A 128x128 1.5% Contrast Sensitivity 0.9% FPN 3μs Latency 4mW Asynchronous Frame-Free Dynamic Vision Sensor Using Transimpedance Preamplifiers

T. Serrano-Gotarredona and B. Linares-Barranco
Instituto de Microelectrónica de Sevilla (IMSE-CNM-CSIC), Av. Américo Vespucio s/n, 41092 Sevilla, Spain.

***Abstract*— Dynamic Vision Sensors (DVS) have recently appeared as a new paradigm for vision sensing and processing. They feature unique characteristics such as contrast coding under wide illumination variation, micro-second latency response to fast stimuli, and low output data rates (which greatly improves the efficiency of post-processing stages). They can track extremely fast objects (e.g., time resolution is better than 100kFrames/s video) without special lighting conditions. Their availability has triggered a new range of vision applications in the fields of surveillance, motion analyses, robotics, and microscopic dynamic observations. One key DVS feature is contrast sensitivity, which has so far been reported to be in the 10-15% range. In this paper, a novel pixel photo sensing and transimpedance pre-amplification stage makes it possible to improve by one order of magnitude contrast sensitivity (down to 1.5%) and power (down to 4mW), reduce the best reported FPN (Fixed Pattern Noise) by a factor of 2 (down to 0.9%), while maintaining the shortest reported latency (3μs) and good Dynamic Range (120dB), and further reducing overall area (down to 30x31μm^2 per pixel). The only penalty is the limitation of intrascene Dynamic Range to 3 decades. A 128x128 DVS test prototype has been fabricated in standard 0.35μm CMOS and extensive experimental characterization results are provided.**

Key Words- dynamic vision sensor, temporal contrast retina, vision sensor, event-based vision, address-event-representation, high-speed vision, contrast sensitivity, frame-free vision sensor.

I. INTRODUCTION

Address-Event-Representation (AER) is an Event-Driven bio-inspired technology for sensing and processing sensory information [1]-[3]. In Event-Driven sensors and systems information is coded and transmitted as pulses (spikes or events), similarly to what happens in biological sensory and processing systems. In an AER sensor, each time a pixel senses relevant information it asynchronously sends an event out, which can be processed event by event by event-based processors [4]-[8]. That way, relevant features pass through all the processing levels almost instantaneously, the only delay being caused by the propagation of pulses along the processing chain. As an additional advantage, only pixels that have relevant information send events out, thus saving power and bandwidth when no relevant information exists. Because of their high speed potential and energy efficient nature, AER sensors have become very popular in recent years. A wide variety of AER vision sensors have recently appeared in literature, including simple luminance to frequency transformation sensors [9]-[10], time-to-first spike coding sensors [11]-[14], foveated sensors [15]-[16], temporal contrast vision sensors [10],[17]-[26], motion sensing and computation systems [27]-[29], and spatial contrast sensors [10], [21]-[22], [30]-[32], to mention just a few.

Of special interest for very high speed processing applications are the so-called “Dynamic Vision Sensors” (DVS), where each pixel autonomously computes the normalized time derivative of the sensed light ($\dot{I}/I$) and provides an output event with its (x, y) coordinate when this amount exceeds a preset contrast [17]-[26]. DVS cameras have unique features such as contrast coding under very wide illumination variation, micro second latency response to fast stimuli, and low output data rate, which greatly improve the efficiency of post-processing stages. They can track extremely fast objects without special lighting conditions, providing timing resolutions better than 100kFrames/s. The availability of DVS cameras is triggering a new range of vision applications in the fields of surveillance, motion analyses, robotics [33], and microscopic dynamic observation [34]. Since they provide frame-free event-driven asynchronous low data rate visual information, very fast and efficient novel vision processing techniques are being developed for performing tasks such as stereo [35]-[37] or optic flow [38] computations, instant feature extraction [6]-[7] and texture or object recognition [39]-[41].

In this paper we present an improved low-power low-latency low-mismatch high-contrast-sensitivity *Dynamic Vision Sensor* (DVS) or temporal contrast sensor. Several prior frame-based temporal difference detector imagers have been published [42]-[47], but these suffer from limited speed response because they are based on the integration of photo currents during consecutive frames and the subsequent computation of the difference between them. Several event-driven (frame-free) temporal contrast vision sensors (DVS) have also been reported in recent years [17]-[23]. The sensor published by Kramer [20] had low contrast sensitivity, while the one described by Zaghloul [21]-[22] suffered from poor FPN (fixed pattern noise). Lichtsteiner et al. [17]-[18] presented the first practical DVS by introducing a self-clocked switched capacitor differencing and amplification stage resulting in low FPN (2.5%), practical contrast sensitivity (15%), reasonable pixel array size (128x128), very good latency (15μs), excellent intrascene dynamic range (120dB), and sufficient maximum event rate throughput (1*Meps*). The sensor was appropriate for high speed vision, since it could follow rotating objects up to a speed of 200 rps (revolutions per second). Posch reported an improved DVS based on the same principles but designed in a more advanced 180nm CMOS technology, with slightly better contrast sensitivity (9%) and FPN (0.9%) and with additional large dynamic range gray scale imaging capability [23]-[25].

Increasing the contrast sensitivity of a DVS is an important challenge, as it improves the quality and richness of the sensed scene. For example, Fig. 1(a) shows an image obtained by collecting events during 30ms from the present DVS when set to a contrast sensitivity of about 1.5% and observing a moving face. Fig. 1(b) shows an image obtained from the same scene, but when setting the DVS to about 10% contrast sensitivity. As can be seen, with improved contrast finer details are captured so that contours and textures are much better defined. On the other hand, improving contrast increases the number of output events by the same factor. Nevertheless, researchers have been persistently pursuing to improve contrast, as this not only improves recognition capability, but also opens new possible applications such as high speed fine texture based recognition. Recently, a single pixel for a DVS sensor that achieves a 0.3% contrast sensitivity (excluding inter-pixel mismatch) was reported [48]. This higher contrast sensitivity is achieved by a two-stage differencing amplifier [49], thus increasing the pixel gain by a factor 50. However, this improvement comes by dramatically reducing pixel bandwidth and therefore maximum pixel event rate. Consequently, the contrast improvement would be only useful for very slow scenes, but DVS cameras are very appealing for very high speed scenes. Another AER DVS with improved contrast sensitivity and lower latency has been reported recently [19]. By introducing a small area non-switched preampliflying stage biased in strong inversion, a slightly better contrast sensitivity (10%) was obtained while reducing the pixel area by one third. By using an alternative photo sensing stage, latency was reduced to 3.6μs. However, this design had a much higher power consumption, its mismatch FPN was also deteriorated, and intra-scene Dynamic Range (DR) was reduced from 120dB to only 3 decades, while overall DR was kept at 120dB. Note that an intra-scene DR of 3 decades is still a

quite good performance for many bright/shadow every day situations. The design presented in this paper uses new low-mismatch transimpedance preamplifying stages operating in subthreshold, coupled to the photo sensor through a source-driven active current mirror [50]. As a result, the achieved contrast sensitivity and power consumption have been improved by about one order of magnitude with respect to previous designs. Pixel area, dynamic range, latency, and bandwidth are similar to the previous designs [17]-[19],[23]-[25], and intra-scene DR is kept at 3 decades. Consequently, this sensor opens new possibilities in low to very high speed scenes with low contrast, allowing to capture a wider range of details not possible with prior DVS cameras.

II. A NEW LOW-POWER LOW-MISMATCH PRE-AMPLIFICATION STAGE FOR DVS

Fig. 2(a) shows the basic block diagram of a typical DVS pixel [17]-[19],[23]-[25]. The first stage transduces photo current to a voltage proportional to the logarithm of light

$$V_{log} = V_{DC} + A_v U_T \ln I_{ph} \tag{1}$$

where U_T is thermal voltage, A_v is a voltage gain factor, and V_{DC} is a light independent DC voltage level with high inter-pixel mismatch. The second stage amplifies V_{log} by C_1/C_2 resetting the charge integrated at C_2 every time $V_{diff} = (C_1/C_2)V_{log}$ varies a fixed threshold V_{th} set by the comparators. This also eliminates the DC component at V_{log}. The result is that each pixel generates a signed asynchronous output "event" every time its light changes by $\ln I_{ph}(t_2) - \ln I_{ph}(t_1) = \theta_{ev}$, with

$$\theta_{ev} = \frac{C_2 V_{th}}{C_1 U_T A_v} \tag{2}$$

Consequently, pixel information is obtained not synchronously at fixed time steps δt (as in conventional video sensors), but asynchronously, driven by data at fixed relative light increments $\theta_{ev} = |\ln(I_{ph}(t_2)/I_{ph}(t_1))|$, as shown in Fig. 2(b). The DVS outputs thus focus on (relevant) information change while reducing data throughput. Parameter θ_{ev} represents the minimum detectable temporal contrast

$$\theta_{ev} = \left|\ln\left(\frac{I_{ph}(t_2)}{I_{ph}(t_1)}\right)\right| = \left|\ln\left(1 + \frac{I_{ph}(t_2) - I_{ph}(t_1)}{I_{ph}(t_1)}\right)\right| \approx \left|\frac{\Delta I_{ph}}{I_{ph}}\right| \tag{3}$$

The compact pixel size imposes the use of simple comparators with offset voltage mismatch standard deviations around $10mV$, resulting in minimum practical values for $V_{th} \approx 50 - 100mV$. Contrast sensitivity must consequently be minimized by maximizing overall voltage gain $A_T = A_v C_1/C_2$.

Fig. 2(c) shows the original photo transduction stage [17]-[18],[23]-[24] with $A_v = n_n$, where n_n is the subthreshold slope factor of NMOS transistor M_{n1}. A reasonable overall voltage gain was obtained by setting $C_1/C_2 = 20$, but this resulted in about 50% pixel area occupation by capacitors (unless a MiM process is used [23]-[25]). Fig. 2(d) shows Leñero's [19] photo transduction and pre-amplification stage, where $A_v \approx 12$ helped to reduce capacitive spread to $C_1/C_2 = 5$ while improving overall voltage gain to about $A_v C_1/C_2 \approx 60$ with smaller area. Having the transistor gate voltage V_G tied to a constant bias voltage removes the Miller capacitance around amplifier A_1. This makes it possible to improve the bandwidth of the photo receptor stage with respect to that in Fig. 2(c) [19]. However, Leñero's [17] scheme has two major drawbacks. The first is that the gain of the pre-amplification stage was dependent on the relative aspect ratios between a PMOS and an NMOS transistor. As transistor aspect ratios suffer from high mismatch, the contrast sensitivity from pixel to pixel had higher mismatch than in Delbrück's implementation [17]-[18]. The second drawback is that for proper operation, the pre-amplification stage has to be biased in strong

inversion. Although transistors were properly sized to operate in strong inversion with moderate currents, the current consumption was still high (2.2μA/pixel), a factor 5 higher than in Delbrück's DVS.

Fig. 2(e) illustrates the novel concept used for the DVS being presented. Current mirror M_{p1}-M_{p2} amplifies photo current to $A_I I_{ph}$ [50] feeding a column of N diode-connected transistors M_{nj} , j = 1, ... N. This transistor column performs a transimpedance amplification from input current $A_I I_{ph}$ to output voltage V_{log}. Assuming each diode-connected transistor is biased in weak inversion and using for its drain-to-source current I_{nj} the approximation

$$I_{nj} = A_I I_{ph} \approx I_{sn} e^{\frac{V_{Gj} - V_{Sj}}{n_n U_T}} \tag{4}$$

results in

$$V_{log} = \sum_{j=1}^{N} (V_{Gj} - V_{Sj}) = V_{DC} + N n_n U_T \log I_p \tag{5}$$

with $V_{DC} = N n_n U_T \log(A_I / I_{sn})$. Consequently, the pre-amplification gain $A_v \approx n_n N$ is improved by a factor N. This improvement introduces no extra inter-pixel mismatch, although the finer analysis of the next Section reveals a second order mismatch component through body effect mismatch $\Delta\gamma_n$ with an increased pre-amplification gain. Voltage headroom limits the practical number of stacked transistors to N = 4.

III. DETAILED PIXEL DESIGN

In the previous Section we assumed that weak inversion transistor current could be approximated by eq. (4). This formulation was very helpful to conceive the circuit in Fig. 2(e) and obtain eq. (5). However, in this Section we will use a more precise expression for the weak inversion transistor current given by [51]-[52]

$$I_{nj} = I_{sn} e^{\frac{(V_{Gj}/n_n) - V_{Sj}}{U_T}} \tag{6}$$

When using eq. (6) instead of eq. (4), the expression for the transimpedance amplifier output V_{log} in Fig. 2(e) changes from that in eq. (5) to

$$V_{log} = \left(\sum_{j=1}^{N} n^j \right) U_T \log\left(\frac{A_I I_{ph}}{I_{sn}} \right) = V_{DC1} + \left(\sum_{j=1}^{N} {n_n}^j \right) U_T \log I_{ph} \tag{7}$$

Since n_n is sightly larger than unity, the result is that the preamplification gain $A_v = \sum {n_n}^j > N$ is further increased. However, n_n is weakly dependent on operating point and slightly sensitive to transistor mismatch. A widely accepted expression for the slope factor is [51]-[52]

$$n_n = 1 + \frac{\gamma_n}{2\sqrt{f(V_G)}} \tag{8}$$

where function $f(V_G)$ is approximately linear with gate voltage V_G and γ_n is the body effect factor. The second term in eq. (8) is smaller than unity (for the CMOS technology we used, it is typically smaller than 0.3), and consequently only weakly influences the value of n_n. Nonetheless, the immediate consequence is that the preamplification gain will be subject to a weak inter-pixel mismatch through body effect mismatch $\Delta\gamma_n$, since

$$\Delta n_n \approx \frac{\Delta \gamma_n}{\gamma_n}(n_n - 1) \tag{9}$$

Fig. 3 shows the detailed schematics of the pixel used. The photoreceptor block is composed of a photodiode connected to a source-driven active-input stage [19] that generates a voltage V_{ph} which is logarithmic with the photocurrent

$$V_{ph} = \frac{V_G}{n_p} + U_T \ln \frac{I_{ph}}{I_{sp}}, \tag{10}$$

where n_p and I_{sp} are, respectively, the substhreshold slope factor and the substhreshold current factor of PMOS transistor M_{p1}. Transistor M_{p3} delivers a copy of the photocurrent I_{ph} to a global summing node V_{avg} for the whole pixel array. As we will explain in the next Section, this current is used to adaptively bias the pixels, keeping them in proper operating conditions regardless of the illumination level.

The new preamplification stage comprises the cascade of two transimpedance amplifiers of the type shown in Fig. 2(e). Optionally, the second amplification stage can be disabled. Transistors M_{p1} and M_{p2} act as a source-driven current mirror [50] amplifying the photo current I_{ph} by factor A_I

$$A_I = e^{\frac{V_G - V_{GA}}{n_p U_T}} \tag{11}$$

As we will show in the next Section, V_{GA} is adjusted automatically for the whole array, dynamically accommodating the current amplification factor with the illumination conditions. This makes it possible to set an average value for the current through the diode-connected transistor chain $\overline{A_I I_{ph}}$ that is independent of average ambient light and that has a high enough value to guarantee a satisfactory speed response in the preamplifiers.

The first transimpedance stage uses 4 diode-connected transistors M_{n11}-M_{n14}. Output voltage of the first amplification stage V_{o1} is thus given by

$$V_{o1} = n_n(1 + n_n + n_n^2 + n_n^3)\left(U_T \log \frac{A_I I_{ph}}{I_{sn}}\right), \tag{12}$$

where the approximate value of the subthreshold slope for the NMOS transistors of our technology is $n_n \approx 1.2$. The first amplifying stage is therefore designed to implement an approximate gain $A_1 = n_n(1 + n_n + n_n^2 + n_n^3) \approx 6.4$.

For further amplification, voltage V_{o1} is converted back to a current I_q by means of transistor M_{nQ}. For proper operation, transistor M_{nQ} has to be biased in subthreshold. To keep transistor M_{nQ} in the subthreshold region, voltage V_Q is adapted globally for the whole array. The arrangement, which will be explained in the next Section, sets the DC mean value of current I_q to a given bias value I_{b2} independent of the illumination. Current I_q is replicated by PMOS current mirror M_{p4}-M_{p5} and enters the second transimpedance amplifier formed by transistors M_{n21}-M_{n23}, thereby obtaining a further amplification factor of $A_2 = 1 + n_n + n_n^2 \approx 3.6$. The complete expression for output voltage V_{log} is given by

$$V_{log} = -n_n A_2 V_Q + A_1 A_2 U_T \log \frac{A_I I_{ph}}{I_{sn}} = V_{DC} + A_1 A_2 U_T \log I_p, \tag{13}$$

where transistors M_{n1x}, M_{n2x} and M_{nQ} are assumed to be identical. The voltage gain factor A_v for this two stage preamplifier is therefore

$$A_v = A_1 A_2 = n_n(1 + n_n + n_n^2)(1 + n_n + n_n^2 + n_n^3) . \tag{14}$$

Voltage V_{log} feeds a capacitive differentiator stage, as proposed by Delbrück [17], so that

$$dV_{diff} = -\frac{C_1}{C_2} dV_{log} = -\frac{C_1}{C_2} A_1 A_2 U_T \frac{dI_{ph}}{I_{ph}}. \tag{15}$$

The capacitive ratio C_1/C_2 is designed to produce a further voltage amplification of 5. The total designed voltage gain in each pixel from V_{ph} to V_{diff} is therefore $A_T = (C_1/C_2)A_1A_2 \approx 120$. Note that this voltage gain depends only on the capacitive ratio C_1/C_2 and the subthreshold slope factors of the diode connected transistors. All of these parameters contribute with low mismatch, while the terms exhibiting high mismatch (such as current mirror gain A_I and specific current factors I_{sn}) vanish during the computation of the temporal difference. The two blocks furthest to the right in Fig. 3 detail the implementation of the capacitive differentiator and comparator stages, which are the same used by Delbrück and Posch [17]-[18],[23]-[25].

Let us call $V_{diffON} < 0$ the voltage excursion at node V_{diff} that generates a single positive event through the ON channel. Similarly, $V_{diffOFF} > 0$ is the voltage excursion at node V_{diff} that generates a single negative event through the OFF channel. If V_f is the voltage change caused by switching feedthrough at node V_{diff} during reset by transistor M_{sw}, we can state that

$$V_{diffON/OFF} = -V_f \pm \frac{n_p}{n_n} V_{\theta R} \tag{16}$$

where $V_{\theta R} = \left| V_{REF} - V_\theta^{+/-} \right|$. 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") $\theta_{ev}^{-} > 0$. By integrating equation (15),

$$\begin{aligned} \theta_{ev}^{+} &= \left| \ln \frac{I_{bright}}{I_{dark}} \right| = \left| \frac{V_{diffON}}{U_T A_T} \right| = \left| \frac{\frac{n_p}{n_n} V_{\theta R} + V_f}{U_T A_T} \right| \\ \theta_{ev}^{-} &= \left| \ln \frac{I_{dark}}{I_{brigth}} \right| = \left| \frac{V_{diffOFF}}{U_T A_T} \right| = \left| \frac{\frac{n_p}{n_n} V_{\theta R} - V_f}{U_T A_T} \right| \end{aligned} \tag{17}$$

The minimum detectable stimulus contrast can be adjusted through contrast sensitivity control voltage $V_{\theta R}$. The slope of θ_{ev} versus $V_{\theta R}$ reveals the effective value of the overall voltage gain A_T.

Inter-pixel contrast threshold mismatch $\Delta\theta_{ev}$ is caused by overall voltage gain mismatch ΔA_T, slope factor mismatch Δn_p and Δn_n (originated by body effect factor mismatch $\Delta\gamma_p$ and $\Delta\gamma_n$), switching offset mismatch ΔV_f, and comparators' offset mismatch (absorbed into term $\Delta V_{\theta R}$). Taking all these causes into account results in

$$\begin{aligned} \Delta\theta_{ev} &= \frac{n_p}{n_n} \frac{V_{\theta R}}{U_T A_T} \Delta a + \frac{1}{U_T A_T} \Delta v \\ \Delta a &= \frac{\Delta\gamma_p}{\gamma_p}(n_p - 1) - \frac{\Delta\gamma_n}{\gamma_n}(n_n - 1) - \frac{\Delta A_T}{A_T} \\ \Delta v &= \Delta V_f + \frac{n_p}{n_n} \Delta V_{\theta R} \pm V_f \frac{\Delta A_T}{A_T} \end{aligned} \tag{18}$$

where, mismatch voltage $\Delta V_{\theta R}$ absorbs the offset voltage of the comparators and ΔV_f represents the inter-pixel mismatch of the switching offset V_f. Note that expressions (17)-(18) are valid for all DVS using the comparator stages shown in the right hand side of Fig. 3, including Delbrück's [17]-[18], Posch's [23]-[25], and the present one. In the present design what differs is the

expression for A_T and its mismatch ΔA_T, which now includes mismatch contributions from capacitors and NMOS transistor slope factors. Eq. (18) has two main terms. The first is linear with control voltage $V_{\theta R}$ and the second is constant. The standard deviation of the contrast threshold mismatch $\Delta\theta_{ev}$ is therefore,

$$\sigma(\theta_{ev}) = \frac{\sigma_v}{U_T A_T}\sqrt{1+\left(\frac{n_p}{n_n}V_{\theta R}\frac{\sigma_a}{\sigma_v}\right)^2} \tag{19}$$

where $\sigma_a = \sigma(\Delta a)$ and $\sigma_v = \sigma(\Delta v)$ are constant. It is interesting to note that overall mismatch is reduced by increasing total voltage gain A_T. The form of eq. (19) is shown in Fig. 4. Mismatch changes with $V_{\theta R}$ setting, showing two asymptotes. For low $V_{\theta R}$ values $\sigma(\theta_{ev}) = \sigma_v/(U_T A_T)$, while for large $V_{\theta R}$ values $\sigma(\theta_{ev})$ is linear with $V_{\theta R}$, and its slope reveals the value of σ_a.

IV. SYSTEM DESCRIPTION AND GLOBAL ILLUMINATION ADAPTATION

The complete system architecture chip block diagram is shown in Fig. 5. The system is composed of a 128x128 pixel array, a global adaptive biasing cell that detects the spatio-temporal average of the illumination over the whole array adapting the DC levels of global voltage biases V_{GA} and V_Q, a set of programmable current sources that allows all the current biases needed by the chip to be fine tuned by accessing a very reduced set of pins [54], and finally the row and column address event communication circuitry that generates the output event addresses and signs. The AER read-out scheme implemented in this design is Boahen's row parallel technique which latches all the events generated simultaneously in a row and reads them out sequentially, significantly speeding up the read out process when high event rates have to be managed [55].

The schematics of the global adaptive biasing cell are detailed in Fig. 6. The cell replicates the average photo current $\overline{I_{ph}}$ obtained from the sum of all pixel photo current replicas. Capacitor C_{avg} at the global summing node V_{avg} sets the adaptation time constant. A spatio-temporal average of photocurrent $\overline{I_{ph}}$ is thus generated and fed to a replica of the pixel photoreceptor circuit. The adaptive biasing cell also contains a replica of the first voltage amplification stage (transistors M_{p2} and M_{n11}-M_{n14}). Transistor M_{pc} is another replica of pixel transistor M_{p2} in Fig. 3. Its role is to adapt the gate voltage V_{GA} in such a way that bias current I_{b1} flows through transistor M_{p2} in Fig. 6. This gate voltage V_{GA} is buffered to all pixels in the array, thus setting the current gain A_I of all pixels. With this arrangement, the average current through the first voltage amplifier in the pixels is set to current I_{b1}, thereby dynamically adapting current gain A_I depending on average photo current $\overline{I_{ph}}$

$$A_I = I_{b1}/\overline{I_{ph}} \tag{20}$$

If array illumination increases, current gain A_I would decrease. On the contrary, if array illumination decreases, current gain A_I would increase.

The global adaptive biasing cell also contains a replica of transistor M_{nQ} (see Fig. 3) with its gate voltage biased to the average output voltage of the first amplification stage $\overline{V_{o1}}$. Transistor M_{nQ} is configured as a source-driven active-input stage [50] biased with a current I_{b2}. The voltage at its source V_Q adapts so that the average current of the pixel second amplifier stages is always set to I_{b2}.

V. EXPERIMENTAL RESULTS

A prototype was fabricated in the double-poly 4-metal 0.35μm CMOS AMS technology, with the OPTO option[1]. The fabricated retina had a resolution of 128x128 pixels and occupied a total area of 4.9x4.9mm^2 including the pads. Fig. 7 shows a

1. The OPTO option adds an EPI substrate layer to reduce dark currents, and an ARC (anti reflective coating) layer.

Table 1. Summary and Comparison of DVS Characteristics

	This work	Leñero et al.[19]	Lichtsteiner et al.[17]-[18]	Posch et al. [23]-[25]
Technology	0.35μm 4M 2P	0.35μm 4M 2P	0.35μm 4M 2P	0.18μm 4M 2P MIM
Resolution	128x128	128x128	128x128	304x240
Chip area (mm^2)	4.9x4.9	5.5x5.6	6x6.3	9.9x8.2
Pixel area (λ^2)	155x150 (30x31μm^2)	175x175 (35x35μm^2)	200x200 (40x40μm^2)	222x222[a] (20x20μm^2)
Fill factor (%)	10.5	8.7	8.1	10-20
Power @ 100keps (mW)	4	132	24[b]	75[c]
Power/Resolution[d] (μW)	0.24	8.06	1.46	1.03
Latency (μs)	3.0	3.6	15	3.0
Contrast Sensitivity (%)	1.5	10	15	9
FPN (% contrast)	0.9	4.0	2.1	1.9
DR	120dB	100dB	120dB	125dB
Intrascene DR	60 dB	56dB	120dB	125dB
Maximum Event Rate (Meps)	20	20	1	N/A

a. only the DVS part of the pixel is considered
b. this chip includes a non-optimum bias generator which by itself consumes 2/3 of total chip power
c. only the DVS functionality was active
d. Power@100keps normalized to total number of pixels

micro photograph of the fabricated prototype, including an inset with the pixel layout. Pixel size was 30x31μm^2. Pixels were arranged in a symmetrical way sharing analog and digital routing channels. Noise coupling between analog and digital parts was therefore minimized. Digital lines were routed horizontally over digital parts using metal 2, while sensitive analog lines were routed horizontally over analog parts. Metal 4 was reserved to cover the full array except for the photodiode. Table 1 summarizes the main design specifications [56] and compares them with previously reported designs. The retina was tested with a 16mm F/1.4 C-mount lens. The minimum achieveable contrast sensitivity and the power consumption were improved by about one order of magnitude with respect to previous DVS designs [17]-[18], [19], [23]-[25]. Pixel area was reduced by about 30-50%, and FPN by 25-50% with respect to previous designs, while latency was kept equal to the best previously reported values. Overall dynamic range (DR) was slightly less than the best reported. Maximum output event rate was kept high at 20*Meps* thanks to Boahen's burst-mode row parallel AER read out scheme [55]. The only penalty was that intra-scene dynamic range was limited to about 3 decades (60dB).

Fig. 8 shows several images generated by collecting n_{ev} events during a time ΔT illustrating contrast sensitivity, high speed capability, and wide intrascene dynamic range. Fig. 8(a-c) illustrate different slow moving scenes. Fig. 8(d) was captured using Edmund's 0.1 contrast density chart. Fig. 8(e) is similar to Fig. 8(d) but with an illumination ratio of 1:1000 produced by inserting neutral density filters with ambient fluorescent illumination. Fig. 8(f) illustrates high speed capability under ambient illumination with ΔT=623 s,when a 52 card deck was browsed in about 0.5s. Fig. 8(g) shows an analog oscilloscope in x/y mode displaying a

decaying amplitude pair of sinewaves, with 90º phase shift and 5KHz oscillation frequency. Thus, one revolution cycle on the screen takes 200μs. The DVS produces events with submicrosecond resolution that track the oscilloscope trace, while ambient light is that produced by the oscilloscope screen. Fig. 8(g) shows the events captured by the DVS during 38μs, which corresponds approximately to an arc of one fifth of a revolution.

A. Uniformity of Response, Minimum Detectable Contrast, and Pixel Gain Characterization.

To characterize the uniformity of the response to a given contrast, a similar procedure to the one developed by Lichtsteiner and Delbrück [17] was followed. In order to stimulate all pixels uniformly, a moving gradient bar (shown in Fig. 9(a)) was presented to the retina and the number of positive and negative events generated by each pixel was recorded. The bar crossed the screen in about 4s. For this we used a TFT monitor providing an illumination of 60 lux, measured at the retina position. The experiment was repeated for different settings of the pixel threshold control voltage $V_{\theta R}$. The bar was swept 30 times for each setting. This way, for each pixel (x, y) we obtain the corresponding number of events $N(x, y)$ fired per edge presentation for the different settings of $V_{\theta R}$. The number of events generated by a pixel for a given stimulus contrast θ and $V_{\theta R}$ setting is given by

$$N(x, y) = \frac{\theta}{\theta_{ev}(x, y)} \tag{21}$$

Consequently, given the applied input contrast θ and the measured values of $N(x, y)$, we obtain for each pixel the contrast sensitivity $\theta_{ev}(x, y) = \theta / N(x, y)$, which will be a function of the threshold voltage setting $V_{\theta R}$ (see eq.(17)). Uniformity in $\theta_{ev}(x, y)$ was characterized for one and two preamplifier stages. For one stage we used a stimulus with a contrast $\theta = \ln(I_{bright}/I_{dark}) = \ln(56\text{lux}/18\text{lux}) = 1.135$, while for two stages we used a lower contrast stimulus $\theta = \ln(I_{bright}/I_{dark}) = \ln(58\text{lux}/45\text{lux}) = 0.254$. This was done to keep the collected number of events similar.

Fig. 9(b) shows several superimposed histograms representing the collected number of events per pixel and contrast edge presentation for different settings of control voltage $V_{\theta R}$, when using two preamplifier stages. By averaging over all the pixels, we obtain the average number of events $\overline{N}$ produced per pixel per contrast edge presentation, and from eq. (21) we obtain the average pixel contrast sensitivity $\overline{\theta_{ev}}$. Standard deviations $\sigma(N)$ and $\sigma(\theta_{ev})$ can be obtained in a similar manner. Fig. 9(c) represents $\overline{\theta_{ev}}$ (in % contrast) versus $V_{\theta R}$, when using one and two preamplifier stages. Similar graphs of previously reported DVS devices have also been included. According to eq. (17), the slopes of these graphs reveal the effective value of $\overline{A_T}$, which was $\overline{A_T} = 95$ for the 2-stage case and $\overline{A_T} = 18$ for the 1-stage case. Fig. 9(d) shows the resulting standard deviation $\sigma(\theta_{ev})$ (in % contrast) versus the average contrast threshold $\overline{\theta_{ev}}$ (in % contrast). Again, the 1-stage and 2-stage values are shown together with those of previously reported DVS devices. It can be seen how these curves resemble the one in Fig. 4. The value of σ_a, which was $\sigma_a = 28\%$ for the 2-stage case and $\sigma_a = 12.5\%$ for the 1-stage case were obtained from the slopes of the diagonal asymptotes. The value of σ_v could not be determined because the horizontal asymptote was not visible. However, we were able to determine an upper bound for σ_v from the 1-stage case, since $\sigma_v < U_T A_T \sigma(\theta_{ev})_{min} = 26mV \text{ x} 18 \text{ x } 2\% = 9.4mV$.

B. Dynamic Range

The operation of the retina was verified for over 120dBs of illumination change. The retina was tested for bright illumination (about 100Klux) down to illuminations below 0.1lux. Maximum laboratory ambient illumination (about 10Klux) was achieved using two 26W compact fluorescent light (CFL) bulbs placed 40*cm* apart from the retina (scene illumination was measured with an RS 180-7133 digital lightmeter located at the retina position). Neutral density filters were then used to progressively darken the

field of view of the retina. We used Newport FS-ND broadband density filters, which guarantee flat transmission from ultraviolet to near infrared.

The lower illumination limit of the retina is limited by photodiode dark current, which is $3fA$. We were able to measure the photodiode dark current with good precision in our chip by collecting the total current flowing through node V_{gnd01} , which is shared by all the pixels in the array. The total current measured in darkness through node V_{gnd01} was $50pA$ which is well above the leakage current for the pads and wiring of $0.76pA$. A 60lux scene illumination (measured with the digital lightmeter at the retina position) produced a measured current through node V_{gnd01} of $69nA$. The measured dark current was therefore equivalent to a scene illumination of 0.04luxes, which is consistent with the measured dynamic range.

The intra-scene illumination range of the retina is limited, since the preamplifying stage self-adapts to one single global illumination level. Nevertheless, an intra-scene illumination dynamic range of up to 60dB was verified. Fig. 8(e) shows the response of the retina with half of the visual field at 0.5lux while the other half was illuminated at 500lux by inserting a neutral density filter vertically over half of the field of view and placing the setup in a dark chamber to avoid reflections.

C. Pixel Bandwidth

To measure the pixel bandwidth a group of pixels located in the center of the array were stimulated with a flashing Kingbright super bright red LED L-793SRC-B (1400mcd@20mA) following a procedure similar to that developed by Lichtsteiner and Delbrück [17]. We created a controlled scene illumination by placing two CFL bulbs near the retina but not directly focused on it. A scene illumination of 2klux was measured by placing the light meter RS-180-7133 in the retina position. The LED was placed in front of the retina. The LED diode was modulated with a sinusoidal signal

$$V(t) = V_{off} + A\sin 2\pi ft \tag{22}$$

We varied the frequency of the sinusoid and counted the number of positive and negative events generated per sinewave period by each stimulated pixel. The measurements were averaged over 30 periods. If the magnitude of the stimulus contrast and the voltage thresholds are kept constant, the pixel gain is directly proportional to the number of generated events (see eqs. (17) and (21)). The measurements were repeated for different values of the illumination by inserting neutral density filters of different attenuation values. These neutral density filters were placed just in front of the retina, thus attenuating both simultaneously the scene and the LED illuminations.

Fig. 10(a) plots the measured values of the events per cycle (averaged over 30 periods) as a function of the frequency of the sinusoidal signal for different illumination values. The curve marked with crosses corresponds to the measurements at the highest 2klux illumination. The curves marked with triangles, circles and diamonds correspond respectively to 200lux, 20lux and 2lux illumination. Each of these curves was fitted to a first order (single pole) transfer function. Fig. 10(a) also shows the fitted first order functions (asterix curves). The resulting fitted poles are plotted in Fig. 10(b) versus illumination. As can be observed, the bandwidth is approximately linear with photocurrent (light) over the measured illumination range. We can therefore conclude that the photoreceptor stage is the one limiting pixel bandwidth. Pixel bandwidth at 2klux illumination is slightly higher than $10KHz$.

D. Latency

Latency is the delay which elapses from the occurrence of an illumination change at the photodiode until the corresponding output event is transmitted off-chip. To measure latency, we used the super bright LED stimulated with a step signal [17]. We measured the latency as the delay between the step in the LED signal and the first output request corresponding to an event with an address from the stimulated pixels. As in the previous experiment, we created an intense scene illumination by placing two CFL bulbs near the retina but not directly focused on it. A scene illumination of 2klux was measured by placing the light meter

RS-180-7133 in the retina position. The LED was placed near the retina. Neutral density filters were placed just in front of the retina, thus simultaneously attenuating both the scene and the LED illuminations. We performed the measurements for different values of the illumination. Each measurement was repeated and averaged over a total of 30 times. Fig. 11 shows the measured latency versus illumination. As can be observed, the latency is inversely proportional to the illumination. For very low illumination (0.2 lux) the measured latency was $6ms$, while for a high level of illumination (2 Klux) the measured latency was 3.2μs. The only part of the pixel whose latency depends on illumination is the photoreceptor stage in Fig. 3. The rest of the circuitry introduces a delay which is independent of illumination but dependent on biasing. We performed the latency measurements for both of the possible configurations of the preamplifying stage: that is, with only one preamplification stage activated and with the two preamplification stages active. We obtained very similar results for both cases.

E. Noise Characterizations

To characterize noise, we followed a similar procedure to that reported by Posch [25]. A 10x10 pixel region of interest was selected. A moving gradient bar of a given contrast was presented to the region and the events generated by those pixels were collected for one hundred repetitions of the stimulus presentation. The experiment was repeated as the contrast of the bar was increased. For each pixel and for each contrast value a probability of generating at least one event was computed over the 100 stimulus presentations. For each pixel, the event generation probability was represented versus the stimulus contrast and an S-shaped curve was obtained [25]. The RMS equivalent input contrast noise for the pixel was obtained by fitting the S-shaped curves with a Gaussian error function. Fig. 13 plots the RMS equivalent contrast noise average value obtained for the pixels in the selected region of interest for different illumination values. As can be observed an equivalent contrast RMS noise between 2.6% and 2.1% was measured. The noise slightly decreases as illumination increases.

Although RMS noise is slightly higher than FPN noise (mismatch), it is interesting to mention that random temporal noise is filtered out when using this type of sensor in event-driven vision processing systems for e.g. object recognition. For example, when feeding the DVS output to event-driven convolution chips programmed with e.g. Gabor kernels [6]-[7] for orientation extraction, noise is not present any more at the filter output. However, event-driven convolution chips do detect spatio-temporal features to perform object recognition, and these features are sensitive to the sensor uniformity (e.g. mismatch or FPN). Therefore, FPN is indeed relevant for event-driven processing, while random temporal noise is not that relevant.

F. Power Consumption

Chip total power consumption changes with output event rate, as shown in Fig. 12. The total chip power consumption was measured for the cases of one and two preamplification stages. It ranged between $2.6mW$ at $1keps$ (kilo events per second) and $95mW$ at a peak output rate of $20Meps$ (mega events per second), from a $3.3V$ power supply. At a nominal output rate of $100keps$, power consumption was $3.96mW$. For high event rates (above $200Keps$) there is a rapid increase in power consumption with event rate because now the contribution of the digital output pads (and all the AER communication circuitry) exceeds that of the internal analog parts. In the present design, thanks to the new low power preamplification stages operating in the subthreshold regime, the nominal rate power consumption has been reduced by a factor of 80 with respect to the previous design [19] which used preamplifiers biased in strong inversion.

VI. CONCLUSIONS

A 128x128 pixel Dynamic Vision Sensor (DVS) has been fabricated in a 0.35μm technology. Its novelty resides in the use of a transimpedance amplification stage, operated in weak inversion, but which introduces low extra mismatch. The internal voltage amplification is improved by a factor of about 6 with respect to Delbrück's and Posch's DVS, while reducing the area

requirements of the capacitive based amplification by a factor of 4. As a result, pixel area and power are reduced and pixel gain is improved, resulting in higher contrast sensitivity. The presented design outperforms previously reported DVS by achieving an input contrast sensitivity as low as 1.5% with a current consumption of 50*nA* per pixel. The achieved 10*kHz* pixel bandwidth and 120dB Dynamic Range are the same as in the previously reported prototypes. The 3.2μ*s* latency is kept equal to the best reported so far. A maximum output event rate of 20Meps is achieved. The only penalty is a reduced intra-scene dynamic range (of 3 decades) with respect to Delbrück's and Posch's DVSs.

VII. ACKNOWLEDGEMENTS

This work has been supported by EU grant FP7-ICT-2007-1-216777 (NABAB), ERA-NET grant PRI-PIMCHI-2011-0768 (PNEUMA) funded by the Spanish Ministerio de Ciencia e Innovación, Spanish research grants (with support from the European Regional Development Fund) TEC2009-10639-C04-01 (VULCANO), and Andalusian research grants P06-TIC-1417 (Brain System) and TIC-6091 (NANONEURO). The authors are very grateful to Tobi Delbrück for his highly valuable indications and for providing the jAER software infrastructure [57], the ATC group of the University of Seville for providing the AER interfacing PCBs developed during the CAVIAR project, Philipp Häfliger for providing the AER test board and lens mount holder, and Christoph Posch for valuable feedback. Authors are also grateful to J. A. Leñero for performing some initial tests on the fabricated prototype.

VIII. REFERENCES

[1] M. Sivilotti, *Wiring Considerations in Analog VLSI Systems with Application to Field-Programmable Networks*, Ph.D. Thesis, California Institute of Technology, Pasadena CA, 1991.
[2] M. Mahowald, *VLSI analogs of neural visual processing: a synthesis of form and function*, Ph. D. Thesis, California Institute of Technology, Pasadena, 1992.
[3] J. Lazzaro, J Wawrzynek, M Mahowald, M Sivilotti, D. Gillespie, "Silicon Auditory Processors as Computer Peripherals," *IEEE Transactions on Neural Networks*, vol. 4, no. 3, pp. 523-528, 1993.
[4] P. Venier, A. Mortara, X. Arreguit, and E. A. Vittoz, "An Integrated Cortical Layer for Orientation Enhancement," *IEEE J. Solid-State Circuits*, vol. 32, no. 2, pp. 177-186, Feb. 1997.
[5] T. Y. W. Choi, P. Merolla, J. Arthur, K. Boahen and B. E. Shi, "Neuromorphic Implementation of Orientation Hypercolumns," *IEEE Tran. on Circ. and Syst. Part I*, vol 52, no 6, pp. 1049-1060, June 2005.
[6] L. Camuñas-Mesa, A. Acosta-Jiménez, C. Zamarreño-Ramos, T. Serrano-Gotarredona, and B. Linares-Barranco, "A 32x32 Pixel Convolution Processor Chip for Address Event Vision Sensors with 155ns Event Latency and 20Meps Throughput," *IEEE Trans. Circ. and Syst. Part-I*, vol. 58, No. 4, pp. 777-790, April 2011.
[7] L. Camuñas-Mesa, C. Zamarreño-Ramos, A. Linares-Barranco, A. Acosta-Jiménez, T. Serrano-Gotarredona, and B. Linares-Barranco, "An Event-Driven Multi-Kernel Convolution Processor Module for Event-Driven Vision Sensors," *IEEE J. of Solid-State Circuits*, in Press, Feb. 2012.
[8] C. Zamarreño-Ramos, A. Linares-Barranco, T. Serrano-Gotarredona, and B. Linares-Barranco, "Multi-Casting Mesh AER: A Scalable Assembly Approach for Reconfigurable Neuromorphic Structured AER Systems. Application to ConvNets," *IEEE Trans. on Biomedical Circuits and Systems*, in Press.
[9] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, "A biomorphic digital image sensor," *IEEE J. of Solid-State Circ.*, vol. 38, pp. 281-294, 2003.
[10] D. Kim and E. Culurciello, "A Compact-pixel Tri-mode Vision Sensor," IEEE Int. Symp. on Circuits and Systems, (ISCAS), Paris, France, pp. 2434 – 2437, 2010.
[11] P. F. Ruedi et al., "A 128x128 pixel 120-dB dynamic-range vision sensor chip for image contrast and orientation extraction," *IEEE J. of Solid-State Circ.*, vol. 38, pp. 2325-33, 2003.
[12] M. Barbaro, P. Y. Burgi, A. Mortara, P. Nussbaum, and F. Heitger, "A 100x100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding," *IEEE J. of Solid-State Circ.*, vol. 37, pp. 160-172, 2002.
[13] S. Chen, and A. Bermak, "Arbitrated time-to-first spike CMOS image sensor with on-chip histogram equalization," *IEEE Trans. on VLSI Systems*, vol. 15, no. 3, 346 - 357. Mar. 2007.
[14] X. G. Qi, J. Harris, "A time-to-first-spike CMOS imager," *Proc. of the IEEE Int. Symp. on Circ. and Syst.* (ISCAS), vol. 4, pp. 824-827, 2004.
[15] M. Azadmehr, H. Abrahamsen, and P. Hafliger, "A foveated AER imager chip," *Proc. of the IEEE Int. Symp. on Circ. and Syst.* (ISCAS), vol. 3, pp. 2751-2754, 2005.
[16] R. J. Vogelstien, U. Mallik, E. Culurciello, R. Etienne-Cummings, and G. Cauwenberghs, "Spatial acuity modulation of an address-event imager," *IEEE Int. Conf. on Electr., Circ. and Syst.* (ICECS), pp. 207-210, 2004.
[17] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128x128 120 dB 15μs latency asynchronous temporal contrast vision sensor," *IEEE J. of Solid-State Circ.*, vol. 43, no. 2, pp. 566-576, Feb. 2008.
[18] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128x128 120dB 30mW asynchronous vision sensor that responds to relative intensity change," in *IEEE Int. Solid-State Circ. Conf.* (ISSCC) *Dig. of Tech. Papers*, pp. 2060-2069, 2006.
[19] J. A. Leñero-Bardallo, T. Serrano-Gotarredona and B. Linares-Barranco, "A 3.6μs latency asynchronous frame-free event-driven dynamic-vision-sensor," in *IEEE Journal of Solid-State Circuits*, vol. 46, No. 6, pp. 1443-55, June, 2011.
[20] J. Kramer, "An integrated optical transient sensor," *IEEE Trans. on Circ. and Syst., Part II*, vol. 49, no. 9, pp. 612-628, Sep. 20002.
[21] K. A. Zaghloul, and K. Boahen, "Optic nerve signals in a neuromorphic chip I: outer and inner retina models," *IEEE Trans. on Biom. Eng.*, vol. 51, no. 4, pp. 657-666, Apr. 2004.
[22] K. A. Zaghloul, and K. Boahen, "Optic nerve signals in a neuromorphic chip II: testing and results," *IEEE Trans. on Biom. Eng.*, vol. 51, no. 4, pp. 667-675, Apr. 2004.
[23] C. Posch, D. Malotin, and R. Wohlgenannt, "A QVGA 143dB Dynamic Range Asynchronous Address-Event PWM Dynamic Image Sensor with Lossless Pixel-Level Video Compression," *IEEE ISSCC Dig. Tech. Papers*, pp. 400-401, Feb. 2010.
[24] C. Posch, D. Matolin, and R. Wohlgenannt, "A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 259-275, January 2011.
[25] C. Posch, and D. Matolin, "Sensitivity and Uniformity of a 0.18mm CMOS temporal contrast pixel array," *Proc. of the IEEE Int. Symp. on Circ. and Syst.* (ISCAS), pp. 1572-1575, May 2011.

[26] R. Berner and T. Delbrück, "Event-Based Pixel Sensitive to changes of Color and Brightness," *IEEE Trans. on Circ. and Syst. Part I*, vol. 58, no. 7, pp. 1581-1590, July 2011.
[27] M. Arias-Estrada, D. Poussart, and M. Tremblay, "Motion vision sensor architecture with asynchronous self-signalling pixels," *Workshop on Computer Architecture for Machine Perception*, pp. 75-83, 1997.
[28] C. M. Higgins and S. A. Shams, "A biologically inspired modular VLSI system for visual measurement of self-motion," *IEEE Sensors Journal*, vol. 2, no. 6, pp. 508-528, Dec. 2002.
[29] E. Ozalevli and C. M. Higgins, "Reconfigurable biologically inspired visual motion system using modular neuromorphic VLSI chips," *IEEE Trans. on Circ. and Syst. Part I*, vol. 52, no. 1, pp. 79-92, 2005.
[30] K. Boahen, and A. Andreou, "A contrast-sensitive retina with reciprocal synapses," *Advances in Neural Information Processing Systems* (NIPS), vol. 4, pp. 764-772, 1992.
[31] J. Costas-Santos, T. Serrano-Gotarredona, R. Serrano-Gotarredona and B. Linares-Barranco, "A spatial contrast retina with on-chip calibration for neuromorphic spike-based AER vision systems," *IEEE Transactions on Circuits and Systems, Part I*, vol. 54, no. 7, pp. 1444-58, 2007.
[32] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, "A five-decade dynamic range ambient-light-independent calibrated signed-spatial-contrast AER retina with 0.1ms latency and optional time-to-first-spike mode," *IEEE Trans. on Circ. and Syst. Part I*, vol. 57, no. 10, pp. 2632-2643, Oct. 2010.
[33] J. Conradt, P. Lichtsteiner, R. Berner, T. Delbruck, RJ. Douglas, M. Cook, "A pencil balancing robot that uses only spike-based visual input," *Proc. IEEE International Symp. Circ. and Syst.* (ISCAS 2009), pp. 781-785, 2009.
[34] D. Drazen, P. Lichtsteiner, P. Häfliger, T. Delbrück, and A. Jensen, "Toward real-time particle tracking using an event-based dynamic vision sensor," *Exp. Fluids*, vol. 51, pp. 1465-1469, 2011.
[35] R. Benosman, S.-H. Ieng, P. Rogister, and C. Posch, "Asynchronous Event-Based Hebbian epipolar Geometry," *IEEE Trans. Neural Networks*, in Press.
[36] S. Schraml and A. N. Belbachir, "A Spatio-Temporal Clustering Method Using Real-Time Motion Analysis on Event-Based 3D Vision," *Proc. of the IEEE Computer Vision and Pattern Recognition Workshops* (CVPRW), pp. 57-63, 2010.
[37] C. Sulzbachner, C. Zinner, and J. Kogler, "An Optimized Silicon Retina Stereo Matching Algorithm Using Time-Space Correlation," *Proc. of the IEEE Computer Vision and Pattern Recognition Workshops* (CVPRW), 2011.
[38] R. Benosman, S.-H. Ieng, C. Clercq, C. Bartolozzi, and M. Srinivasan, "Asynchronous Frameless Event-based Optical Flow," *Neural Networks*, 2011 in Press.
[39] S. Chen, P. Akselrod, B. Zhao, J. A. Pérez-Carrasco, B. Linares-Barranco and E. Culurciello, "Efficient feedforward categorization of objects and human postures with address-event image sensors," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, in Press.
[40] J. A. Pérez-Carrasco, B. Acha, C. Serrano, L. Camuñas-Mesa, T. Serrano-Gotarredona, and B. Linares-Barranco, "Fast Vision through Frame-less Event-based Sensing and Convolutional Processing. Application to Texture Recognition," *IEEE Trans. Neural Networks*, vol. 21, No. 4, pp. 609-620, April 2010.
[41] J. Pérez-Carrasco, T. Serrano-Gotarredona, C. Serrano, B. Acha, and B. Linares-Barranco, "High-speed character recognition system based on a complex hierarchical AER architecture," *Proc. IEEE Int. Symp. on Circ. and Syst.* (ISCAS), pp. 2150-2153, May 2008.
[42] U. Mallik, M. Clapp, E. Choi, G. Cauwenberghs, and R. Etienne-Cummings, "Temporal Change Threshold Detection Imager," in *IEEE Int. Solid-State Circ. Conf.* (ISSCC) *Dig. of Tech. Papers, vol. 1*, pp. 362-603, 2005.
[43] Y, M. Chi, U. Mallik, M.A. Clapp, E. Choi, G. Cauwenberghs, and R. Etienne-Cummings, "CMOS camera with in-pixel temporal change detection and ADC", *IEEE J. of Solid-State Circ.*, vol. 43, no. 10, 2187-2196, Oct. 2008.
[44] V. Gruev, and R. Etienne-Cummings, "A pipelined temporal difference imager," *IEEE J. of Solid-State Circ.*, vol. 39, no. 3, pp. 538-543, Mar. 2004.
[45] D. Kim, Z. Fu, J. H. Park, and E. Culurciello, "A 1-mW CMOS temporal-difference AER sensor for wireless sensor networks," *IEEE Trans. on Elec. Devices*, vol. 56, no. 11, pp. 2586-2593, Nov. 2009.
[46] M. Gottardi, N. Massari, and S. A. Jawed, "A 100mW 128x64 pixels contrast-based asynchronous binary vision sensor for sensor networks applications," *IEEE J. of Solid-State Circ.*, vol. 44, no. 4, pp. 1582-1592, May 2009.
[47] S. Chen, W. Tang, X. Zhang, and E. Culurciello, "A 64 64Pixels UWB Wireless Temporal-Difference Digital Image Sensor," *IEEE Trans. on VLSI Systems*, vol. 20, No. 12, pp. 2232-2240, Dec. 2011.
[48] T. Delbruck, and R. Berner, "Temporal contrast AER pixel with 0.3%-contrast event threshold," *Proc. of the IEEE Int. Symp. on Circ. and Syst.* (ISCAS), pp. 2442-2445, 2010.
[49] C. Posch, D. Matolin, and R. Wohlgenannt, "A two-stage capacitive-feedback differencing amplifier for temporal contrast IR sensors," *Int. J. of Analog Int. Circ. and Signal Proc.*, vol. 64, no.1, pp. 45-54, July 2010.
[50] T. Serrano-Gotarredona, B. Linares-Barranco, and A. G. Andreou , "Very Wide Range Tunable CMOS/Bipolar Current Mirrors with Voltage Clamped Input," *IEEE Trans. Circuits and Systems (Part I): Fundamental Theory and Applications*, pp. 1398-1407, November 1999.
[51] C. C. Enz, F. Krummenacher and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low current applications," *Int. J. of Analog Int. Circ. and Signal Proc.*, no. 8, pp. 83-114, 1995.
[52] A.I.A. Cunha, M.C. Schneider, and C. Galup-Montoro, "An MOS transistor model for analog circuit design," *IEEE J. of Solid-State Circ.*, vol. 33, no. 10, 1510-1519, Oct. 1998.
[53] R. Serrano-Gotarredona, T. Serrano-Gotarredona, A. Acosta-Jiménez and B. Linares-Barranco, "A Neuromorphic Cortical Layer Microchip for Spike Based Event Processing Systems," *IEEE Trans. on Circ. and Syst., Part I*, vol. 52, no. 12, pp. 2548-2566, Dec. 2006.
[54] R. Serrano-Gotarredona, L. Camuñas-Mesa, T. Serrano-Gotarredona, J. A. Leñero-Bardallo, and B. Linares-Barranco, "The stochastic I-Pot: A circuit building block for programming bias currents," *IEEE Trans. on Circ. and Syst., Part II*, vol. 19, no. 7, pp. 1196-1219, July 2008.
[55] K Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Trans. on Circ. and Syst., Part II*, vol. 47, no. 5 pp. 416-434, May 2000.
[56] T. Delbrück, B. Linares-Barranco, E. Culurciello, and C. Posch, "Activity-driven, event-based vision sensors," *Proc. of the IEEE Int. Symp. on Circ. and Syst.* (ISCAS), pp. 2426 - 2429, 2010.
[57] jAER Open Source Project, available at http://sourceforge.net/apps/jaer/wiki.

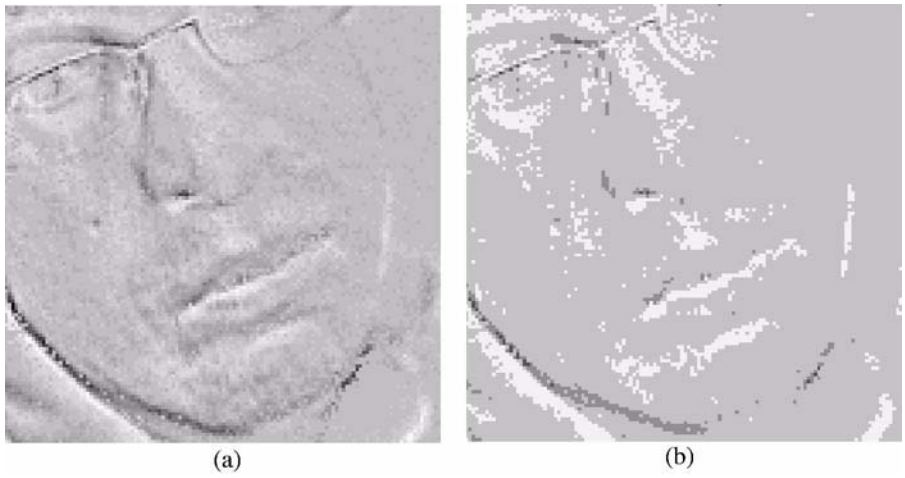

Fig. 1: Illustration of effect of improving contrast sensitivity in a DVS when observing a moving face. (a) Reconstructed image when collecting during 30ms the events produced by the present DVS when contrast sensitivity is set to 1.5%. (b) Same scene, but now the sensor is set to a contrast sensitivity of about 10%.

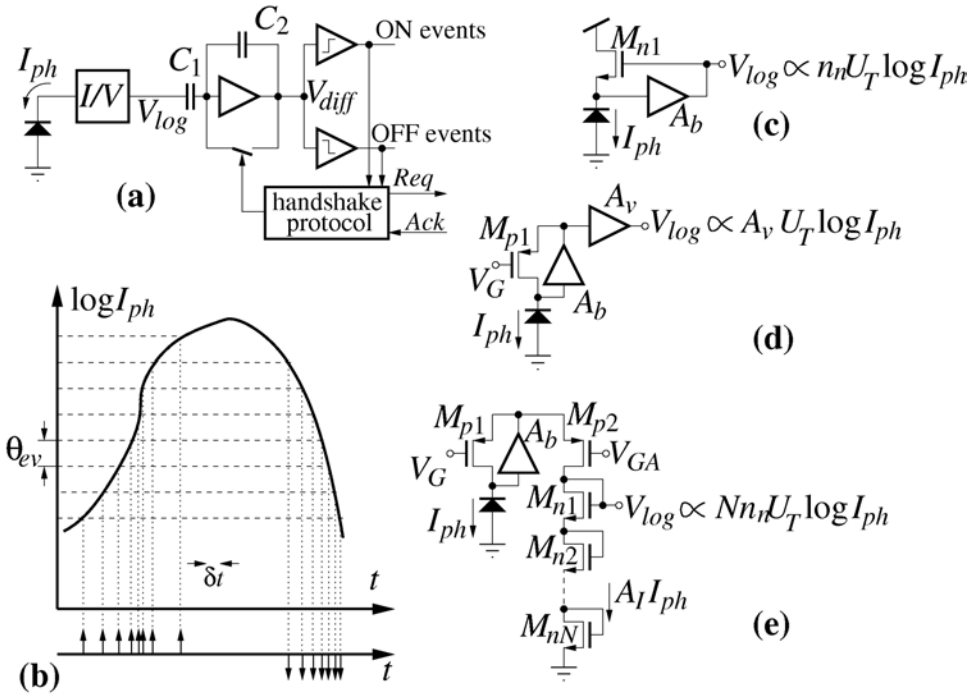

Fig. 2: (a) Pixel block diagram, (b) data driven asynchronous event generation, (c) Delbrück's original photo current transduction circuit, (d) Leñero's transduction with mismatch sensitive pre-amplification, (e) proposed transduction with mismatch insensitive pre-amplification.

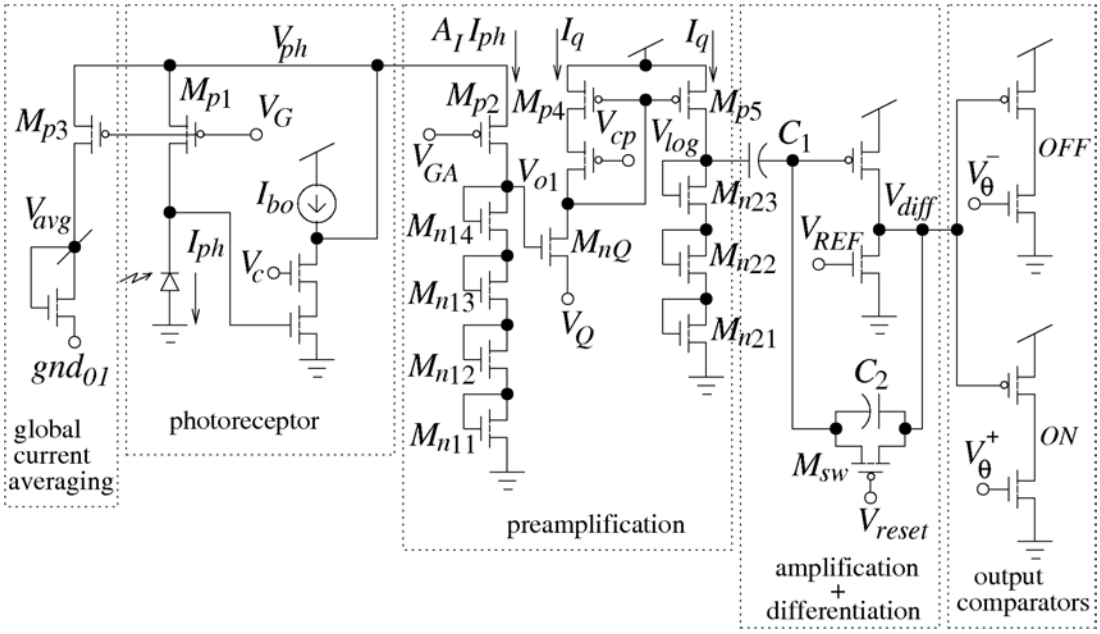

Fig. 3: Detailed pixel schematics with two diode-connected MOS chain pre-amplification stages

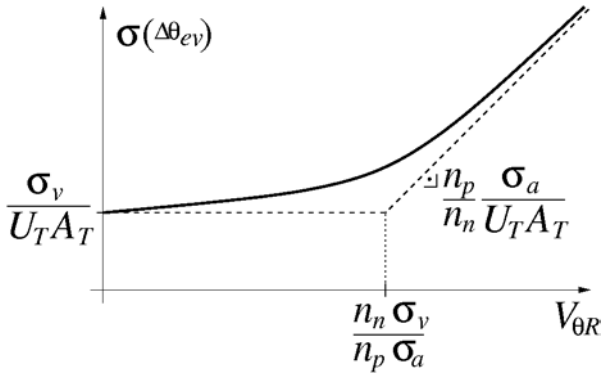

Fig. 4: Theoretical representation of contrast sensitivity mismatch versus control voltage $V_{\theta R}$.

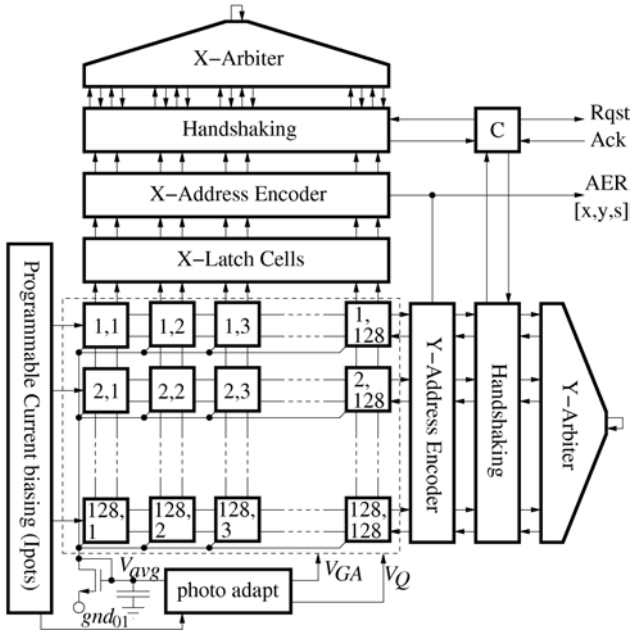

Fig. 5: Block Diagram of complete Dynamic Vision Sensor Chip

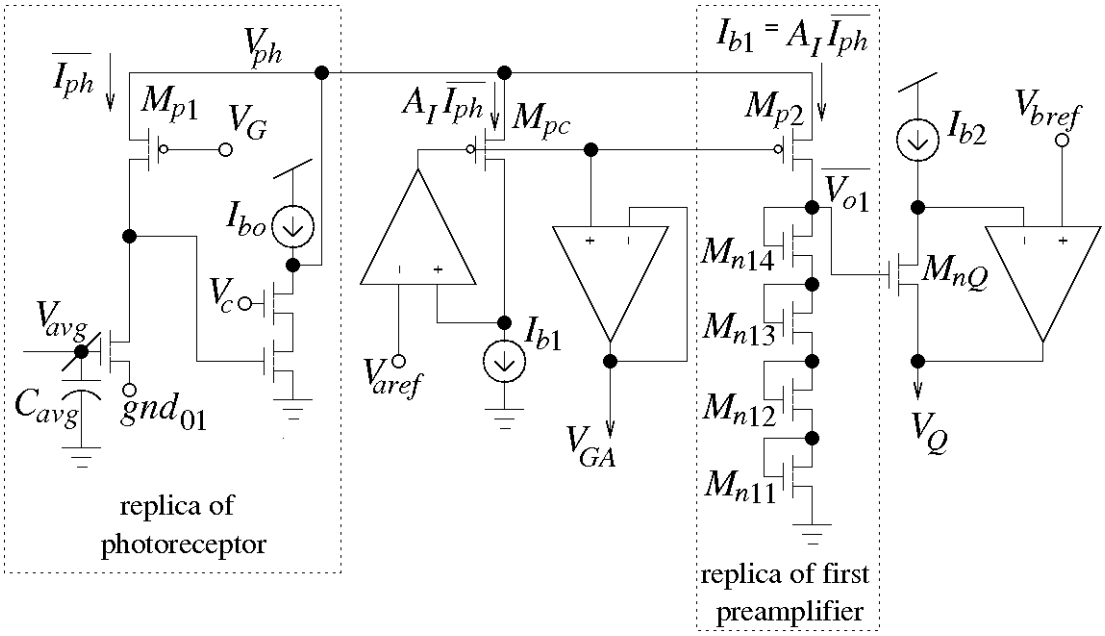

Fig. 6: Global adaptive biasing circuitry

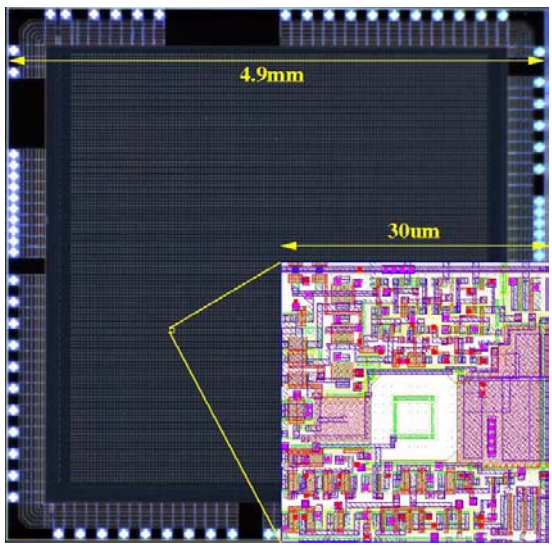

Fig. 7: Chip micro photograph and pixel layout

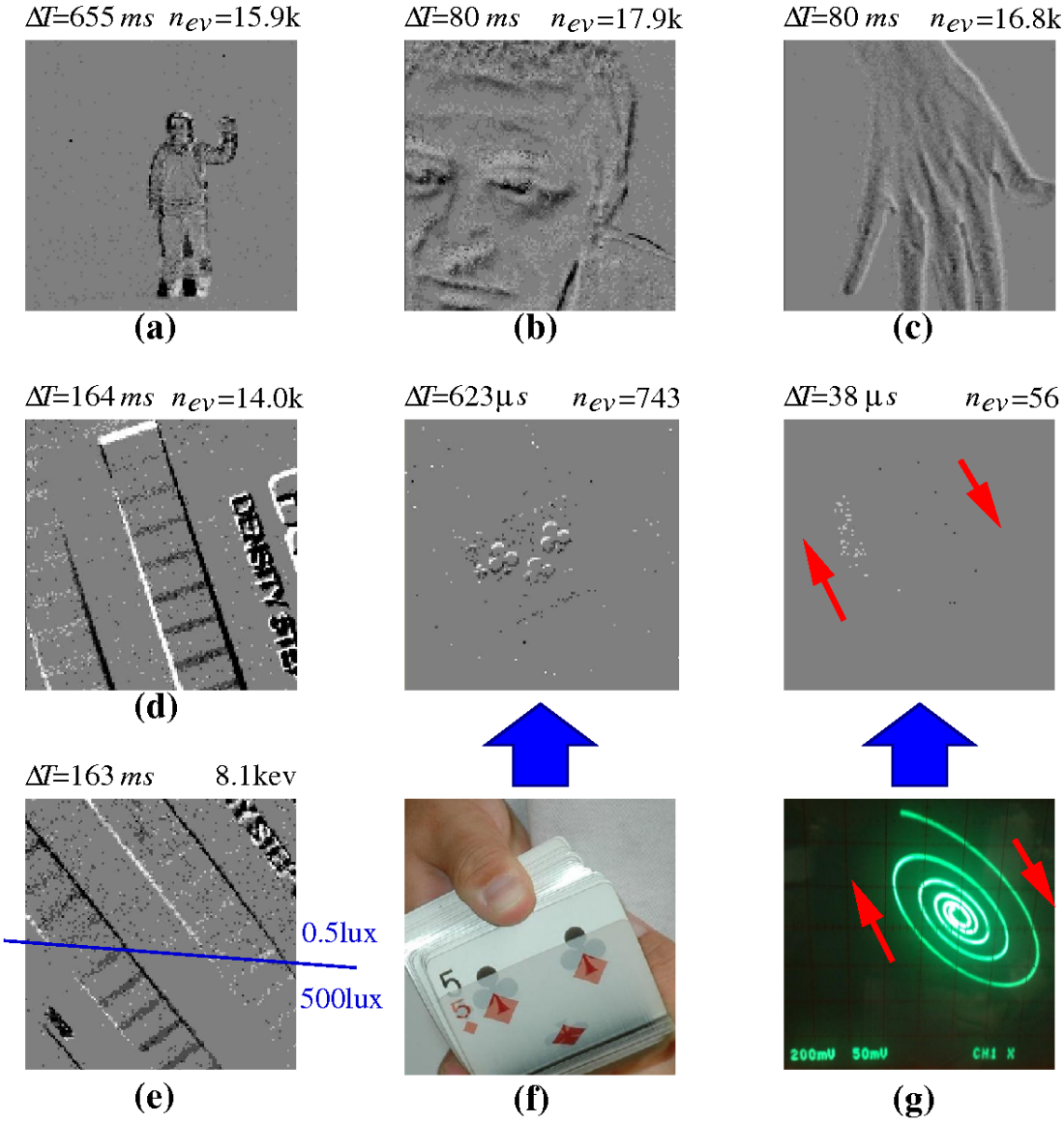

Fig. 8: Example images obtained by collecting new DVS events during a time ΔT. Background gray level represents pixels with no output events. Brighter pixels represent positive events and darker pixels negative events.

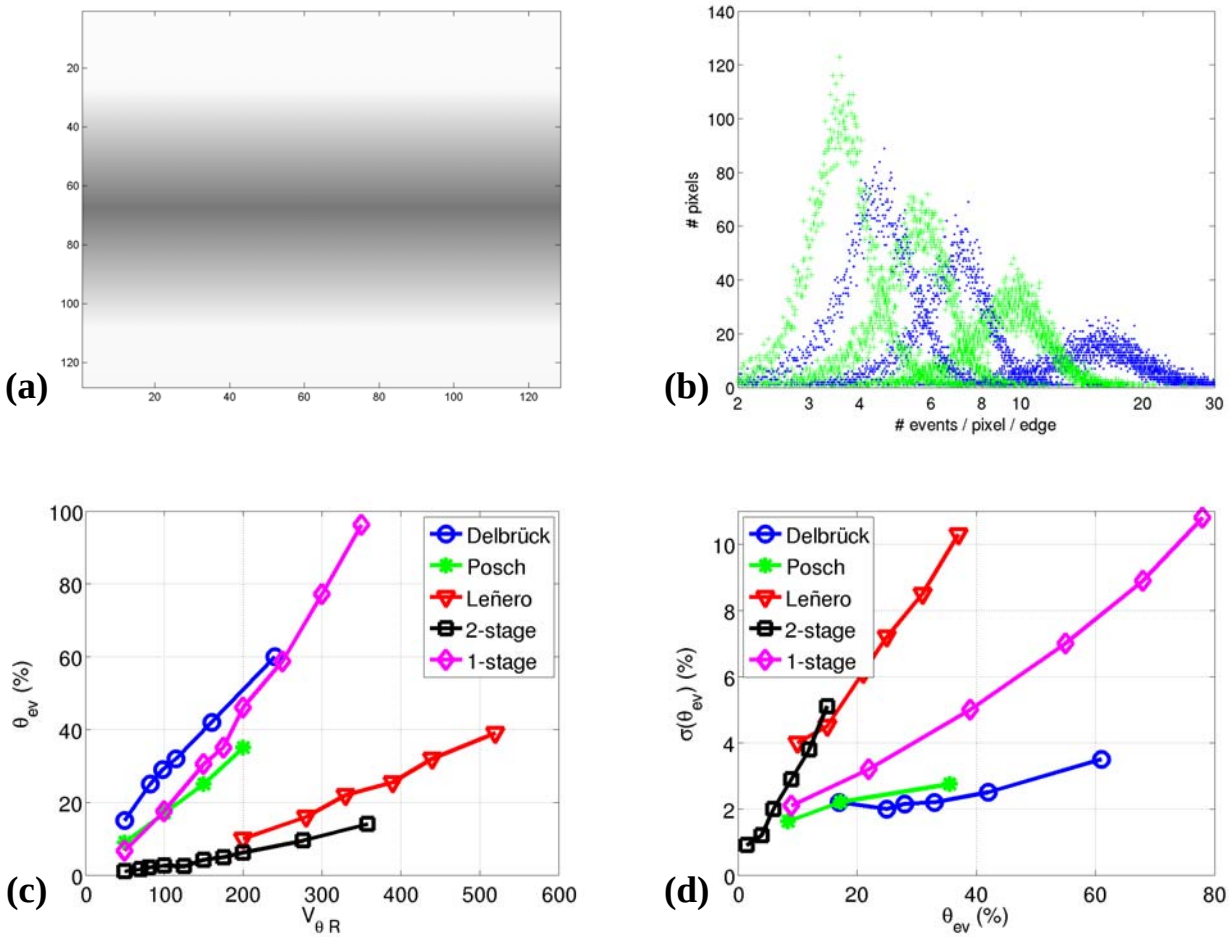

Fig. 9: Characterization of contrast sensitivity and uniformity. (a) Stimulus top-down moving bar used to produce a constant vertical contrast gradient on the pixel array. (b) Histogram of collected average number of events per pixel and edge presentation for different settings of contrast control voltage $V_{\theta R}$, when using two preamplifier stages and setting the gradient bar to a contrast of θ=0.254. (c) Contrast sensitivity θ_{ev} versus contrast control voltage $V_{\theta R}$ for the presented DVS with one and two preamplifier stages, and for other reported DVS devices. (d) Measured standard deviation of contrast sensitivity (FPN) versus contrast sensitivity θ_{ev} .

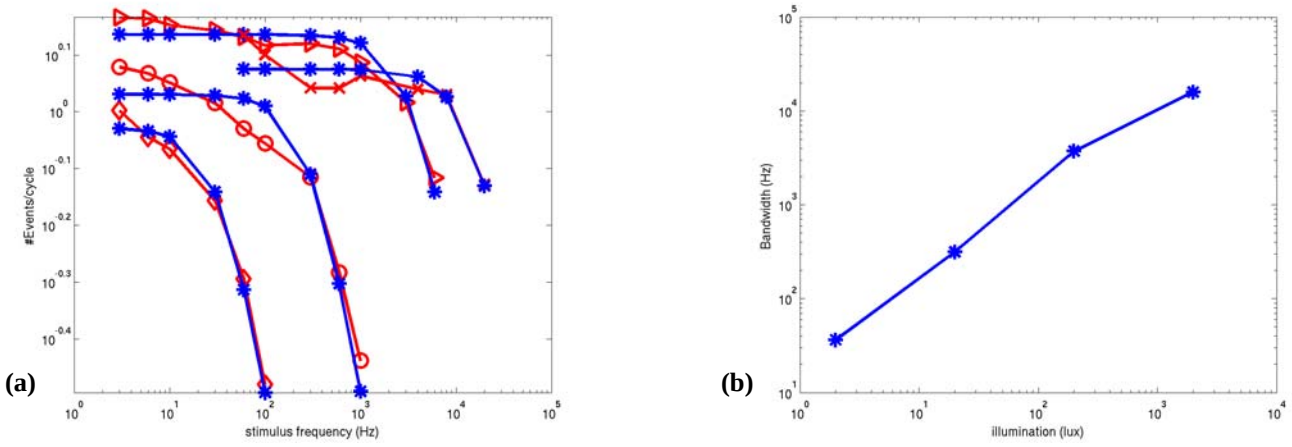

Fig. 10: (a) Diamonds (2lux), Circles (20lux), Triangles (200lux), and Crosses (2klux): experimentally measured response (in events per cycle) versus frequency of sinusoidal light stimulus, for different levels of ambient illumination; Asterixes: fit of experimental data to an ideal single pole response to extract 3dB frequency. (b) Extracted 3dB frequency versus ambient illumination.

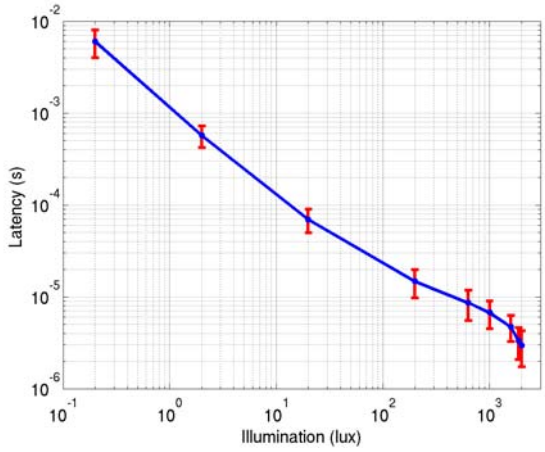

Fig. 11: Measured average pixel latency versus illumination. Inter pixel spread is indicated with vertical bars.

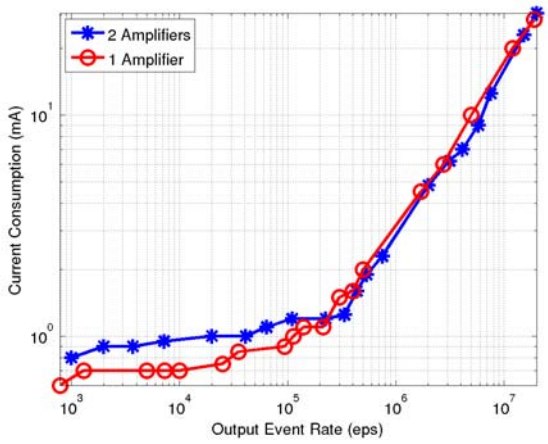

Fig. 12: Total DVS chip current consumption as a function of output event rate with both or only one pre-amplifier stages active.

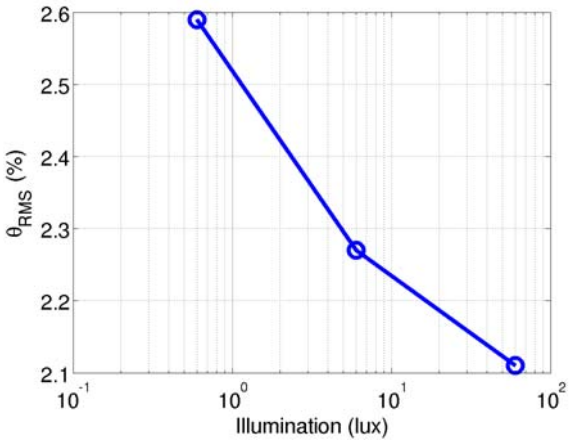

Fig. 13: Measured RMS equivalent input contrast noise versus illumination