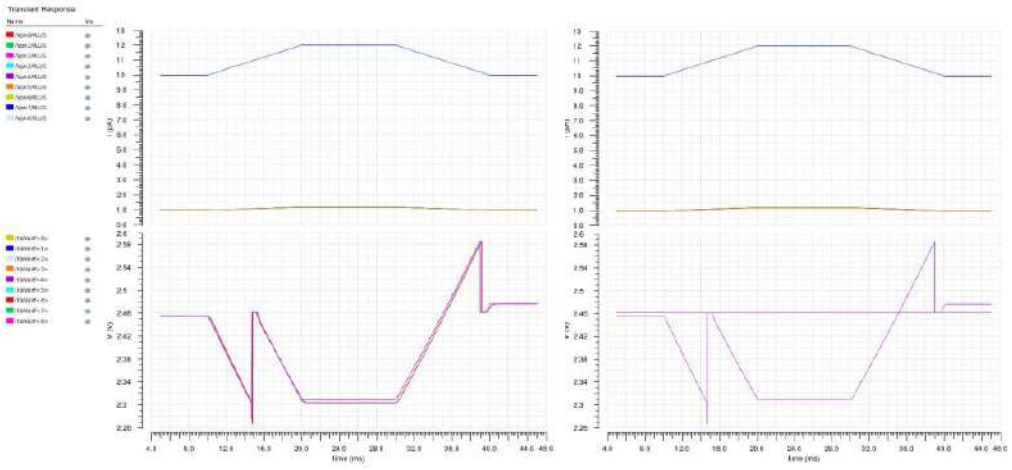
in the group) follows the same V_{diff} vs. time variation than the individual pixels in the LR configuration, while the rest of the pixels in the group remain silent.



(a) (b)

Figure 7. Transient simulation showing the response the voltage at node V_{diff} of a group of 3x3 pixels when the injected photo current goes through a up and down voltage ramp (a) The group is in a high-resolution configuration and (b) The group is in low resolution mode.

Figure 8 shows a simulation similar to the one of Figure 7, but in this case, 4 of the 9 group pixels are receiving a photocurrent 10 times higher than the rest of pixels. Figure 8(a) shows the voltages at node V_{diff} of all the pixels in the group in HR configuration, while in Figure 8 (b) the group is in low resolution configuration. As can be observed the response of voltages V_{diff} is independent of the absolute value of the photocurrent but depends on the relative variation of the stimulus.



(a) (b)

Figure 8. Simulation of the transient response of voltages at node V_{diff} of a group of 3x3 pixels when the injected photo current goes through a up and down voltage ramp (a) The group is in a high-resolution configuration and (b) The group is in low resolution mode.

3.3 Digitally Reconfigurable Foveated Architecture 3

A sensor foveated architecture that combines the higher flexibility to configure the foveation regions of the first architecture shown in Figure 5 with the reduction in the number of foveation control lines and sharing of the configuration memory among pixels in a group is proposed in Figure 9(a). In this proposed architecture, neighboring pixels group in a

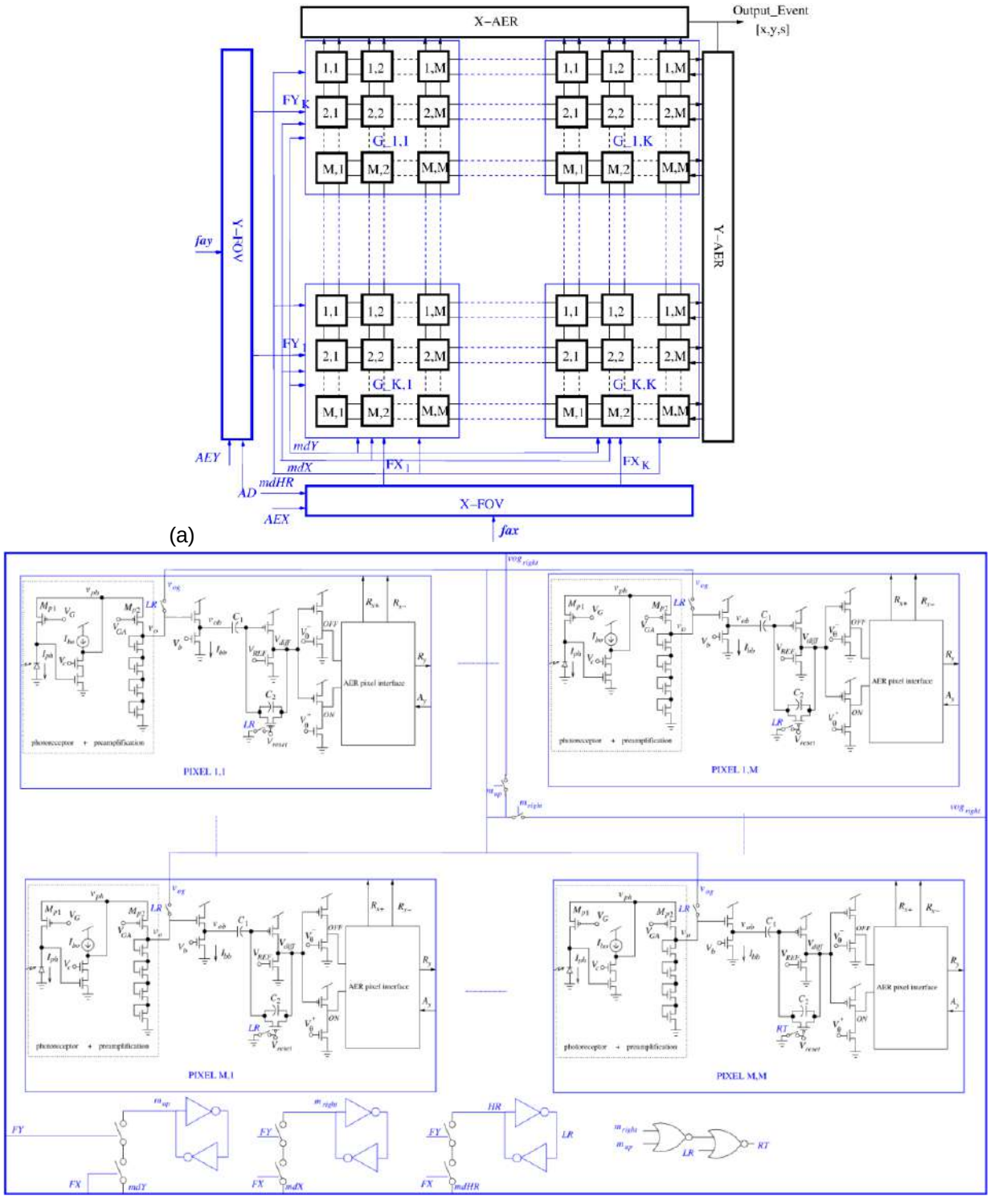
predefined form when configured in low resolution as proposed in the second foveated architecture. However, in this case the possibility of forming supra-groups with arbitrary shape is added. To connect neighboring groups, a similar scheme to the one considered in the first foveated architecture can be used as depicted in Figure 9(b). In this case, three bits define the resolution configuration of each pixel group.

Signal HR configures each group to operate individually (high resolution mode) or in a low-resolution group interconnected manner. In case HR is active only one pixel in the group emits output events.

Two additional bits m_{up} and m_{right} can define interconnectivity with upper and right group pixels whenever signal HR is low. In that case, only one pixel in the upper right corner of the supra-group would have its RT signal deactivated allowing it to generate output events. Table 1 contains the logic map implemented to control reset signal RT.

m_{up} m_{right} \ HR	0	1
0 0	0	0
0 1	1	0
1 1	1	0
1 0	1	0

Table 1. Logic map implemented by RT reset signal



(a)

(b)

Figure 9. EF-DVS architecture 3. (a) Foveated DVS Array Architecture, (b) X-Foveation control block

4 EF-DVS AMS0.35 microns prototype

An EF-DVS implementing the first proposed foveated architecture has been designed in AMS 0.35 OPTO technology. The prototype contains an array of 128x128 pixels with can be operated individually or can be combined in LR regions following the foveation scheme explained in section 3.1.

4.1 EF-DVS Schematic

Figure 10(a) shows the top-level schematic of the implemented sensor. It includes two sub-blocks: (Figure 10(b)) a biasing circuitry to generate on-chip programmable current biases and one main block (Figure 10(c)) that includes an array of 128x128 pixels, circuitry for global illumination adaptation and the arbitrating and decoding circuitry to generate the output address events.

Figure 11 plots the blocks of the 128x128 array. It contains an array of 128x128 DVS pixels and the foveal control blocks (as already shown in Figure 5(b)) including a 7-bit decoder. Each DVS pixel contains the photo sensor amplification and differentiation circuitry shown in Figure 3(a), the foveation control circuitry shown in Figure 3(b), and additional circuitry for driving the row and column event request and receiving acknowledge signals from peripheral arbitration and decoding circuitry. Signals *py<0:127>* and *sy_c<0:127>* (see Figure 11) are the event row request and acknowledge signals, respectively. *pulsep<0:127>* and *pulsem<0:127>* are column positive and negative event request signals. The *periph128* block (see Figure 10(c)) contains row and column arbitrating, selecting and decoding circuitry that encode the corresponding selected event in the ***x*** and ***y*** external buses and generate the Acknowledge and Request signals for external 4-phase handshaking protocol communication. Signal *x<7>* bit encodes the event sign.

Output I_{ave} of the array delivers a current equivalent to the total accumulated current of the 128x128 photodiodes in the array. A copy this current is sent to an external I_{ave} pad for testing purposes. Another copy of the current divided by a 128 factor is sent to peripheral circuitry where an 128 times replica of the photoreceptor and preamplification pixel circuitry can be used for characterization and automatic global adaptation of the pixel internal current gain A_I that can be adjusted depending on the global average illumination through the voltage difference $\Delta V_P = V_{GPH} - V_{GPA}$. Additionally, some decoding circuitry and two analogue buffers have been added to select and observe without disturbing a selection of analogue internal nodes in the circuit.

Figure 10(b) shows the schematic of the biasing circuitry. Programmable current *Ipot* circuits are used to generate on-chip the sensor biasing currents[19]. A selection circuitry controlled by signal *sel_Iref* allows commuting the input reference current for *Ipot* block between an on-chip generated current (generated by a specific current extractor block[20] *is_extcasc*) and an externally provided current through pad Iref. *Ipot* block and additional digital selection signals are configured using a shift register with digital input signals *in, clock, Copy*. *Ipot* circuits are externally measured and calibrated. For this purpose, each *Ipot* circuit output current can be selected and measured for each configuration word using an external output *Itest*. To be able to measure off-chip very low currents additional controllable amplifying current mirrors (MP5-MP8 and MN5-MN6) have been added that can be also selected using signal *sel_test_ampC*.

Figure 12 details the schematic of the *Ipot* block. It contains five *Ipot* circuits[19]; one per bias current. Each *Ipot* circuit receives a copy of the input reference current and generates a bias current which is controlled by 24 configuration bits. The configuration bits are fed by a concatenated shift register and stored in latches as detailed in Figure 12(b). Each Ipot (see Figure 12(c)) contains two concatenated current splitters[21]. The first current splitter divides the input reference current *Iref* in decades and is controlled by *exp<0:5>*, a 6 bit digital control word which selects the current range. In this exp<0:5> word, only one bit is simultaneously active. The second current splitter contains 16 branches with a fine grained current division factor. *Mant<0:16>* is a 16 bit control word with a combination of the 16 output branches for fine tuning the bias current value. Another *Ipot* controlling bit selects the current sign and a final test bit selects between the normal biasing output or the test output for externally characterizing the *Ipot* current. Table 2 contains the meaning of the 24 bit configuration word of each *Ipot* circuit.

As can be observed in Figure 10(b), the output of the Ipot block shift register is concatenated with an additional digital control block that stores 13 configuration bits. The total bit-length of the chip configuration shift register is 5X24+13= 133

bits. The role of the different configuration bits is detailed in Table 3. Signals *Bit<121:130>* are used to control the internal nodes that should be connected to the input of the test analogue buffers.

Table 4 contains the correspondence between the selection words in the decoder selecting the signals connected to the buffer inputs and the corresponding signal names. The first six signals correspond to internal signals of pixel<0,0> in the array. In those cases two names are used: the first name corresponds to the name used in the schematic of Figure 2, and the second name is the corresponding name in the upper level shown in Figure 10(c). Signal *Test_buffers* is connected to an external analogue pad intended for testing the buffers.

Register order	Bit<0:5>	Bit<6:21>	Bit<22>	Bit<23>
Ipot control	Exp<0:5>	Mant<0:16>	sign	test
	1-hot code	Fine-tuned code	0–N-type current 1-P-type current	0- bias current 1- test current

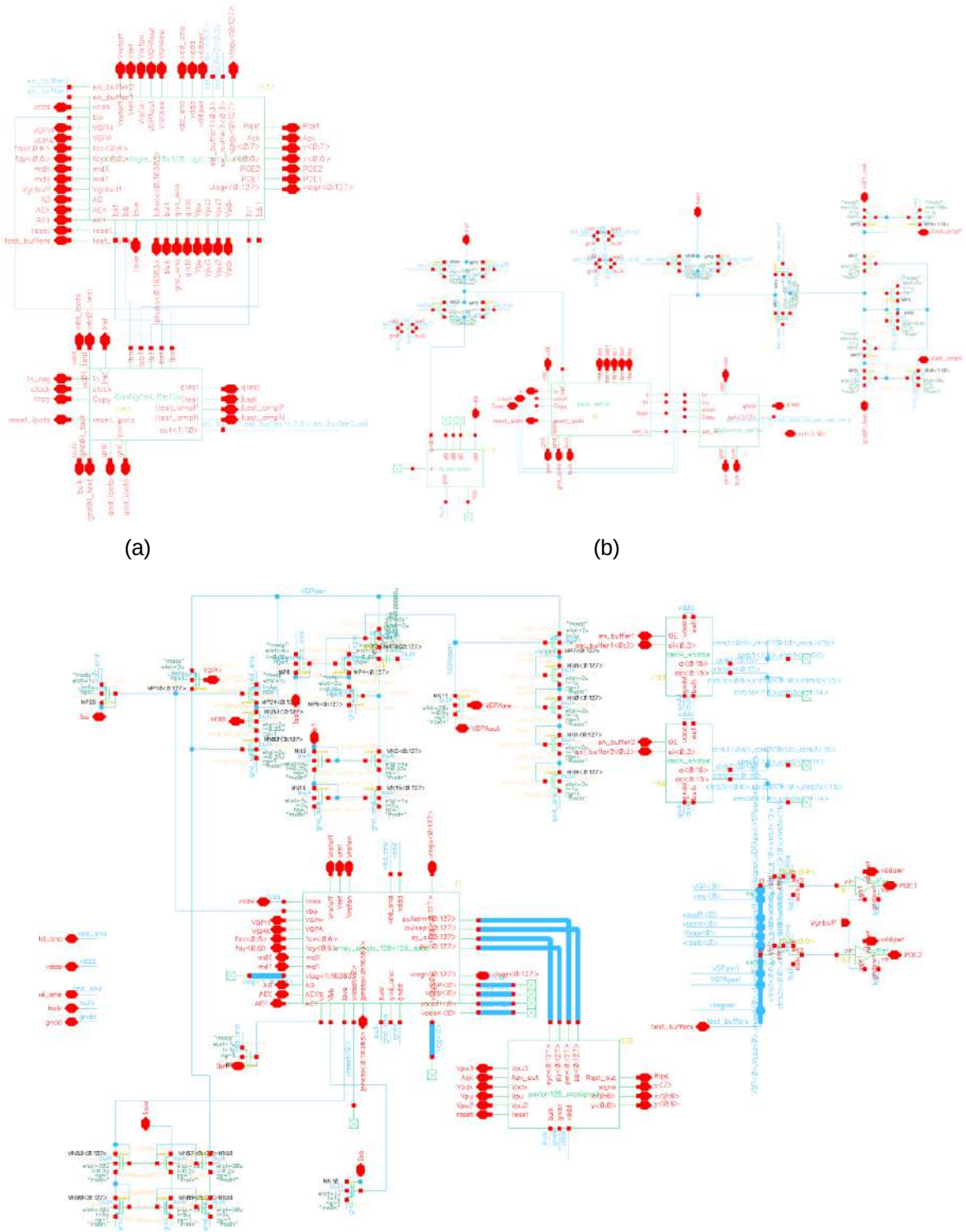
Table 2. Single Ipot configuration

Register Bit	Signal Name	Explanation
Bit<0:23>	Ibb	Bias Current of Pixel Buffer Stage
Bit<24:47>	Iab1	Bias Current for Peripheral Adaptative Control
Bit<48:71>	Ib1	Bias Current for Adaptative Control of Pre-amplifiers current gain
Bit<72:95>	Ibrf	Bias Current for Refractory Period Control
Bit<96:119>	Ibo	Bias Current of Photoreceptor source-driven active mirror
Bit<120>	Sel_Iref	‘1’- Iref external; ‘0’ –on-chip current reference
Bit<121>	En_buffer1	‘1’- Decoder buffer1; ‘0’- Disable decoder buffer1
Bit<122>	Sel_buffer1<3>	Selection bit decoder buffer 1
Bit<123>	Sel_buffer1<2>	Selection bit decoder buffer 1
Bit<124>	Sel_buffer1<1>	Selection bit decoder buffer 1
Bit<125>	Sel_buffer1<0>	Selection bit decoder buffer 1
Bit<126>	En_buffer2	‘1’- Decoder buffer2; ‘0’- Disable decoder buffer2
Bit<127>	Sel_buffer2<3>	Selection bit decoder buffer 2
Bit<128>	Sel_buffer2<2>	Selection bit decoder buffer 2
Bit<129>	Sel_buffer2<1>	Selection bit decoder buffer 2
Bit<130>	Sel_buffer2<0>	Selection bit decoder buffer 2
Bit<131>	Sel_test_out	‘1’- Itest out connected to internal node, ‘0’-Itest out disconnected
Bit<132>	Sel_test_amp	‘1’-Ipots out connected to Itest_ampN/P, ‘0’- disconnected Itest_ampN/P

Table 3. Configuration role of shift-register bits

Sel_buffer<3>	Sel_buffer<2>	Sel_buffer<1>	Sel_buffer<0>	Signal connected to buffer input
0	0	0	0	v_{ph}, *VSP<0>*
0	0	0	1	V_{diff}, *vcap<0>*
0	0	1	0	Vreset<0>
0	0	1	1	v_o, *vlog<0>*
0	1	0	0	*ON, vocon<0>*
0	1	0	1	*OFF, vocoff<0>*
0	1	1	0	vlogperi
0	1	1	1	None
1	0	0	0	None
1	0	0	1	VGPAperi
1	0	1	0	VSPperi
1	0	1	1	None
1	1	0	0	None
1	1	0	1	None
1	1	1	0	None
1	1	1	1	*Test_buffers*

Table 4. Configuration bits for controlling signal external visualization through analogue buffers



(a) (b)

(c)

Figure 10. (a): two main blocks in the sensor schematic: (b) biasing block, and (c) 128x128 sensor array with address event readout blocks.

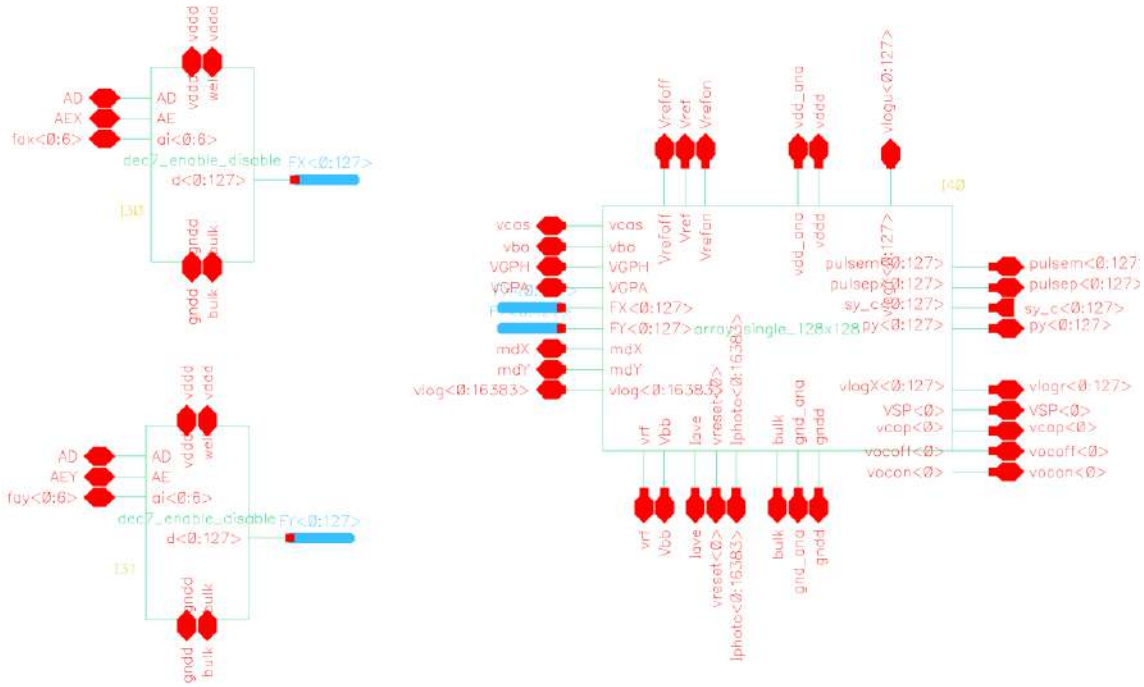
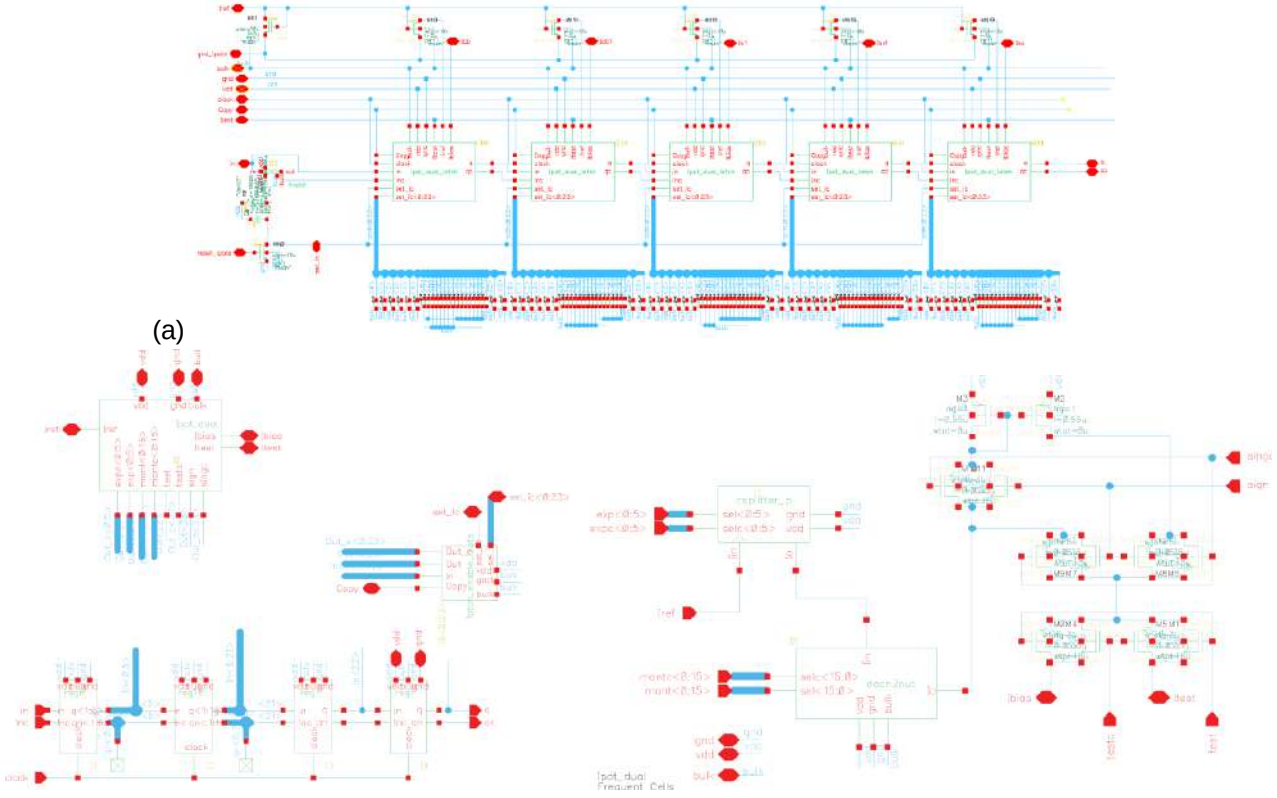


Figure 11. Schematic of 128x128 pixel array and peripheral foveation control blocks



(a)

(b)

(c)

Figure 12 (a) Schematic of Ipots chain for current biases control, (b) schematic of an Ipot including the digital shift register and the programmable current splitting ladders and (c) schematic of the programmable current splitting ladders

4.2 EF-DVS Chip Layout

Figure 13 plots the full layout of the EF-DVS sensor in the AMS 0.35 OPTO technology including the pads. The total chip area is 5,02x5,05 mm^2 including the pads. In Figure 14, the area dedicated to the main circuit blocks is indicated. The sensor is fully covered with a shielding of metal 4 except for the photodetector areas. The chip has been packaged in JLCC84 package. The chip pinout is detailed in Table 5.

Figure 14 details a plot of the layout of a 2x2 assembly of two foveated pixels. The area of an assembly of 2x2 pixels is 62x62,2□m^2. In a 2x2 pixel assembly, the pixels are disposed in a mirroring position in the top-down and left-right directions, so that sensitive analogue parts and analogue biases areas are shared by neighbouring pixels and are isolated as much as possible from digital noisy parts.

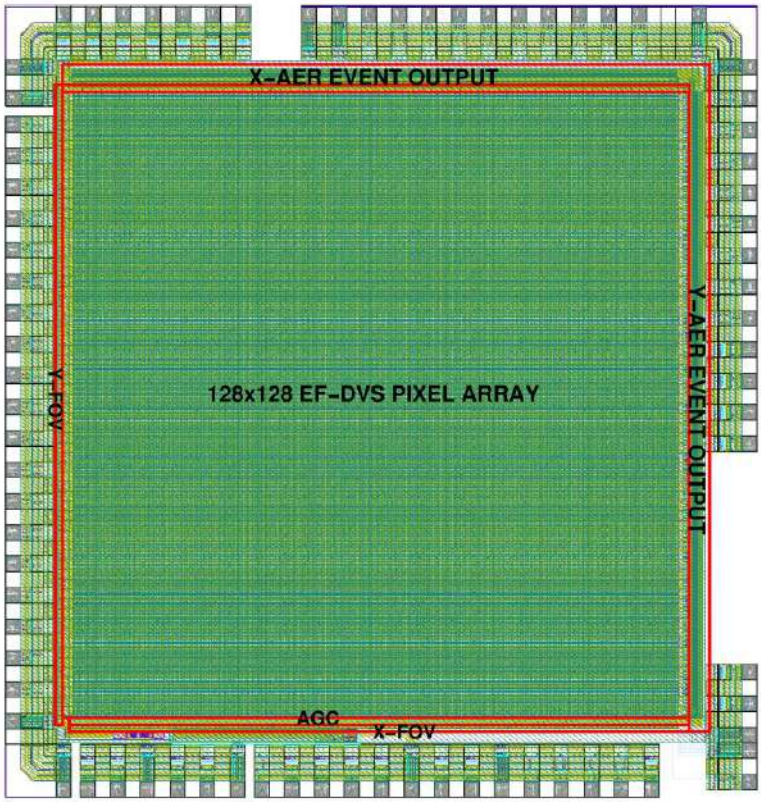


Figure 13. EF-DVS sensor layout in AMS-0.35 OPTO Technology packaged in a JLCC84

Pin No	Schematic Name	Type	Description
1	fax<6>	I – digital input	Bit of x-word input address for foveal configuration
2	fax<5>	I – digital input	Bit of x-word input address for foveal configuration
3	fax<4>	I – digital input	Bit of x-word input address for foveal configuration
4	fax<3>	I – digital input	Bit of x-word input address for foveal configuration
5	fax<2>	I – digital input	Bit of x-word input address for foveal configuration
6	fax<1>	I – digital input	Bit of x-word input address for foveal configuration
7	fax<0>	I – digital input	Bit of x-word input address for foveal configuration
8	AEX	I – digital input	Enable all column selection in foveal config when high
9	mdX	I – digital input	Signal configuring horizontal foveal grouping

10	mdY	I – digital input	Signal configuring vertical foveal grouping
11	gndd	power	Ground connection
12	vddd	power	Power Supply for digital blocks
13	test_buffers	I – analog in	Analog input to test the analog output buffers
14	POE1	O –analog out	Ouput of the first analog test buffer
15	vddpwr	power	Power Supply for analog output buffers
16	Vgnbuff	I – analog in	Voltage bias of analog test buffers. Nom value=0.5V
17	POE2	O –analog out	Ouput of the second analog test buffer
18	bulk	power	Ground connection
19	qtest	O – digital out	Output of configuration shift register
20	Itest_ampN	O- analog out	Output current of N type current mirror-test of Ipots
21	Itest_ampP	O- analog out	Output current of P type current mirror-test of Ipots
22	gnd01_test	I – analog in	Analog voltage to set gain in test of low Ipot currents
23	vdd01_test	I – analog in	Analog voltage to set gain in test of low Ipot currents
24	gnd_ipots	power	Ground connection
25	Itest	O- analog out	Output current test of Ipots – both signs
26	Iref	I – analog in	Input current – usable as reference biasing current
27	vdd_ipots	power	Power Supply for ipots
28	in_reg	I – digital input	Input of the configuration shift register
29	clock	I – digital input	Clock of the configuration shift register
30	copy	I – digital input	Latch of the configuration shift register
31	reset_ipots	I – digital input	Reset the configuration shift register to default values
32	VGPAout	I/O-analog out	Output voltage of global adaptation/Voltage test input
33	NC	NC	Non Connected
34	NC	NC	Non Connected
35	NC	NC	Non Connected
36	VGPAsw	I-digital in	Control of switch connecting VGPAout
37	gnd_ana	power	Ground connection
38	vdd_ana	power	Analog power supply
39	Vrefoff	I – analog in	Voltage off threshold of pixel
40	Iave	0 –analog out	Current output of N-type measuring illumination level
41	bulk	power	Ground connection
42	Vpu3	I – analog in	Distributed-gate pull-up voltage. Optional.Default=gnd
43	Vpdx	I – analog in	Distributed-gate pull-down voltage. Opt.Default=vdd
44	Vpu	I – analog in	Distributed gate pull-up voltage. Optional.Default=gnd
45	Ack	I-digital in	Acknowledge for output address event
46	y<6>	O-digital out	Bit of Y-word of output address event
47	y<5>	O-digital out	Bit of Y-word of output address event
48	y<4>	O-digital out	Bit of Y-word of output address event
49	y<3>	O-digital out	Bit of Y-word of output address event
50	y<2>	O-digital out	Bit of Y-word of output address event
51	y<1>	O-digital out	Bit of Y-word of output address event
52	y<0>	O-digital out	Bit of Y-word of output address event
53	vddd	power	Power Supply for digital blocks
54	gndd	Power	Ground connection
55	Rqst	0-digital out	Request for output address event

56	x<7>	O-digital out	Bit of X-word of output address event
57	x<6>	O-digital out	Bit of X-word of output address event
58	x<5>	O-digital out	Bit of X-word of output address event
59	x<4>	O-digital out	Bit of X-word of output address event
60	x<3>	O-digital out	Bit of X-word of output address event
61	x<2>	O-digital out	Bit of X-word of output address event
62	x<1>	O-digital out	Bit of X-word of output address event
63	x<0>	O-digital out	Bit of X-word of output address event
64	reset	I-digital input	Reset of address event peripheral latches
65	gndd	Power	Ground connection
66	vddd	power	Power Supply for digital blocks
67	Vpu2	I – analog in	Distributed-gate pull-up voltage. Optional.Default=gnd
68	Vrefon	I – analog in	Voltage on threshold of pixel
69	Vref	I – analog in	Voltage reference threshold of pixel
70	VGPH	I – analog in	Voltage bias. Nominal value = 2.7V
71	VGPA	I – analog in	Voltage bias. Nominal value = 2.7V
72	vcas	I – analog in	Voltage bias. Nominal value = 1.65V
73	bulk	power	Ground connection
74	vdd_ana	power	Analog power supply
75	gnd_ana	power	Ground connection
76	fay<0>	I – digital input	Bit of y-word input address for foveal configuration
77	fay<1>	I – digital input	Bit of y-word input address for foveal configuration
78	fay<2>	I – digital input	Bit of y-word input address for foveal configuration
79	fay<3>	I – digital input	Bit of y-word input address for foveal configuration
80	fay<4>	I – digital input	Bit of y-word input address for foveal configuration
81	fay<5>	I – digital input	Bit of y-word input address for foveal configuration
82	fay<6>	I – digital input	Bit of y-word input address for foveal configuration
83	AEY	I – digital input	Enable all row selection for foveal config when high
84	AD	I – digital input	Disable signal of vert&horiz decoders for foveal config

Table 5. Pinout description of EF-DVS JLCC84 package

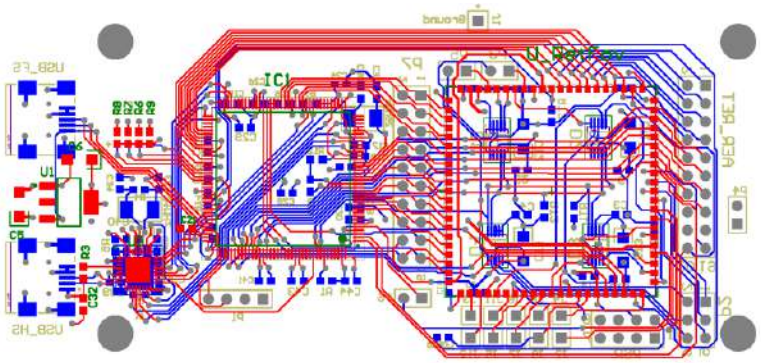


Figure 15. PCB designed for the EF-DVS sensor

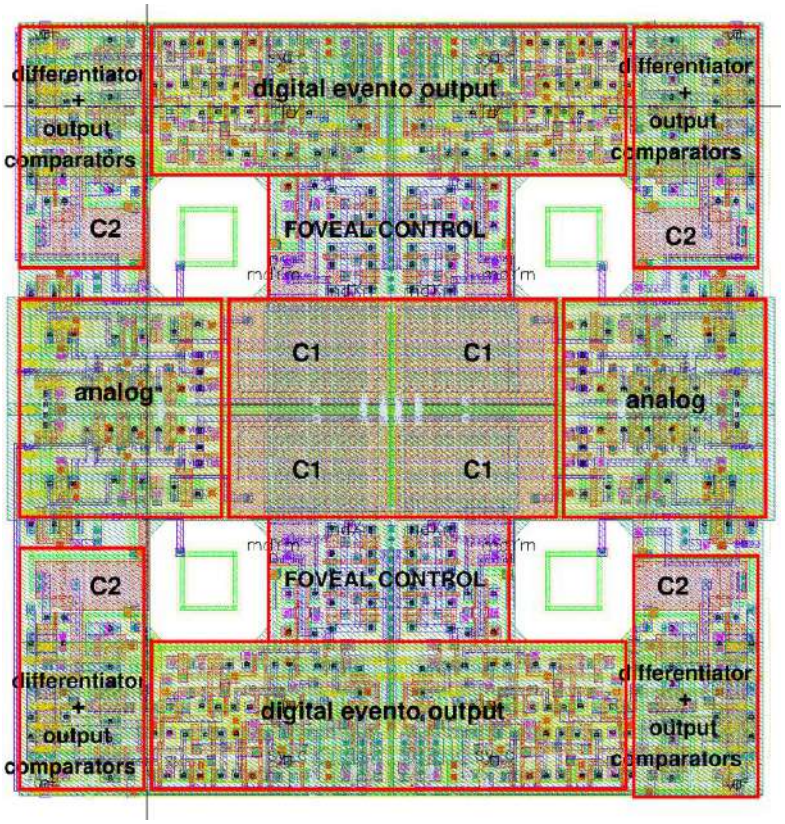


Figure 14. Layout of an array of 2x2 foveated EF-DVS pixels in AMS-0.35 OPTO Technology

4.3 Test PCB Design

A PCB has been designed to test the EF-DVS sensor as shown in Figure 15. The PCB contains a STM microcontroller, some DACs, the EF-DVS and connectors for USB communication, a parallel AER output connector, as well as connectors for signal output and signal testing.

A schematic representation of the full board is shown in Figure 16, where "C" means to connect to an output connector pin and "M" means to connect to the microcontroller input pin.

Figure 17 contains a schematic representation of the main connectors, where "R" means the connector pins are connected to output pins of the Retina and "M" means the connector pins are connected to the output pins of the microcontroller.

The microcontroller serves to communicate to a host computer through a USB connector and:

1. Bias voltages and currents which are set through the STMicrocontroller. There are four DACs on the PCB which are set through I2C protocol and also the initial value of the programmable biasing currents of the retina can be changed through the microcontroller, if needed.
2. Digital outputs of the microcontroller connect to digital foveation control inputs in the EF-DVS. The operating mode of the retina (low or high resolution), as well as the region of interest, can be determined by the fay<0:6>, fax<0:6>, AEY, AD, AEX, mdX, and mdY digital signals which are generated by the microcontroller. These signals are also connected to a separate connector for having external access (shown in Figure 16).
3. AER<0:15> is the digital output word of the retina which represents the address event. This address is connected to the separate connector to have direct and external access; it is also connected to inputs of the microcontroller to have the possibility of using the address for some real-time decisions.

fay<0:6>	I – digital input	Bits of input address for foveal configuration
AEY	I – digital input	Enable all row selection for foveal config when high
AD	I – digital input	Disable signal of vert&horiz decoders for foveal config
fax<0:6>	I – digital input	Bits of address for foveal configuration
AEX	I – digital input	Enable all column selection in foveal config when high
mdX	I – digital input	Signal configuring horizontal foveal grouping
mdY	I – digital input	Signal configuring vertical foveal grouping
AER<0:15>	O-digital out	Bit of Y-word of output address event

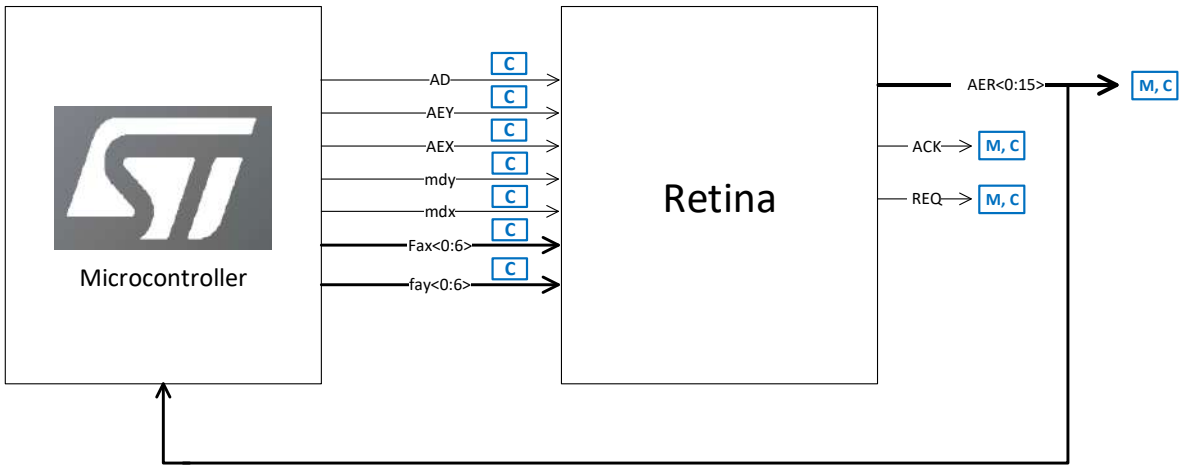


Figure 16 Board-Level diagram of the Retina and microcontroller connection

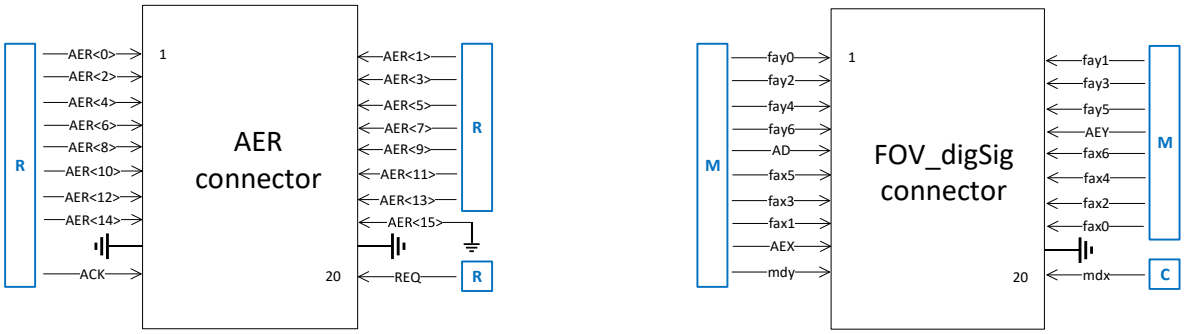


Figure 17 Board-Level diagram of the main connectors. DVS event output connector (Left). Connector for digital input signals to control the region of interest(Right).

5 Future Work

The design of the 128x128 EF-DVS sensor was sent out for fabrication in the AMS 0.35micron technology in April 2021. The packaged dies were received in November 2021. In parallel, with the sensor fabrication, the test PCB was designed. After checking the chip correct pinout and package, and refining the PCB design, the test PCB fabrication order was placed in November 2021. The PCB with the mounted components was received on February 22nd 2022. We are now ready for testing the EF-DVS prototype.

After April 2021, a new EF-DVS sensor in an advanced XFAB 180nm technology has been designed. The new foveated pixel occupies a reduced area. A 2x2 array of dvs pixels in the xfab 0.18 technology occupies 35.1 um x35.85um. A preliminar version of 32x32 a EF-DVS is being designed to be sent out for fabrication in May 2022. Once the correct functionality in the new technology has been checked a higher resolution 256x256 EF-DVS sensor will be fabricated in XFAB 180nm to be used in the project final demonstrators.

6 Publications

T. Serrano-Gotarredona and B. Linares-Barranco, *patent - Electronically Foveated Dynamic Vision Sensor*, ES1641.1671, October 5th 2021, CSIC (90%) and Universidad de Sevilla (10%).

7 References

1. Finateu, T. *et al.* A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86µm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline. in *Digest of Technical Papers - IEEE International Solid-State Circuits Conference* vols 2020-Febru 112–114 (Institute of Electrical and Electronics Engineers Inc., 2020).
2. Chen, S. & Guo, M. Live demonstration: CeleX-V: A 1M pixel multi-mode event-based sensor. in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* vols 2019-June 1682–1683 (IEEE Computer Society, 2019).
3. Brandli, C., Berner, R. & Delbruck, T. A 240 × 180 130 dB 3 µs Latency Global Shutter Spatiotemporal Vision Sensor. *IEEE J. Solid-State Circuits* **49**, 2333–2341 (2014).
4. Li, C., Longinotti, L., Corradi, F. & Delbruck, T. A 132 by 104 10µm-Pixel 250µW 1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression. *2019 Symp. VLSI Circuits* **2**, C216–C217 (2019).
5. Son, B. *et al.* A 640×480 dynamic vision sensor with a 9µm pixel and 300Meps address-event representation. *Dig. Tech. Pap. - IEEE Int. Solid-State Circuits Conf.* **60**, 66–67 (2017).
6. Bhattacharya, B. Sen & Furber, S. B. Biologically inspired means for rank-order encoding images: A quantitative analysis. *IEEE Trans. Neural Networks* **21**, 1087–1099 (2010).

7. Schwartz, E. L. Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception. *Biol. Cybern.* **25**, 181–194 (1977).
8. Shibata, T., Vijayakumar, S., Conradt, J. & Schaal, S. Biomimetic oculomotor control. *Adapt. Behav.* **9**, 189–207 (2001).
9. Ude, A., Gaskett, C. & Cheng, G. Foveated vision systems with two cameras per eye. *Proc. - IEEE Int. Conf. Robot. Autom.* **2006**, 3457–3462 (2006).
10. Thunuguntla, S. S. Object tracking using log-polar transformation. *PhD Thesis, Univ Lousiana* (2005).
11. Dauc, E., Albiges, P., Perrinet, L. & Neurosciences, I. De. A dual foveal-peripheral visual processing model implements efficient saccade selection. *J. Vis.* 1–21 (2020) doi:10.1101/725879.
12. Azadmehr, M., Abrahamsen, J. P. & Hafliger, P. A Foveated AER Imager Chip. *Int Symp. Circuits Syst.* 2751–2754 (2005).
13. Xia, S., Sridhar, R., Scott, P. & Bandera, C. All CMOS foveal image sensor chip. *Proc. Annu. IEEE Int. ASIC Conf. Exhib.* 409–413 (1998) doi:10.1109/asic.1998.723085.
14. Wodnicki, R., Roberts, G. W. & Levine, M. D. A log-polar image sensor fabricated in a standard 1.2-µm ASIC CMOS process. *IEEE J. Solid-State Circuits* **32**, 1274–1277 (1997).
15. Etienne-Cummings, R., Van Spiegel, J. Der, Mueller, P. & Zhang, M. Z. A foveated silicon retina for two-dimensional tracking. *IEEE Trans. Circuits Syst. II Analog Digit. Signal Process.* **47**, 504–517 (2000).
16. Constandinou, T. G., Degenaar, P. & Toumazou, C. An adaptable foveating vision chip. *Proc. - IEEE Int. Symp. Circuits Syst.* 3566–3569 (2006).
17. Serrano-Gotarredona, T. & Linares-Barranco, B. A 128× 128 1.5% contrast sensitivity 0.9% FPN 3 µs latency 4 mW asynchronous frame-free dynamic vision sensor using transimpedance preamplifiers. *IEEE J. Solid-State Circuits* **48**, 827–838 (2013).
18. Lichtsteiner, P., Posch, C. & Delbruck, T. A 128x128 120 dB 15 us latency asynchronous temporal contrast vision sensor. *IEEE J. Solid-State Circuits* **43**, 566–576 (2008).
19. Serrano-Gotarredona, R., Camuñas-Mesa, L., Serrano-Gotarredona, T., Leñero-Bardallo, J. A. & Linares-Barranco, B. The stochastic i-pot: A circuit block for programming bias currents. *IEEE Trans. Circuits Syst. II Express Briefs* **54**, 760–764 (2007).
20. Heim, P., Schultz, S. & Jabri, M. A. Technology-independent biasing technique for CMOS analogue micropower implementations of neural networks. in *Proceedings of the 4-th International Worshoop on Cellular Neural Networks and their Applications* (1995).
21. Enz, C. C. & Vittoz, E. A. CMOS low-power analog circuit design. *Proc. - IEEE Int. Symp. Circuits Syst.* 79–133 (1996) doi:10.1109/etlpds.1996.508872.

8 Documentation

8.1 Applicable and Referenced Documents

#	Id	Description	Identifier (Ed Rev)	Date
AD1	FPP	Full Project Proposal	1.0	15.01.2019

8.2 Glossary and Terminology

Acronym	Definition

DVS	Dynamic Vision Sensor
EF-DVS	Electronically Foveated Dynamic Vision Sensor
HR	High Resolution
LR	Low Resolution
ROI	Region of Interest
DAC	Digital Analogue Converter
USB	Universal Serial Bus
PCB	Printed Circuit Board
I2C	Inter Integrated Circuit