

A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS

Christoph Posch, *Member, IEEE*, Daniel Matolin, and Rainer Wohlgenannt

Abstract—The biomimetic CMOS dynamic vision and image sensor described in this paper is based on a QVGA (304×240) array of fully autonomous pixels containing event-based change detection and pulse-width-modulation (PWM) imaging circuitry. Exposure measurements are initiated and carried out locally by the individual pixel that has detected a change of brightness in its field-of-view. Pixels do not rely on external timing signals and independently and asynchronously request access to an (asynchronous arbitrated) output channel when they have new grayscale values to communicate. Pixels that are not stimulated visually do not produce output. The visual information acquired from the scene, temporal contrast and grayscale data, are communicated in the form of asynchronous address-events (AER), with the grayscale values being encoded in inter-event intervals. The pixel-autonomous and massively parallel operation ideally results in lossless video compression through complete temporal redundancy suppression at the pixel level. Compression factors depend on scene activity and peak at ~ 1000 for static scenes. Due to the time-based encoding of the illumination information, very high dynamic range—*intra-scene DR of 143 dB static and 125 dB at 30 fps equivalent temporal resolution*—is achieved. A novel time-domain correlated double sampling (TCDS) method yields array FPN of $<0.25\%$ rms. SNR is >56 dB (9.3 bit) for >10 Lx illuminance.

Index Terms—Address-event representation (AER), biomimetics, CMOS image sensor, event-based vision, focal-plane processing, high dynamic range (HDR), neuromorphic electronics, time-domain CDS, time-domain imaging, video compression.

I. INTRODUCTION

BIOLOGICAL sensory and information processing systems appear to be much more effective in dealing with real-world tasks than their artificial counterparts. Humans still outperform the most powerful computers in routine functions involving, e.g., real-time sensory data processing, perception tasks and motion control and are, most strikingly, orders of magnitude more energy-efficient in completing these tasks. The reasons for the superior performance of biological systems are still only partly understood, but it is apparent that the hardware architecture and the style of computation in nervous systems are fundamentally different from what is state-of-the-art in artificial

synchronous information processing. It has been demonstrated [1]–[3] that modern silicon VLSI technology can be employed in the construction of *biomimetic* or *neuromorphic* artefacts that mimic biological neural functions. Neuromorphic systems, as the biological systems they model, process information using energy-efficient, asynchronous, event-driven methods.

The greatest successes of neuromorphic systems to date have been in the emulation of peripheral sensory transduction, most notably in vision. Since the seminal attempt to build a “silicon retina” by Mahowald and Mead in the late 1980s [4], a variety of biomimetic vision devices has been proposed and implemented [5]. In the field of imaging and vision, two observations are crucial: biology has no notion of a frame, and the world—the source of most visual information we are interested in—works in continuous-time and asynchronously. The authors are convinced that biomimetic asynchronous electronics and signal processing have the potential—also in fields that are historically dominated by synchronous approaches such as artificial vision, image sensing and image processing—to reach entirely new levels of performance and functionality, comparable to the ones found in biological systems. Future artificial vision systems, if they want to succeed in demanding applications such as, e.g., autonomous robot navigation, high-speed motor control, visual feedback loops, etc. must exploit the power of the asynchronous, frame-free, biomimetic approach.

Studying biological vision, it has been noted that there exist two different types of retinal ganglion cells and corresponding retina–brain pathways in, e.g., the human retina: The “Magno”-cells are at the basis of what is named the transient channel or the Magno-cellular pathway. They have short latencies and respond transiently when changes—movements, onsets, offsets—are involved. The “Parvo”-cells are at the basis of what is called the sustained channel or the Parvo-cellular pathway. Parvo-cells are mainly concentrated in the fovea, the center of the retina. They have longer latencies, respond in a sustained way, and are most probably involved in the transportation of detailed pattern, texture and color information. It appears that these two parallel pathways in the visual system are specialized for certain types of visual perception [6].

- The Magno-cellular system is more oriented toward general detection or alerting and is referred to as the “where” system.
- Once an object is detected, the detailed visual information (spatial details, color) seems to be carried primarily by the Parvo-system. It is hence called the “what” system.

Practically all conventional frame-based image sensors can functionally be attributed to the “what” system side, thus ne-

Manuscript received April 23, 2010; revised June 25, 2010; accepted September 08, 2010. Date of current version December 27, 2010. This paper was approved by Guest Editor Kofi Makinwa.

The authors are with the Department of Safety and Security, AIT Austrian Institute of Technology GmbH, 1220 Vienna, Austria (e-mail: christoph.posch@ait.ac.at).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSSC.2010.2085952

glecting the dynamic information provided by the natural scene and perceived in nature by the Magno-cellular pathway. Attempts to mimicking the Magno-cellular pathway has recently been a line of activity in neuromorphic vision and has led to the development of the Dynamic Vision Sensor (DVS) [7]–[9]. This type of visual sensor is sensitive to the dynamic information provided by the scene, however it disregards the sustained information perceived in nature by the Parvo-cellular “what” system.

A logical next step appears to be a combination of “where” and “what” system functionalities. A visual device implementing this paradigm could open up a whole new field of sensor functionality and performance, and image processing techniques. A first attempt towards this goal is presented here. ATIS, an asynchronous time-based image sensor, is the first visual sensor to combine several functionalities of the biological “where” and “what” systems and other bio-inspired approaches, namely event-based time-domain imaging, temporal contrast dynamic vision and asynchronous, event-based information encoding and data communication.

Technically the imager incorporates an array of asynchronous, fully autonomous pixels, each containing event-based change detection and pulse-width-modulation (PWM) exposure measurement circuits. The operation principle ideally results in highly efficient lossless video compression through temporal redundancy suppression at the focal-plane while the asynchronous, time-based exposure encoding yields exceptional dynamic range, SNR, fixed-pattern noise (FPN) performance and temporal resolution along with the possibility to flexibly optimize trade-offs in response to differing application demands.

This paper describes the ideas behind the sensor concept, discusses selected design and implementation issues, and summarizes some important results from chip laboratory tests and application-oriented characterization. Section II discusses relevant aspects of the state-of-the-art in solid-state imaging and reviews known approaches to overcome certain limitations and prior work related to the presented device. In Section III, the implemented sensor concept is functionally described and in Sections IV and V, details of pixel and imager design are discussed, starting from basic theoretical considerations. Section VI contains test and measurement results from the fabricated imager.

II. LIMITATIONS TO SOLID-STATE IMAGING AND RELATED WORK

Continuous advances in deep-submicron CMOS process technology allow building high-performance single-chip cameras, combining image capture and advanced on-chip processing circuitry in the focal plane. Despite all progress, several problems with solid-state imaging remain unresolved and the performance is limited, mostly due to physical constraints of fabrication technology and operating principles.

A. Temporal Redundancy

Conventional image sensors acquire the visual information time-quantized at a predetermined frame rate. Each frame

carries the information from all pixels, regardless of whether or not this information has changed since the last frame had been acquired. This approach obviously results, depending on the dynamic contents of the scene, in a more or less high degree of redundancy in the recorded image data. Acquisition and handling of these dispensable data consume valuable resources and translate into high transmission power dissipation, increased channel bandwidth requirements, memory size and post-processing power demands.

One fundamental approach to dealing with temporal redundancy in video data is frame difference encoding. This simplest form of video compression includes transmitting only pixel values that exceed a defined intensity change threshold from frame to frame after an initial key-frame. Frame differencing is naturally performed off-sensor at the first post-processing stage [10], [11], yet a number of image sensors with focal-plane frame differencing have been reported.

In an early reference from 1997, Aizawa *et al.* report pixel-parallel frame-differencing in a 32×32 prototype array [12]. A 189×182 pixel array with frame differencing is reported in [13]. The sensor outputs normal image frames and difference images for, e.g., motion detection. In [14], current mode frame differencing for activity estimation is described. The device, on detecting activity beyond a certain threshold, switches to high DR imaging mode based on time-domain PFM with residual voltage readout. A CMOS imager with pixel-level temporal change detection is described in [15]. A 90×90 pixel array performs motion/change detection by pixel-level frame differencing. The pixels store the previous frame brightness level and provide an output of the intensity change polarity.

All these frame differencing imagers still rely on acquisition and processing of full frames of image data and are not able to self-consistently suppress temporal redundancy and provide real-time compressed video output. Furthermore, even when the processing and difference quantization is done at the pixel-level, the temporal resolution of the acquisition of the scene dynamics, as in all frame-based imaging devices, is still limited to the achievable frame rate and is time-quantized to this frame rate. The main obstacle for sensor-driven video compression lies in the necessity to combine a pixel identifier and the corresponding grayscale value and implement conditional readout using the available array scanning readout techniques.

Autonomous suppression of temporal redundancy, and consequently real sensor-driven video compression, can be achieved through pixel-individual exposure on-demand, based on asynchronous, pixel-autonomous change detection. The problem of efficient, combined transmission of pixel addresses and intensity values can be resolved by using time-based PWM exposure measurement and asynchronous, event-based AER information encoding and data communication [5], [16], [17]. This approach, in addition, avoids unnatural time-quantization in all stages of image data acquisition and early processing. The device described in this paper implements this approach and achieves highly efficient sensor-driven ideally lossless video compression, delivering high-quality streaming video with compression factors depending essentially only on scene activity.

B. Dynamic Range

The dynamic range (DR) of an image sensor is defined as the ratio of the maximum processible signal and the noise floor under dark conditions. Conventional CMOS active pixel sensors (APS) are based on some variation on the 3T or 4T voltage-mode pixel. In the standard APS scheme, the exposure time and the integration capacitance are held constant for the pixel array. For any fixed integration time, the analog readout value has a limited signal swing that determines the maximum achievable DR as

$$\text{DR} = 10 \log \left(\frac{V_{\text{sat}}^2}{V_{\text{dark}}^2 + V_{\text{reset}}^2 + V_{\text{out}}^2} \right) \quad (1)$$

where V_{sat} is the maximum allowed voltage at the integration node and V_{dark} , V_{reset} , and V_{out} are darkcurrent, reset (kTC) and readout noise voltages, respectively. Most voltage (and current) mode image sensors exhibit a saturating linear response with a DR limited to 60–70 dB. Both the signal saturation level and the noise floor are eventually constrained by the fabrication process.

Light from natural scenes can span up to 140 dB of DR, ranging from 1 mlx up to 10 klx and more. According to notable experts in the field, it is clear that high dynamic range imaging will dominate the market in the near future [18].

C. Signal-to-Noise Ratio

The signal-to-noise ratio (SNR), as an important criterion for image quality, is defined as the quotient of the signal power and the average noise power:

$$\text{SNR} = 10 \log \left(\frac{V_{\text{sig}}^2}{V_{\text{dark}}^2 + V_{\text{photo}}^2 + V_{\text{reset}}^2 + V_{\text{out}}^2} \right) \quad (2)$$

with V_{photo} representing the photocurrent shot noise. Since the photocurrent shot noise is the dominant noise source for moderate and high light illumination conditions, (2) can be approximated as

$$\text{SNR} \approx 10 \log \left(\frac{C_D \cdot V_{\text{sig}}}{q} \right) \quad (3)$$

where C_D is the photodiode integration capacitance and q the elementary charge. Because the SNR is proportional to the integration voltage V_{sig} , in conventional APS image sensors with a fixed integration time for all pixels, the image quality strongly depends on the illuminance, even for moderate or high light illumination conditions (compare Fig. 2).

D. Time-Domain Imaging

To overcome the standard image sensor's DR and SNR limitations, several approaches have incorporated the dimension of time, in one form or another, as a system variable. While some designs use either variable integration times or time-dependent well capacities to increase dynamic range [19]–[21], other designs are based on directly measuring the time it takes the photocurrent to produce a given voltage change at the sense node. This technique is commonly called time-domain or pulse modulation (PM) imaging.

In PM imaging the incident light intensity is not encoded in amounts of charge, voltage, or current but in the timing

of pulses or pulse edges. Dual to the voltage-mode pixel, the “integration-time”-mode pixel connects the sense node to a comparator which toggles state when V_{int} goes beyond some reference value V_{ref} . The state is reflected in the binary signal V_{out} , which may be connected to an output bus and/or fed back to the reset transistor. If an external signal V_{reset} is used to reset the sense node, the pixel operates as a timer. If the loop is closed to connect the comparator output to the reset transistor, the pixel becomes an oscillator which generates pulses on the V_{out} node at a frequency related to the instantaneous photocurrent (integration rate) and hence to the pixel illumination. PM imaging can thus be coarsely classified into two basic techniques, namely pulsewidth modulation (PWM) encoding, and pulse frequency modulation (PFM) encoding (Fig. 1). References [22], [23], and [24] dating from 1996 onwards report early PWM image sensor implementations. The first PFM circuit was reported by Frohmader *et al.* [25] in 1982. The first PFM-based image sensor was proposed in 1993 [26] and demonstrated in 1994 [27].

Both schemes allow each pixel to autonomously choose its own integration time. By shifting performance constraints from the voltage domain into the time domain, DR is no longer limited by the power supply rails.

DR in PWM exposure encoding is given by the simple relation

$$\text{DR} = 20 \log \frac{I_{\text{max}}}{I_{\text{min}}} = 20 \log \frac{t_{\text{int,max}}}{t_{\text{int,min}}} \quad (4)$$

where the maximum integration time $t_{\text{int,max}}$ is limited by the darkcurrent (typically seconds) and the shortest integration time by the maximum achievable photocurrent and the sense node capacitance (typically microseconds). DR values of the order of 100–120 dB have been reported for various PWM imagers [28], [29].

Also the sensor's SNR benefits from the time-domain approach. In time-based image sensors, every pixel reaches the maximum integration voltage V_{sig} in every integration cycle. Consequently the achievable SNR is essentially independent of illuminance and photocurrent [compare (3)]. Fig. 2 plots SNR for a voltage-mode APS and a time-domain pixel as a function of illuminance. The strong light dependency of APS SNR is apparent while the time-based pixel reaches full SNR already at low light conditions. The main boundary conditions to these data are: readout noise $V_{\text{out}} = 100 \mu\text{V}$ rms, photodiode capacitance $C_D = 30$ fF, photodiode area $A_D = 150 \mu\text{m}^2$, darkcurrent $I_{\text{dark}} = 1.5$ nA/cm², photocurrent $I_{\text{ph}} = 0.1$ pA/lx, integration swing (PM) $V_{\text{int}} = 2$ V.

III. ATIS IMAGER CONCEPT

As touched on above, the adverse effects of data redundancy, common to all frame-based image acquisition techniques, can be tackled in different ways. The biggest conceivable gain however is achieved by simply not recording the redundant data in the first place and directly reducing data volume at the sensor output. The immediate benefits are reductions in bandwidth, memory and computing power requirements for data transmission and post-processing, hence decreasing system power, complexity and cost.

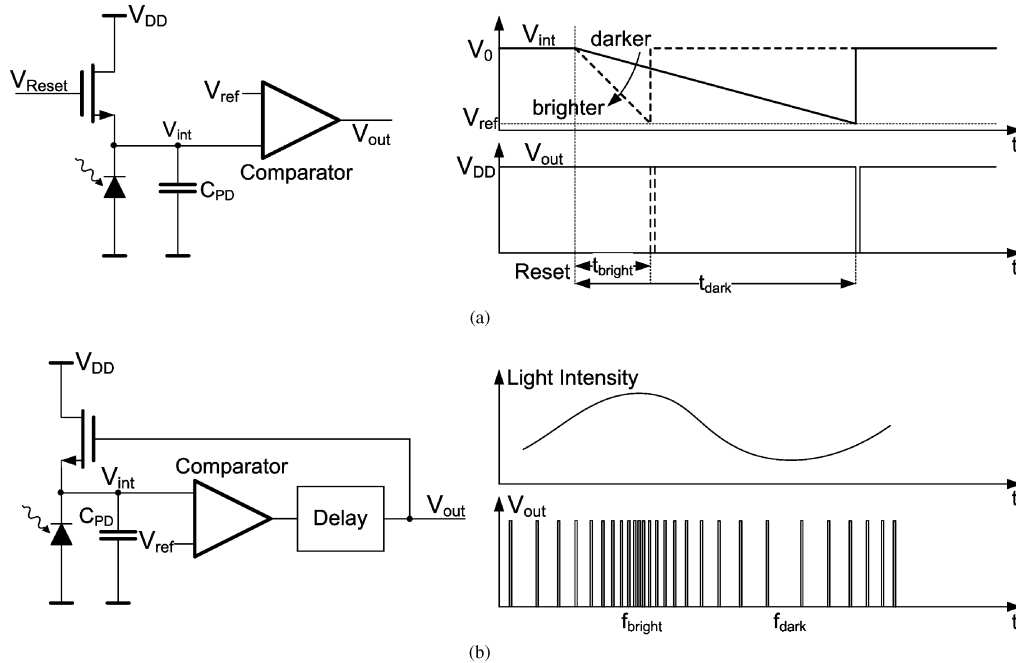


Fig. 1. (a) PWM and (b) PFM encoding of the exposure information.

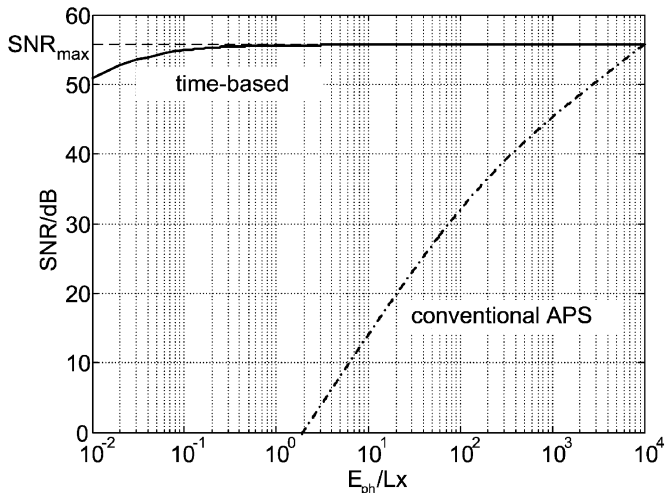


Fig. 2. SNR of a standard voltage mode APS and a time-based pixel as a function of illuminance.

A fundamental solution for achieving complete temporal redundancy suppression is using an array of fully autonomous pixels that combine a *change detector* and a *conditional exposure measurement* device. The change detector individually and asynchronously initiates the measurement of a *new* exposure/grayscale value only if—and immediately after—a brightness change of a certain magnitude has been detected in the field-of-view of the respective pixel. Such a pixel does not rely on external timing signals and independently requests access to an (asynchronous and arbitrated) output channel only when it has a new grayscale value to communicate. Consequently, a pixel that is not stimulated visually does not produce output. In addition, the asynchronous operation avoids the time quantization of frame-based acquisition and scanning readout.

Fig. 3 shows simplified schematics of the implemented pixel and typical signal waveforms. The pixel is composed of two main blocks, change detector (CD) and exposure measurement (EM) block [Fig. 3(a)].

A. Change Detection

As the change detector a fast continuous-time logarithmic photoreceptor with asynchronous event-driven signal processing is used [Fig. 3(c)]. The circuit has originally been developed for temporal contrast dynamic vision sensors [7]–[9] and combines an active, continuous-time, logarithmic photo-front-end ($PD1, M1, A1$) with a well-matched, self-timed, self-balancing switched-capacitor amplifier ($C1, C2, A2$). It continuously monitors photocurrent I_{ph} for changes and responds with an ON or OFF event that represents a fractional increase or decrease in intensity that exceeds tunable thresholds. The occurrence of these events is sensed by one of two voltage comparators. The circuit responds to relative temporal contrasts of a few percent over 6 decades of illumination. Fig. 3(d) shows typical signal waveforms of the change detector circuit. The upper trace represents an arbitrary voltage waveform at the node V_p tracking the photocurrent through $PD1$. The event generation circuitry responds with pulse events of different polarity to positive and negative gradients of the photocurrent; the rate of change is encoded in the inter-event intervals. Each of these change events is used to trigger a reset signal to the exposure measurement part. The polarity information is not required for the conditional exposure measurement functionality but is useful in various machine vision applications that rely on high temporal resolution event-based change information [30], [31].

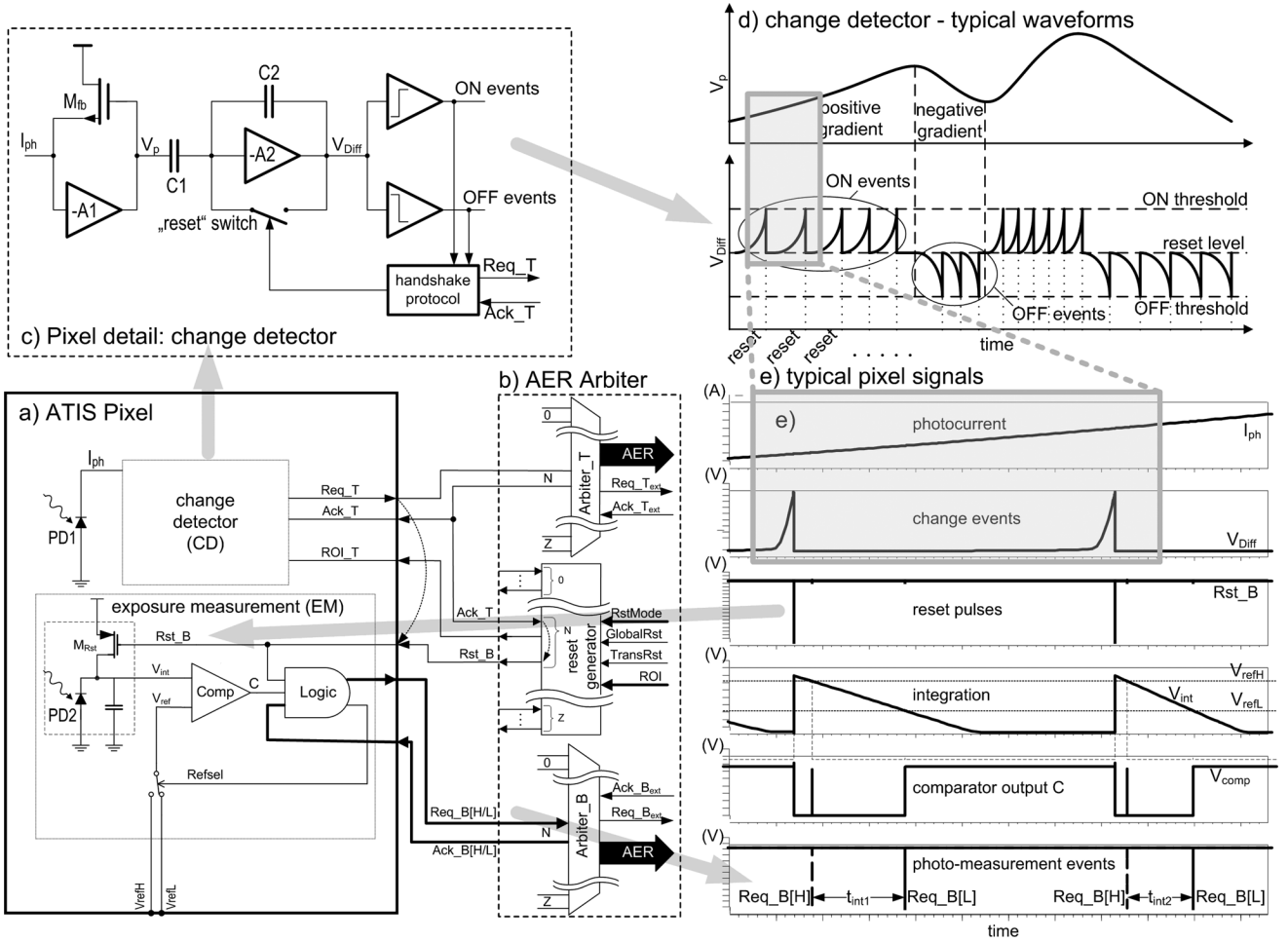


Fig. 3. (a) ATIS pixel connected to (b) arbiters and reset generator/mode-of-operation/ROI controller, (c) change detector schematic, (d) typical change detector signals, and (e) exemplary pixel waveforms illustrating two change event triggered exposure measurements cycles (taken from SPICE simulation).

B. Exposure Measurement

The exposure-measurement device is realized as a time-based PWM circuit (compare Section II-D). Additional state and control logic in the pixel allows the implementation of time-domain true correlated double sampling (TCDS) [32] based on two global integration thresholds (V_{refH}/V_{refL}). The true differential operation within one integration cycle eliminates both comparator offset FPN and reset kTC noise [Fig. 3(a)].

In the following, one cycle of transient change detection and exposure measurement is explained and illustrated by typical signal waveforms, taken from transistor-level simulation results [Fig. 3(e)] [33]. The change detector responds to a relative change in illumination by triggering the transmission of an address-event via **Arbiter_T**, and simultaneously delivers a pulse on the reset line Rst_B (via row and column reset/mode-control circuits), which initiates an exposure-measurement cycle. The Rst_B signal briefly closes the switch M_{Rst} , connecting the sense node to V_{DD} . The pixel state control logic ensures that, at this point, the higher threshold voltage V_{refH} is connected as the reference voltage by setting the $RefSel$ signal accordingly. By releasing the Rst_B signal, the integration process starts and the voltage V_{int} on the photodiode decreases proportionally to the photocurrent and thus proportionally to the illumination

at the photodiode. When the photodiode voltage V_{int} reaches V_{ref} , the comparator output C toggles, causing the state logic to trigger the transmission of an address-event via **Arbiter_B** by activating the $Req_B[H]$ signal, and to toggle the $RefSel$ signal—now V_{refL} is set as the voltage reference and the comparator output C toggles back. The integration continues in the meantime. V_{int} reaching V_{refL} marks the end of the measurement cycle, C toggles again and the logic releases another address-event by sending a $Req_B[L]$ signal to **Arbiter_B**. The time between the two address events, triggered by $Req_B[H]$ and $Req_B[L]$, is inversely proportional to the average pixel illumination during the integration. Transient sensor (change detector) and exposure-measurement operation, once started, are completely detached and do not influence each other (in particular do not share a common output channel), with one important exception: If the change detector senses another change before the integration process has finished, the current measurement cycle is aborted and the integration is restarted. In this case, the $Req_B[H]$ address-event is discarded by the post processor (which it detects in receiving two consecutive $Req_B[H]$ events from the same pixel address). This behavior is intentional and does not imply information loss (depending on the observation time-scale), because, as a further change in illumination had taken place, the initial exposure measurement

result would have been already obsolete. This conduct ensures that each transmitted exposure measurement result is as accurate and recent as possible.

It also implies that if the scene (or a part of it) constantly changes at a rate that produces CD events with a frequency higher than the inverse integration time (hundreds of microseconds to tens of milliseconds for typical illumination conditions, compare Section VI-B1), the involved pixels, for the time of sustained change/oscillation, will potentially never finish integration and deliver a new grayscale value. A conventional frame-based image sensor, however, would neither be able to deliver useful information under these circumstances. If desired ATIS can, at any time, be set to frame mode and acquire a snapshot of the scene independently of CD operation (see Section IV).

C. System Considerations

The asynchronous change detector and the time-based exposure measurement approach harmonize remarkably well, mainly for two reasons: On the one hand because both reach a dynamic range of >120 dB—the first is able to detect relative changes in illumination of a few percent over the full range, the latter is able to resolve the associated grayscales independently of the initial light intensity. On the other hand because both circuits operate event-based, namely the events of detecting illumination or reflectance changes and the events of pixel integration voltages reaching reference thresholds. Consequently an asynchronous, event-based communication scheme (Address Event Representation, AER [5], [16], [17]) is used in order to provide efficient allocation of the transmission channel bandwidth to the active pixels. Along with the pixel array address, the relevant information is inherently encoded in the event timing. Time-to-digital conversion of the event timings and the calculation of grayscale values from integration times are done off-chip (see Section V-C for details).

The ATIS dynamic vision and image sensor is built around a QVGA (304×240) pixel array and uses separate bus arbiters and event-parallel AER channels for communicating change events and grayscale encoding events independently and in parallel [Fig. 3(b)]. Furthermore, the sensor features a flexible column/line-wise reset/trigger scheme for various modes-of-operation. Besides the (default) self-triggered mode, there are, e.g., external trigger modes for “snapshot” frame acquisition with “time-to-first-spike” (TTFS) encoding [34], or column-parallel relay readout [35]. Change detector and externally triggered imager operation can be fully decoupled and used independently; programmable regions-of-(non)-interest (ROI/RONI) are available separately for both parts.

IV. PIXEL DESIGN

A. Change Detector

The change detector circuit has been adapted from, and is functionally equivalent to the one described in detail in [8]. The goal for the design of this circuit was to achieve temporal contrast sensitivity of a few percent with low mismatch, wide dynamic range and low latency [33]. The circuit consists of a photoreceptor front-end, a differencing switched-capacitor amplifier and a comparator-based event generator (Fig. 4).

1) *Logarithmic Photoreceptor*: The photoreceptor responds logarithmically to intensity, thus implementing a gain control mechanism that is sensitive to temporal contrast or *relative* change. The circuit comprises a photodiode whose photocurrent is sourced by a saturated nMOS transistor M_{fb} . The gate of M_{fb} is connected to the output of an inverting amplifier (M_{pr} , M_{cas} , M_n) whose input is connected to the photodiode. This transimpedance configuration converts the photocurrent logarithmically into a voltage and also holds the photodiode clamped at virtual ground. As a result, the bandwidth of the photoreceptor is extended by the factor of the loop gain in comparison to a simple passive logarithmic photoreceptor. At low-light conditions the bandwidth of photoreceptor is limited by the photocurrent and can be approximated by a first-order low-pass filter with corner frequency

$$f_{3dB} = \frac{1}{2\pi C_{DMfb}} \frac{I_{ph}}{U_T} \quad (5)$$

where C_{DMfb} is the gate-drain capacitance of transistor M_{fb} and U_T is the thermal voltage.

The bandwidth increase of the feedback configuration effects a corresponding reduction in SNR, which is given by

$$\text{SNR} \approx 10 \log \left(\frac{C_{DMfb} \cdot U_T}{\kappa_{Mfb} \cdot q} \right) \quad (6)$$

where κ_{Mfb} is the subthreshold slope factor of transistor M_{fb} , [37]. SNR levels of about 30 dB can be reached with this configuration.

The DR of the continuous-time logarithmic photoreceptor is given by the same expression as the one of time-based image sensors (4), where I_{max} is the photocurrent at maximum illumination and I_{min} is the darkcurrent. Assuming equal darkcurrent densities in both photodiodes, both pixel circuits, CD and EM, exhibit a very similar DR.

2) *Event Generator*: The photoreceptor output is buffered by a source follower and then differentiated by capacitive coupling to a floating node at the input of a common-source amplifier stage with switched capacitor feedback. The source follower isolates the sensitive photoreceptor from the rapid transients in the differencing amplifier. The amplifier is balanced using a reset switch that shorts input and output, yielding a reset voltage level depending on the amplifier operating point. Transients sensed by the photoreceptor circuit appear as an amplified deviation from this reset voltage at the output of the inverting amplifier. The closed loop differencing amplifier gain is determined by the capacitor ratio C_1/C_2 .

The comparators (M_{ONn} , M_{ONp} , M_{OFFn} , M_{OFFp}) compare the output of the inverting amplifier V_{diff} against global thresholds. The two thresholds, set by $V_{b,ON}$ and $V_{b,OFF}$, are offset from the reset voltage in both directions to detect increasing and decreasing changes. A change in V_{diff} that triggers one of the comparators leads to a corresponding “event” and a reset of the differencing amplifier. The events are communicated by implementing a 4-phase AE handshaking with the peripheral AE circuits [8]. Transistor M_s implements the externally controlled ROI/RONI functionality by conditionally applying a permanent reset to the circuit.

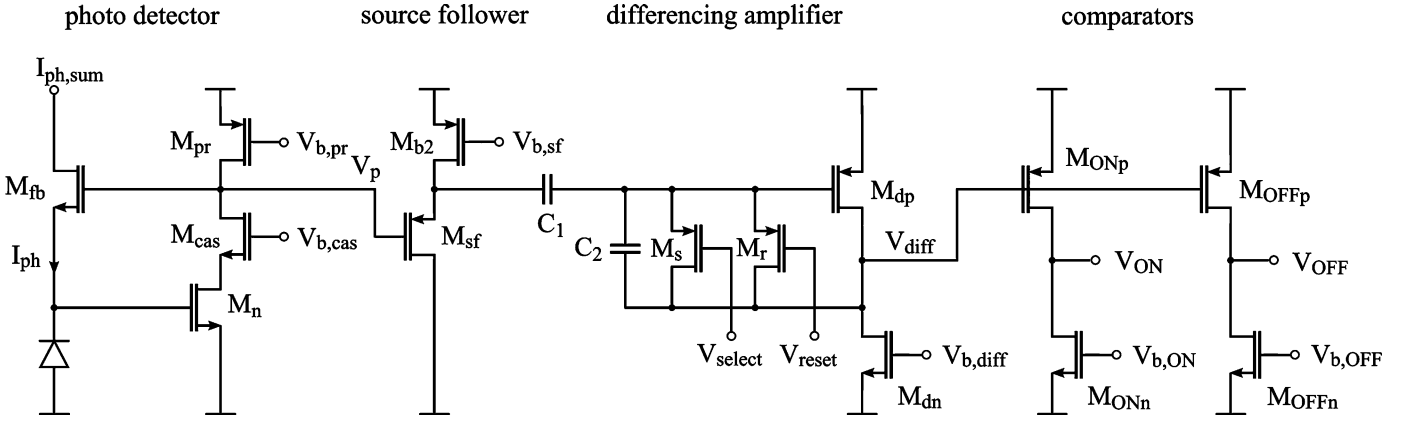


Fig. 4. Transistor-level change detector schematic corresponding to the abstract schematic in Fig. 3(c). (Circuitry for AE handshake is omitted).

B. PWM Exposure Measurement

The time-domain approach to exposure measurement has been chosen for reasons of DR and SNR performance, as discussed in Section II-D, and its affinity to event-based information encoding and data communication as outlined in Section III-C. In the following, the main design considerations concerning the time-based exposure measurement circuit are summarized.

1) *Circuit Basics*: For the PWM circuit, an n-well/p-sub photodiode with pMOS reset transistor is used [Fig. 3(a)]. The standard CMOS mixed-mode/RF fabrication process along with relatively relaxed area restrictions allow realizing the reset transistor M_{Rst} as p-type, thus maximizing integration swing. The sense node is directly coupled to the voltage comparator input. An analog switch permits to choose between two externally applied threshold voltages at the comparator's reference input terminal for TCDS operation as described in Section IV-B2). A logic block stores the instantaneous pixel state and controls the reference switch accordingly (refer to Section IV-B5). Furthermore, it is responsible for the event-based communication with the AER arbiter.

2) *Time-Domain Correlated Double Sampling (TCDS)*: For low light illumination conditions, the SNR is determined by the darkcurrent shot noise and the kTC noise. The darkcurrent shot noise strongly depends on the fabrication technology while the kTC noise can be effectively reduced using double sampling techniques. Correlated double sampling (CDS) of the pixel voltage, as a widely-used method for conventional APS imagers, eliminates the reset noise by sampling the pixel voltage twice, once immediately after the photodiode reset and once after the integration is finished. Subtracting the voltages cancels out the reset noise and DC noise components such as fixed-pattern noise (FPN) [38]. In order to suppress reset noise and FPN in time-based image sensors, dual to voltage-mode CDS, a time-domain differential method can be inferred based on a dual-threshold arrangement.

In [39] a time-based image sensor using a differential technique is presented. To determine the integration time within one integration cycle, two comparators connected to different reference voltages are used. This approach suppresses reset noise, however FPN noise power due to comparator offsets is doubled. The authors of [40] show a differential approach using only one

comparator, but two integration cycles with different reference voltages to determine the time difference for the voltage drop. This method corresponds to non-true correlated double sampling in the voltage domain and provides a suppression of FPN, but duplicates the effect of kTC noise.

The differential TCDS approach implemented in the ATIS pixel is based on a comparator and pixel-level state and control logic that conditionally applies different threshold voltages V_{refH} and V_{refL} within one integration cycle. Consequently, this method eliminates both kTC reset noise and comparator offset FPN. The noise reduction performance of the proposed TCDS scheme has been derived theoretically and compared to measurement results from the fabricated chip in [32]. A brief summary is given here.

With the integration time t_{int} for the differential approach

$$t_{int} = \int_{V_{refH}}^{V_{refL}} \frac{C_D(V_{pix})}{I_{ph}(V_{pix})} dV_{pix} \quad (7)$$

and a first order approximation of the error ε_{tint} in time measurement due to comparator offset V_{off}

$$\varepsilon_{tint} = \left(\frac{dt_{int}}{dV_{refL}} - \frac{dt_{int}}{dV_{refH}} \right) V_{off} \quad (8)$$

the relative error $\varepsilon_{tint}/t_{int}$ for the TCDS case can be expressed as

$$\frac{\varepsilon_{tintCDS}}{t_{int}} \approx 2 \frac{C_D(V_{refL}) - C_D(V_{refH})}{C_D(V_{refL}) + C_D(V_{refH})} \cdot \frac{V_{off}}{V_{refH} - V_{refL}} \quad (9)$$

and without TCDS using only one threshold:

$$\frac{\varepsilon_{tint}}{t_{int}} \approx \frac{2 \cdot C_D(V_{ref})}{C_D(V_{DD}) + C_D(V_{ref})} \cdot \frac{V_{off}}{V_{DD} - V_{ref}} \quad (10)$$

In Fig. 5, (9) and (10) are plotted as a function of threshold voltages, with both curves normalized to V_{off} (solid lines). The circular markers are results from measurements on the fabricated chip. Integration times were measured with and without an externally applied voltage offset of 20 mV at the reference input of the comparator and for both techniques. To achieve results independent of the exact offset voltage, the relative error numbers were normalized accordingly.

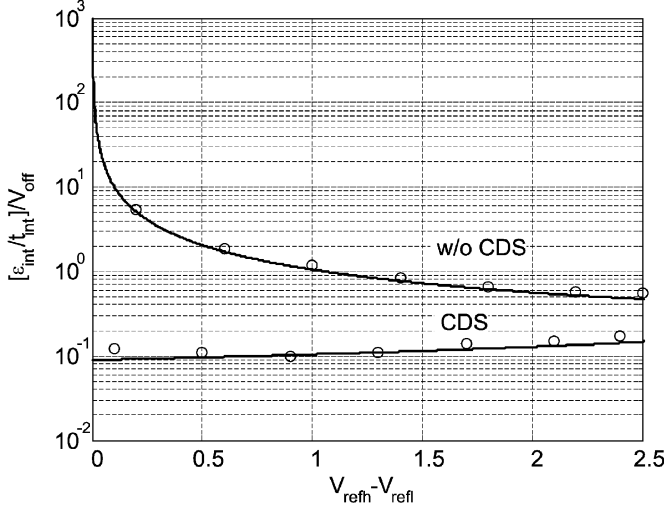


Fig. 5. Relative error $\varepsilon_{t_{\text{int}}}/t_{\text{int}}$ for time measurement with and without CDS. The error is normalized to V_{off} .

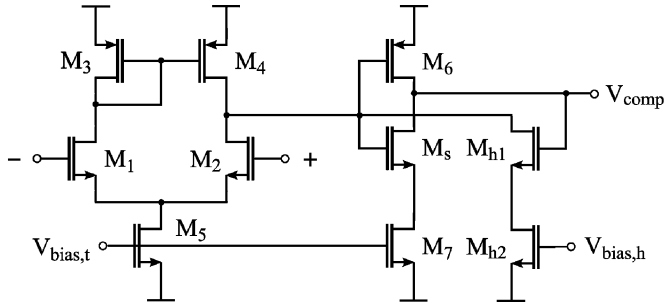


Fig. 6. Two-stage comparator circuit with hysteresis and dynamic power control.

The correspondence between calculation and measurement is very good. It can be concluded that the proposed technique reduces the relative error in the time measurement by a factor of about 5 to 20 for voltage swings $V_{\text{refH}} - V_{\text{refL}}$ between 0.5 and 2.5 volts. The effect is more manifest for lower swings making this method increasingly attractive when progressing towards modern CMOS processes with lower V_{DD} .

3) *Comparator*: Fig. 6 shows the voltage comparator circuit. The comparator is based on a standard two-stage operational amplifier consisting of transistors M_1 to M_7 plus three additional transistors M_8 , M_{h1} , and M_{h2} whose functions are discussed below and in the following sub-section [32].

Usually comparator design is a trade-off between comparator's speed and gain and its noise immunity. To render the comparator switching operation insensitive to input signal noise while achieving high switching speed and gain, an adjustable hysteresis was added to the circuit. The hysteresis is realized using only two additional transistors M_{h1} and M_{h2} where M_{h1} is a switch and M_{h2} operates as a current source.

When the input voltage V_{int} (gate M_1) passes the threshold V_{ref} (gate M_2), the output of the comparator changes, transistor M_{h1} is turned on, the current I_{h2} is subtracted from the drain node of M_4 , and the threshold point is subsequently increased. So the input voltage must return beyond the previous threshold plus the voltage V_{hyst} before the comparator's output switches

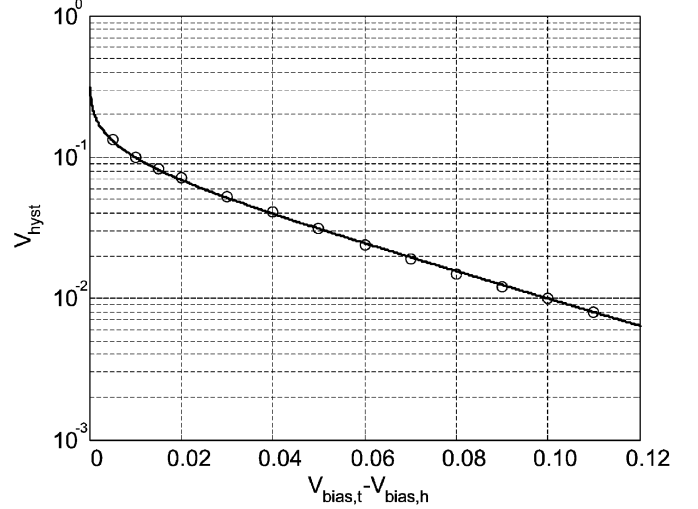


Fig. 7. Hysteresis as a function of voltage difference $V_{\text{bias,t}} - V_{\text{bias,h}}$.

again. For the transistors operating in weak inversion, the hysteresis voltage can be calculated as

$$V_{\text{hyst}} = \frac{1}{\kappa_1} U_T \ln \left(\frac{\frac{I_5}{I_{h2}} + 1}{\frac{I_5}{I_{h2}} - 1} \right) \quad (11)$$

where κ_1 is the subthreshold slope factor of M_1 and U_T is the thermal voltage [32]. The hysteresis is set by the ratio of the tail current I_5 and the hysteresis current I_{h2} flowing through transistor M_{h2} . Due to variable bias voltages at transistors M_5 and M_{h2} , V_{hyst} can be adjusted over a wide range. With same transistor dimensions for M_5 and M_{h2} , V_{hyst} is given by

$$V_{\text{hyst}} = \frac{1}{\kappa_1} U_T \ln \left(\frac{1 + e^{-\kappa_5 (V_{\text{bias,t}} - V_{\text{bias,h}})/U_T}}{1 - e^{-\kappa_5 (V_{\text{bias,t}} - V_{\text{bias,h}})/U_T}} \right). \quad (12)$$

Using a first order Taylor series approximation and assuming $\kappa_5 (V_{\text{bias,t}} - V_{\text{bias,h}})/U_T \gg 0$, (12) can be simplified to

$$V_{\text{hyst}} = \frac{2}{\kappa_1} U_T e^{-\kappa_5 (V_{\text{bias,t}} - V_{\text{bias,h}})/U_T}. \quad (13)$$

In Fig. 7, (12) is plotted for $1/\kappa_1 = 1/\kappa_5 = 1.77$ and $U_T = 25.6$ mV as a function of voltage difference $V_{\text{bias,t}} - V_{\text{bias,h}}$. The circular markers are results from measurements on the fabricated chip with a tail current $I_5 = 100$ nA. The correspondence between calculation and measurement is very good. The proposed solution offers an area and power efficient hysteresis implementation.

4) *Power Consumption*: The static power consumption of an image sensor with pixel-level signal processing is mainly determined by the power consumption of the pixel-level comparator. To significantly decrease overall power consumption of the comparator, the power consumption during both exposure measurement and idle state phases has to be reduced.

During exposure measurement ($V_{\text{int}} > V_{\text{refL}}$) the comparator's current consumption is dominated by I_5 , as turned-off transistor M_6 inhibits current flow in the output stage. For the basic technique of time measurement (one reference voltage) the minimal value of I_5 is determined by the acceptable switching

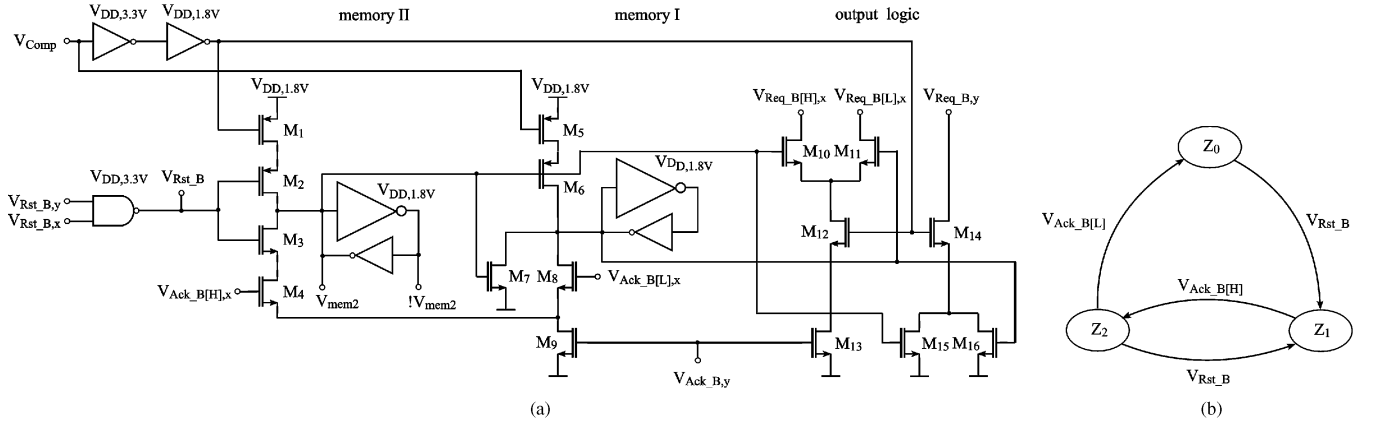


Fig. 8. (a) Schematic of the pixel-level logic circuit and (b) state diagram.

delay at the reference voltage. This acceptable switching delay depends on the desired precision in time measurement. A higher slope of V_{int} requires a higher value of I_5 . Besides offset and noise suppression, the application of the TCDS technique also yields a significant reduction in power consumption. Because the delay time of the comparator approximately cancels out for both thresholds, the tail current I_5 can be noticeably reduced, especially for fast integration. Experimental results show a tail current of only 50 nA is sufficient to realize 8-bit resolution with an accuracy of 0.5LSB even for fast integration slopes of 100 mV/s. To achieve the same precision of time measurement without TCDS would require a current I_5 more than 100 times higher. For integration slopes slower than 1 V/ms, currents down to 10 nA are sufficient using the dual-threshold approach.

With the voltage V_{int} passing the lower reference V_{refL} , the pixel changes to idle state. In this state, the pixel is waiting for a new reset, which, for the change event-triggered ATIS operation mode, can take an arbitrarily long time. Thus, it is highly desirable that also in this state, the comparator power consumption is minimized. With the pixel entering idle state, the continuing photocurrent integration leads to a further reduction of V_{int} and eventually tail current I_5 is cut off completely. To effectively minimize idle state power consumption, the current flow in the output stage also has to be switched off. Because of the pixel's autonomous operation, a global current control—like, e.g., in [43]—cannot be used. In order to switch off the current in the output stage individually, transistor M_s has been placed between the output node and the current sink transistor M_7 . M_s switches off I_7 when the pixel enters the idle state. As a consequence of the described methods, the current flow in the comparator in normal operation can be chosen to be of the order of 50 nA, even for fast integration slopes, and is completely turned off in idle state.

5) *Digital Logic Circuit*: The in-pixel state and control logic circuitry essentially consists of three parts: Asynchronous digital output circuits for event communication to the bus arbiters, two 1-bit memory elements for storing the current pixel state and for controlling pixel and TCDS operation, and level adapters for interfacing the 3.3 V analog and 1.8 V digital supply voltage domains. Fig. 8 shows the circuit implementation and the corresponding state diagram.

During the reset phase of the photodiode the digital logic is set to state $Z_1 = \{1, 0\}$ by the column and row reset signals $V_{\text{Rst}_B,y}$ and $V_{\text{Rst}_B,x}$. The comparator input is connected to V_{refH} . When the voltage V_{int} reaches V_{refH} , the output logic transmits the row request signal $V_{\text{Req}_B,y}$ to the peripheral AE circuitry. After the corresponding row acknowledge signal $V_{\text{Ack}_B,y}$ switches, the column request signal $V_{\text{Req}_B[H],x}$ is activated. With the subsequent column acknowledge $V_{\text{Ack}_B[H],x}$ going high, the state logic switches to state $Z_2 = \{0, 1\}$, the request signals turn off and the comparator input voltage is switched to V_{refL} (V_{mem2} is going low). When the lower threshold is reached, the communication process resumes, signaling the x -address by asserting $V_{\text{Req}_B[L],x}$. After the column acknowledge signal has been received, the logic switches to the idle state $Z_0 = \{0, 0\}$ where it remains until a new exposure measurement cycle is started.

6) *Pixel Layout*: The chip has been implemented in a standard 0.18 μm one-poly-six-metal (1P6M) mixed-mode/RF CMOS process. Fig. 9 shows the layout of the pixel with the main circuit parts and the CD transistors annotated. The square pixel covers 900 μm^2 of silicon area (30 μm pitch). The two photodiodes for continuous time operation of the change detector (PD2) and integrating PWM exposure measurement (PD1) are placed side by side at the top edge of the pixel area. The fill factor of the pixel is 10% of total pixel area for the change detector and 20% of total pixel area for the exposure measurement part.

V. IMAGER DESIGN

The ATIS sensor is built around a QVGA (304 \times 240) array of pixels. A block schematic of the sensor system architecture, consisting of the pixel array, column and row-wise reset and trigger control, address encoders and AER periphery for asynchronous data transfer [16], is shown in Fig. 10. Other building blocks like bias generator, global exposure measurement and test structures are omitted in the diagram. Fig. 11 shows a microphotograph of the fabricated sensor chip.

A. Data Readout

The pixels in the sensor array communicate with column and row arbiters via 4-phase AER handshaking as described in de-

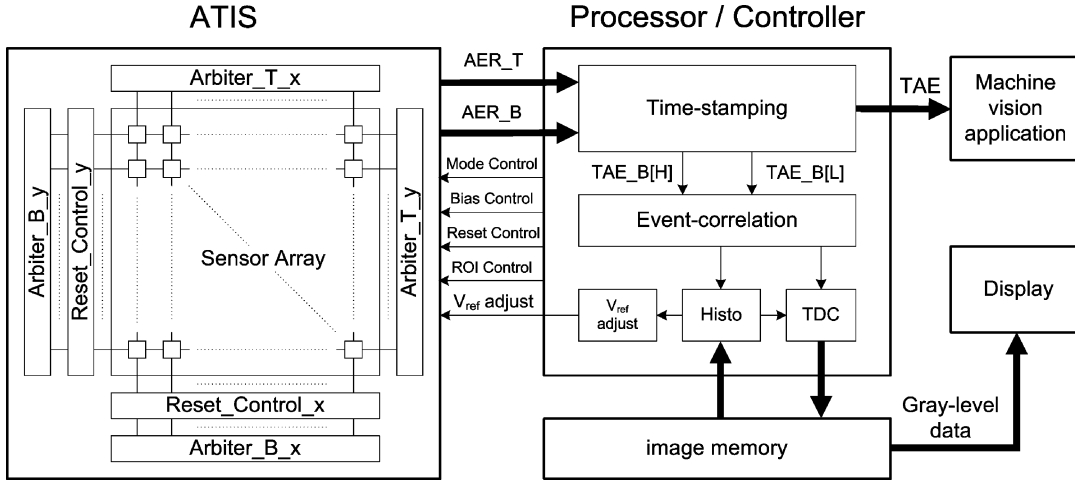


Fig. 12. Block diagram of the ATIS imager system.

signal. The global reset mode, functionally identical to the time-to-first-spike (TTFS) scheme [34], can run concurrently to the normal operation mode, and allows the quick acquisition of a reference frame during normal operation or an overlaid synchronous video mode.

- In *sequential reset mode (ACPR, asynchronous column-parallel readout)*, the reset of one pixel (N) in one column is triggered by the $Ack_B[H/L]$ signal of the preceding pixel ($N - 1$). After triggering the pixels of the first (top) row of the pixel array, the trigger runs in parallel along the columns, each pixel triggering its bottom neighbor when it has reached the first integration threshold V_{refH} . This asynchronous “rolling shutter” mode is intended to avoid a familiar problem of event collisions in TTFS imagers seeing highly uniform scenes (having many pixels finishing integration and trying to send their event at the same time), at the cost of slower image acquisition. Multiple pixel rows across the array (e.g., rows 1 and 121, or 1, 61, 121, 181) can simultaneously be selected as starting rows to decrease frame acquisition time at the cost of higher event collision probability. Also this mode can run concurrently to the normal operation (ATIS) mode. The ACPR mode has been described in detail and analyzed in [35].

C. Imaging System

Fig. 12 shows a block diagram of the complete imaging system, containing the ATIS sensor, a processor/controller, and image memory. In the current implementation, processor and memory is off-chip and is realized as an ASIC [44], however it is planned to integrate all building blocks into one Vision-System-on-Chip in the future.

The processor/controller contains a high-resolution (≥ 10 ns) digital counter based time-stamping device for the incoming address-events and an event-correlator, which matches the timed address-event pairs ($TAE_B[H] - TAE_B[L]$) for “time-to-digital conversion” (TDC). The time resolution of the system is adequate to allow a minimum of 8-bit grayscale resolution for the image data over the full dynamic range.

The grayscale data can be transmitted on-demand (i.e., when they occur), e.g., as UDP packets, to a remote receiver and can be displayed, or can be worked on *in situ* or remotely using event-based image processing algorithms, while, e.g., simultaneously the “timed address-events” (TAE_T) from the CD part can, e.g., drive a machine vision application.

VI. CHARACTERIZATION

The following subsections contain results of selected laboratory tests and measurements and application-oriented characterization. Measurement results are put into context of corresponding theoretical considerations wherever feasible.

A. Change Detector

1) *Event Latency*: The latency of the change detector response, an important parameter for dynamic vision, is the time from the occurrence of an illumination change at the change detector photodiode to the corresponding event output (circuit in Fig. 4). This variable delay time is, at low illumination, limited by the photoreceptor bandwidth which can be approximated as a first-order low-pass with corner frequency according to (5). The corner frequency is proportional to the photocurrent I_{ph} and hence to pixel illumination. The resulting latency for a step transient from I_{ph1} to I_{ph2} can be written as

$$T_{Lat} \approx -\ln \left(1 - \frac{K_{Min}}{\left(\frac{I_{ph2}}{I_{ph1}} \right)} \right) \cdot \frac{C_{DMfb} \cdot U_T}{I_{ph1}} \quad (14)$$

with K_{Min} being the minimum photocurrent step for reaching the event threshold and $I_{ph2}/I_{ph1} \gg K_{Min}$.

At high light levels, the photoreceptor is fast and the latency is dominated by the speed of amplifier A2 and the comparators [compare Fig. 3(c)], approaching a constant level with regard to illumination. With amplifier A2 and comparators operating in the subthreshold region, the latency (e.g., for OFF polarity events) is given by

$$T_{Lat} = T_{Lat,Diff} + T_{Lat,Comp} \quad (15)$$

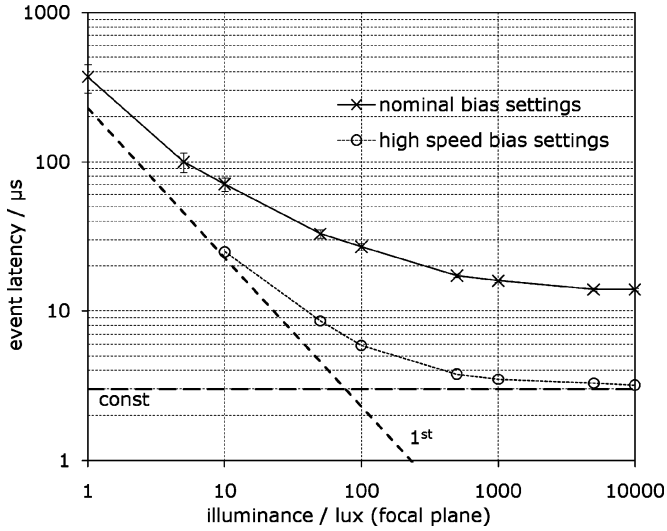


Fig. 13. Pixel response latency versus illumination for two change detector circuit operating points.

with

$$T_{\text{Lat,Diff}} \approx -\ln \left(1 - \frac{K_{\text{Min}}}{\ln \left(\frac{I_{\text{ph2}}}{I_{\text{ph1}}} \right)} \right) \cdot \frac{C_1}{C_2} \cdot \frac{C_{A2\text{out}}}{g_{m,dp}}, \quad (16)$$

$$T_{\text{Lat,Comp,OFF}} \approx \frac{L_{\text{OFFn}} C_{\text{OFFout}}}{W_{\text{OFFn}} I_{0,\text{OFFn}}} \cdot e^{-\kappa_n \cdot V_{\text{OFF}}/U_T} (V_{\text{DD}} - V_{\text{S,Inv}}). \quad (17)$$

The parameters in the equations are: $C_{A2\text{out}}$ the capacitive load at the output of amplifier A2, L_{OFFn} and W_{OFFn} the transistor dimensions of the load transistor M_{OFFn} in the OFF-polarity comparator, C_{OFFout} is the capacitive load at comparator output, and $V_{\text{S,Inv}}$ the threshold level of the following digital stage.

Fig. 13 shows measured event latency as a function of illumination for different operating points. The two boundary conditions, first-order roll-off and approaching constant w.r.t. illuminance, are marked in the plot. For high-current bias settings, the pixel latency goes below $10 \mu\text{s}$ for illumination above $\sim 30 \text{ Lx}$ and approaches $3 \mu\text{s}$ for bright light conditions. The temporal resolution of the CD is of the order of 100 k fps.

2) *Contrast Sensitivity*: The sensitivity to temporal contrast, or relative change in illumination, is the second important performance parameter of the change detector. A temporal contrast of

$$\frac{I_{\text{ph2}} - I_{\text{ph1}}}{I_{\text{ph1}}} = -\frac{C_2}{C_1} \frac{\kappa_{\text{Mfb}} \kappa_{\text{Mn}}}{\kappa_{\text{Mp}}} \cdot \left(\frac{V_{\text{b,diff}} - V_{\text{b,on}}}{U_T} - \ln 2 \right) \quad (18)$$

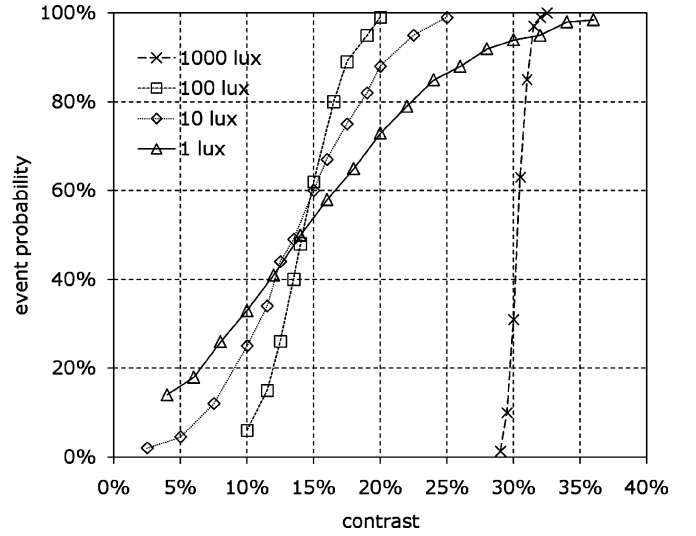


Fig. 14. Contrast sensitivity: Response probability versus stimulus contrast for four orders of magnitude of illumination. The 50% point of contrast sensitivity is constant from 1 lx to 100 lx at 13% contrast and moves towards lower sensitivity (higher contrast for equal response probability) at higher illumination. The (maximum) slope of the S-curve is inversely proportional to circuit noise (dominated by photocurrent shot noise) and hence to pixel illumination, increasing steadily from 1 lx to 1000 lx.

ideally results in the generation of one (ON polarity) event. The symbols used are κ for the subthreshold slope factor of the respective transistors, $V_{\text{b,diff}}$ and $V_{\text{b,on}}$ for gate bias voltages at the referred nodes and U_T for the thermal voltage (compare Fig. 4). This simplified equation assumes unity gain of the source follower buffer and weak-inversion operation of all amplifier and comparator stages and shows light independent contrast sensitivity, depending only on device parameters and bias settings [8].

One way to evaluate and depict contrast sensitivity is to measure and plot the event response probability¹ as a function of increasing contrast at identical initial illuminance. In an ideal, noise-free world, this would result in a step (0% to 100% probability with infinite slope) at a given threshold contrast $(I_{\text{ph2}} - I_{\text{ph1}})/I_{\text{ph1}}$. In reality, noise turns the ideal step into an “S”-shaped curve.

Fig. 14 shows S-curves illustrating the pixel’s contrast sensitivity. Plotted is the response probability as a function of temporal contrast given in % of $(I_{\text{ph2}} - I_{\text{ph1}})/I_{\text{ph1}}$ where I_{ph1} and I_{ph2} are the photocurrents before and after a step change in pixel illumination. The 50%-response probability point in the shown region of highest contrast sensitivity—between 1 Lx and 100 Lx—is constant at around 13% contrast for the chosen operating point settings. The (maximum) slope of the S-curve is inversely proportional to the circuit noise (dominated by photocurrent shot noise) and hence to pixel illumination. For further increasing illumination, the slope steepens further but starts moving towards higher stimulus contrasts. The decrease of contrast sensitivity for increasing photocurrent is not yet fully understood.

¹An event probability of 50% means that the change detector generates on average 50 events in response to 100 identical stimuli.

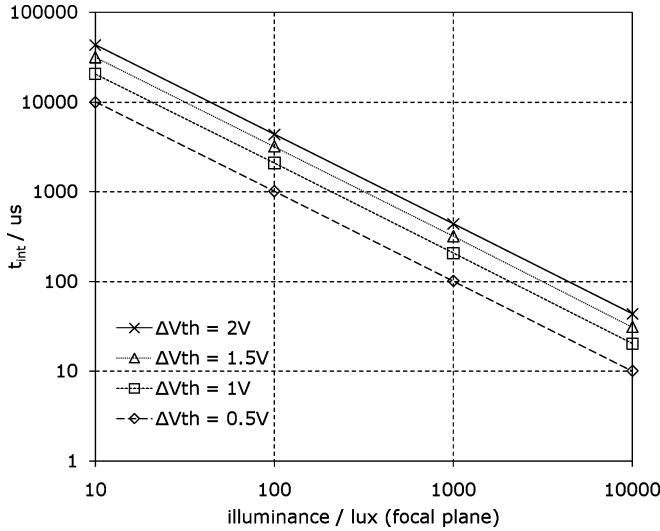


Fig. 15. PWM transfer function (integration time versus lux) for four different values of integration swing ΔV_{th} .

B. PWM Imaging

The charge capacity of the integration node depends on operation voltages and approaches $Q_{int,max} \approx 450,000$ electrons at maximum (2 V) integration swing ΔV_{th} . Sense node capacitance is 36 fF, yielding a conversion gain of $4.4 \mu V/e^-$. Photodiode darkcurrent has been measured at ~ 3 fA, darkcurrent shot noise evaluates to $670 e^-$, corresponding to 3 mV rms, for an integration swing ΔV_{th} of 2 V.

1) *PWM Imaging—Transfer Function*: What is exposure time to conventional voltage-mode image sensors is the integration voltage swing in time-domain imaging. Fig. 15 plots measured integration times for integration swings ΔV_{th} between 0.5 V and 2 V as a function of pixel illumination. The theoretically asserted $1/x$ -relation is accurately satisfied. Integration times range from, e.g., 10 ms @ 10 Lx to 10 μs @ 10 kLx for an integration swing of 500 mV.

2) *Image Sensor Signal-to-Noise Ratio*: In Fig. 16 measured imager SNR as a function of integration swing for different light intensity is shown. SNR is >56 dB at an integration swing ΔV_{th} of 2 V and light levels above 10 Lx. Standard 8 bit grayscale resolution (48 dB) is achieved for very low illuminations and small integration swings. For $\Delta V_{th} = 100$ mV and 10 Lx, SNR is still at 42 dB, allowing for 7 bit-resolution imaging at very short integration times (<2 ms @ 10 Lx). The result is 500 fps equivalent temporal resolution imaging and video at low-light conditions. The weak dependence of SNR on illuminance for time-based image sensors, as illustrated in Fig. 2, is well observable in the measured data.

C. Imager Dynamic Range

The usable sensor DR entails a trade-off with the temporal resolution required to capture scene dynamics and is limited by maximum allowable integration time at the dark end and AER communication channel data throughput at the bright end. Fig. 17 shows image data acquired with the ATIS sensor from a static high-DR scene in one exposure. The picture was taken in an otherwise dim room with the sensor pointing towards a high-power LED, using the externally triggered “snapshot” mode

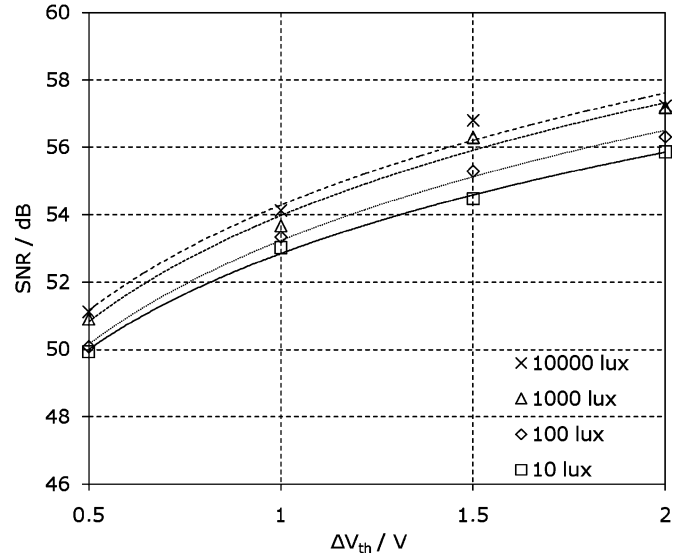


Fig. 16. Measured SNR as functions of integration swing ΔV_{th} and light intensity. SNR is >56 dB for an integration swing of 2 V for light levels above 10 Lx. For $\Delta V_{th} = 100$ mV and 10 Lx SNR is still at 42.3 dB.

[35]. The first seven images, (a)–(g), show different scalings of the exposure data with the scaling shifted by a factor of 10 between images, while (h) is an illustrative attempt to generate an 8 bit (48 dB) composite image by equalizing the data using a histogram method [45]. The acquired data—exhibiting a minimum integration time of 350 ns and a maximum integration time of 4.5 s—demonstrate that the ATIS sensor is capable of reproducing an intra-scene DR of at least 143 dB [compare (4)].

A maximum integration time of several seconds however is inadequate for many applications. To increase the temporal resolution without trading-off much DR, a method that is complementary to the multiple-exposure technique, used for DR improvement in standard voltage-mode imagers, is proposed. Due to the presence of two independent thresholds, normally used for differential TCDS, it is possible to apply *two* integration swings (the complementary parameter to exposure time) during *one* image acquisition. With DR still at 143 dB, the longest integration times can be reduced by a factor of 20 (to $<\sim 200$ ms, equivalent to 5 fps) if the time between *change event* and *upper TCDS threshold* is used to determine pixel exposure in the dark parts of the scene (with $V_{DD} - V_{refH} = (V_{refH} - V_{refL})/20$). Consequently for a temporal resolution of 33 ms (30 fps video speed equivalent temporal resolution), a DR of 125 dB is achieved. The penalty to pay for this high DR at high-speed operation is reduced SNR and higher FPN (no TCDS) for pixels that reach only the first threshold. From a system operation point of view, each pixel, depending on individual illumination, practically chooses for itself which threshold to use. The second threshold event is either simply ignored by the post-processor when arriving too late, or will never appear since a new exposure has been started before (compare Section III-B).

D. Video Compression

The temporal redundancy suppression of the ATIS change-detector controlled operation ideally yields lossless focal-plane

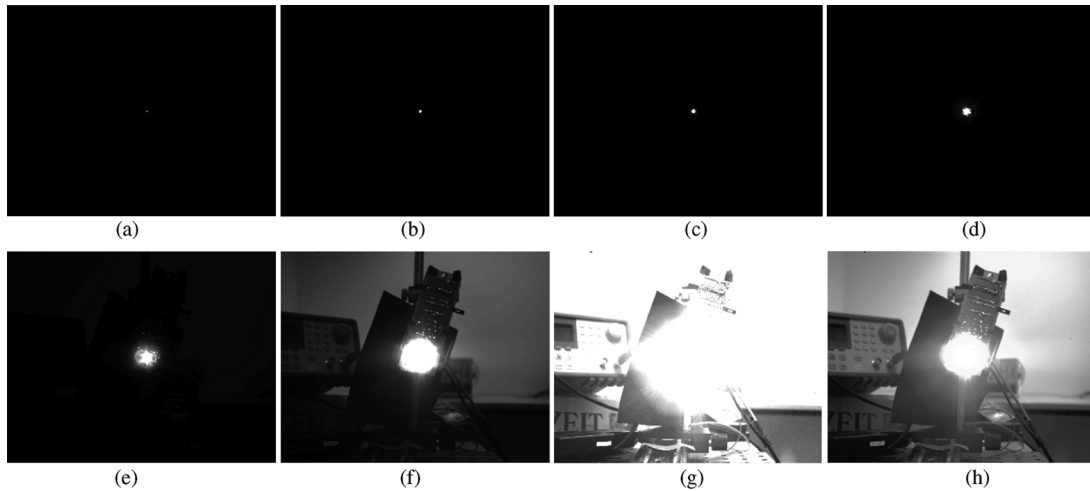


Fig. 17. High-DR imaging: Seven scalings, (a)–(g), of the same exposure image data and composite image (h). In order to print or display (on a conventional screen) the data in a meaningful way, one illustrative approach is to create a sequence of images, each exploiting the available 8 bit (48 dB) grayscale resolution. Starting at (a), the shortest integration time is mapped to the grayscale value 255 (white), while integration times of 255 times the shortest value and longer are mapped to the grayscale value 0 (black). In (b) the shortest integration time multiplied by a factor of 10 is mapped to white and, as before, 255 times the integration time of white is mapped to black. This procedure is repeated until arriving at image (g) with a multiplication factor of 1,000,000. The sensor data thereby remain unaltered. The shortest integration time in this exposure example is 350 ns while the longest is 4.7 s, yielding a static scene dynamic range of 143 dB reproduced by the sensor. Image (h) is an illustrative attempt to generate an 8 bit composite image by equalizing the data using a histogram-based method [45].

video compression with compression factors depending only on scene dynamics. Theoretically approaching infinity for static scenes, in practice, due to sensor non-idealities, the achievable compression factor is limited and appears to be of the order of 1000 for bright static scenes as compared to a conventional, frame-based imager of the same resolution delivering raw 8 bit grayscale data at video speed of 30 fps.

Fig. 18 shows a typical surveillance scene generating a 2.5 k to 50 k events/s @ 18 bit/event continuous-time video stream. The actual event rate depends on instantaneous scene activity. Comparing corresponding bit rates—45 k to 900 k bit/s—to the raw data rate of a QVGA 8 bit grayscale sensor at 30 fps of 18 Mbit/s demonstrates lossless video compression with compression factors between 20 and 400 for this example scene. Fig. 18(a) contains a still frame taken from a continuous-time video sequence; Fig. 18(b) shows the same frame assuming video transmission has started from an empty image. The effect of objects triggering exposure measurement in the pixel they hit while moving across the focal plane becomes visible (e.g., a white car moving from the bottom-left corner of the image towards the center).

Static background does not produce data apart from the odd CD noise event, also triggering exposure measurement in the respective pixel. This effect reduces the achievable video compression factor to about 1000 for bright, static scenes (from infinity in an ideal, noise-free world). On the other hand this effect is useful for capturing very slow changes in the scene (like, e.g., varying scene illumination from sun light due to passing clouds) through a continuous statistically distributed slow update of the entire image. Typical background noise activity is of the order of 1 k–3 k pixels per second, about 3 orders of magnitude below raw data rate from a conventional, frame-based sensor of same array size running at 30 fps.

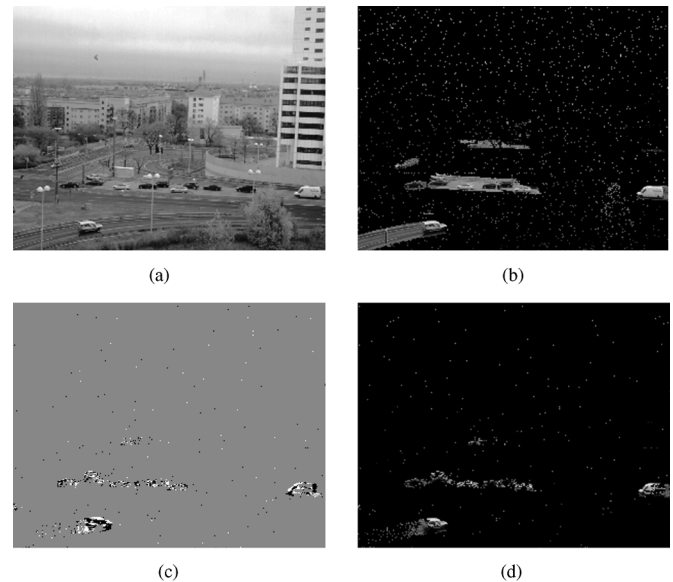


Fig. 18. Traffic scene generating between 2.5 k and 50 k events/s, depending on instantaneous scene activity. The video compression factor w.r.t. raw data from a QVGA 30 fps 8 bit grayscale sensor was measured to be 20–400 for this example scene. The video compression is essentially lossless, no dynamic image errors or artefacts are visible in the video (a).

Fig. 18(c) shows change detector events collected during a time slice of 33 ms (ON events in white and OFF events in black) and Fig. 18(d) shows the grayscale data generated in response to the change events in (c). The compression gain is even larger when the much higher temporal resolution (e.g., 500 fps equivalent for scenes > 10 Lx) of the ATIS sensor is taken into account.

E. Fixed Pattern Noise

Due to the lack of appropriate laboratory equipment, the imager FPN has not yet been measured rigorously. However an

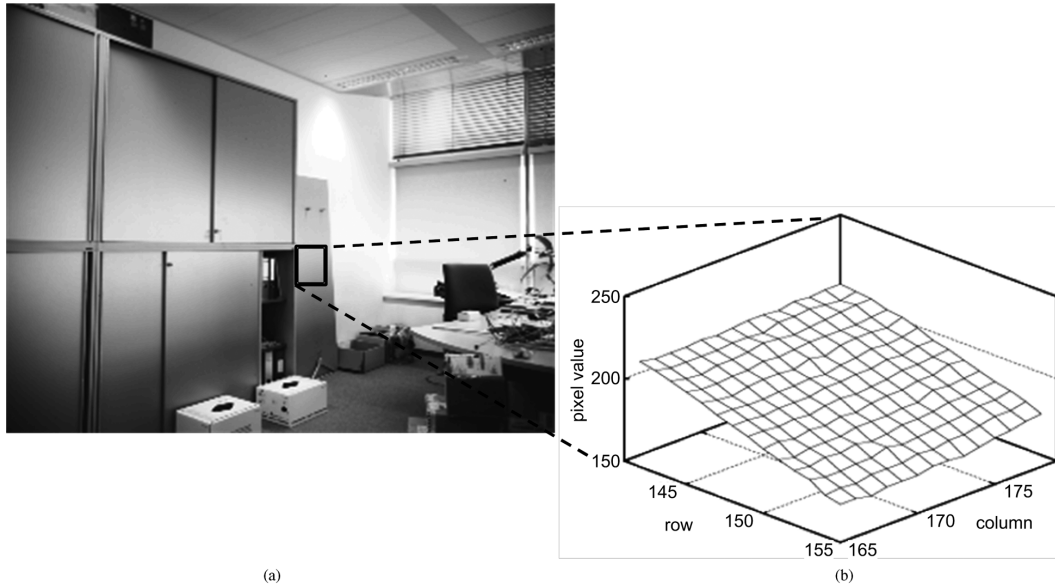


Fig. 19. (a) Indoor scene at ~ 100 Lx. (b) Mesh plot of grayscale values of the gradient in the marked area of (a).

upper bound of 0.25% was established from evaluating different homogenous parts of recorded image series, similar to the one shown in Fig. 19(a). Fig. 9(b) shows gray scale values of an approximately planar gradient taken from the image in Fig. 19(a) displayed as a mesh plot for illustrating pixel response uniformity. Assuming correctness of mismatch parameters published by the process vendor, the theoretically determined array FPN of below 0.2% seems in reach. Also owing to the SNR of 56 dB (9.3 bit), grayscale gradients are resolved smoothly without visible artefacts [Fig. 19(a)].

F. Summary Table

Table I provides an overview of the ATIS sensor's main specifications and important test and measurement results.

VII. CONCLUSION

A biomimetic, frame-free, wide DR vision and image sensor with pixel-level video compression is presented. The sensor comprises an array of autonomous pixels that individually detect illumination changes and asynchronously encode in inter-event intervals the instantaneous pixel illumination after each detected change, ideally realizing highly efficient video compression through temporal redundancy suppression at the pixel-level. Familiar deficiencies of time-based imagers have been remedied (a) using a novel time-domain correlated double sampling (TCDS) technique and (b) by realizing illumination-dependent temporal readout load spreading. Intra-scene DRs of 143 dB static and 125 dB @ $t_{\text{int}} < 33$ ms have been achieved, in line with or better than recent high-DR developments reported, e.g., in [46] and [47]. Target application areas are high-speed high-temporal-resolution dynamic machine vision, e.g., robotics or visual feedback loops, low-data rate video for wireless or TCP based applications, and wide DR, high-quality high-temporal-resolution imaging and video, e.g., scientific applications.

TABLE I
SUMMARY SENSOR CHARACTERISTICS

Parameter	Value
Fabrication process	UMC L180 MM/RF 1P6M Standard CMOS
Supply voltage	3.3V (analog), 1.8V (digital)
Chip size	$9.9 \times 8.2 \text{mm}^2$
Optical format	2/3"
Array size	QVGA (304×240)
Pixel size	$30 \mu\text{m} \times 30 \mu\text{m}$
Pixel complexity	77T, 3C, 2PD
Fill factor	30% (20% EM, 10% CD)
Integration swing ΔV_{th}	100mV to 2.3 V (adjustable)
SNR typ.	$> 56 \text{dB}$ (9.3bit) @ $\Delta V_{\text{th}} = 2\text{V}$, $> 10 \text{Lx}$
SNR low	42.3dB (7bit) @ $\Delta V_{\text{th min}}$ (100mV), 10Lx
t_{int} @ $\Delta V_{\text{th min}}$ (100mV)	2ms @ 10Lx (500 fps equ. temp. resolution)
Temporal resolution EM	500fps equ. (@ 10Lx), 50kfps equ. (@ 1000Lx)
Temporal resolution CD	100kfps equ. (@ $> 100 \text{Lx}$)
DR (static)	143dB
DR (30fps equivalent)	125dB
FPN	$< 0.25\%$ @ 10Lx (with TCDS)
Sense Node Cap	36fF
Conversion gain	$4.4 \mu\text{V}/e^-$
Darkcurrent	$1.6 \text{nA}/\text{cm}^2$ (@ 25°C)
Power consumption	50mW (static), 175mW (high activity)
Readout format	Asynchronous address-events (AER), $2 \times 18 \text{bit-parallel}$

ACKNOWLEDGMENT

The authors would like to acknowledge T. Delbrück and P. Lichtsteiner of ETH/UNI Zürich for original change detector design, numerous discussions and joint efforts in the realization of neuromorphic dynamic vision devices and applications.

REFERENCES

- [1] C. Mead, *Analog VLSI and Neural Systems*. Boston, MA: Addison-Wesley, 1989.
- [2] C. Mead, "Neuromorphic electronic systems," *Proc. IEEE*, vol. 78, pp. 1629–1636, Oct. 1990.
- [3] M. A. C. Maher, S. P. Deweerth, M. A. Mahowald, and C. A. Mead, "Implementing neural architectures using analog VLSI circuits," *IEEE Trans. Circuits Syst.*, vol. 36, no. 5, pp. 643–652, May 1989.
- [4] M. A. Mahowald and C. A. Mead, "The silicon retina," *Scientific American*, May 1991.
- [5] K. Boahen, "Neuromorphic microchips," *Scientific American*, vol. 292, pp. 55–63, May 2005.
- [6] A. H. C. Van Der Heijden, *Selective Attention in Vision*. New York: Routledge, 1992, ISBN: 0415061059.
- [7] P. Lichtsteiner and T. Delbruck, "A 64×64 AER logarithmic temporal derivative silicon retina," *Research in Microelectronics and Electronics, 2005 Ph.D.*, vol. 2, pp. 202–205, Jul. 2005.
- [8] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μ s latency asynchronous temporal contrast vision sensor," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, Feb. 2008.
- [9] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 30 mW asynchronous vision sensor that responds to relative intensity change," in *IEEE ISSCC 2006 Dig. Tech. Papers*, Feb. 6–9, 2006, pp. 2060–2069.
- [10] F. W. Mounts, "A video coding system with conditional picture-element replenishment," *Bell Syst. Tech. J.*, pp. 2545–2554, Sep. 1969.
- [11] Y. Chin and T. Berger, "A software-only videocodec using pixelwise conditional differential replenishment and perceptual enhancements," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 3, pp. 438–450, Mar. 1999.
- [12] K. Aizawa, Y. Egi, T. Hamamoto, M. Hatori, M. Abe, H. Maruyama, and H. Otake, "Computational image sensor for on sensor compression," *IEEE Trans. Electron Devices*, vol. 44, no. 10, pp. 1724–1730, Oct. 1997.
- [13] V. Gruiev and R. Etienne-Cummings, "A pipelined temporal difference imager," *IEEE J. Solid-State Circuits*, vol. 39, no. 3, pp. 538–543, Mar. 2004.
- [14] J. Yuan, H. Y. Chan, S. W. Fung, and B. Liu, "An activity-triggered 95.3 dB DR –75.6 dB THD CMOS imaging sensor with digital calibration," *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2834–2843, Oct. 2009.
- [15] 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. Solid-State Circuits*, vol. 42, no. 10, pp. 2187–2196, Oct. 2007.
- [16] K. Boahen, "A burst-mode word-serial address-event link-I: Transmitter design," *IEEE Trans. Circuits Syst. I*, vol. 51, no. 7, pp. 1269–1280, 2004.
- [17] K. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Trans. Circuits Syst. II*, vol. 47, no. 5, pp. 416–434, 2000.
- [18] G. Ward, "The hopeful future of high dynamic range imaging," in *2007 SID Int. Symp.*, May 22–25, 2007.
- [19] S. J. Decker, R. D. McGrath, K. Brehmer, and C. G. Sodini, "A 256×256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, Dec. 1998.
- [20] T. Lulé, B. Schneider, and M. Böhm, "Design and fabrication of a high dynamic range image sensor in TFA technology," *IEEE J. Solid-State Circuits*, vol. 34, no. 5, pp. 704–711, May 1999.
- [21] D. X. D. Yang, A. E. Gamal, B. Fowler, and H. Tian, "A 640×512 CMOS image sensor with ultrawide dynamic range floating-point pixel-level ADC," *IEEE J. Solid-State Circuits*, vol. 34, no. 12, pp. 1821–1834, Dec. 1999.
- [22] V. Brajovic and T. Kanade, "A VLSI sorting image sensor: Global massively parallel intensity-to-time processing for low-latency adaptive vision," *IEEE Trans. Robot. Automat.*, vol. 15, no. 1, pp. 67–75, 1999.
- [23] J.-E. Eklund, C. Svensson, and A. Astrom, "VLSI implementation of a focal plane image processor—a realization of the near-sensor image processing concept," *IEEE Trans. VLSI*, vol. 4, no. 3, pp. 322–335, 1996.
- [24] M. Nagata, J. Funakoshi, and A. Iwata, "A PWM signal processing core circuit based on a switched current integration technique," *IEEE J. Solid-State Circuits*, vol. 33, no. 1, pp. 53–60, Jan. 1998.
- [25] K. Frohmader, "A novel MOS compatible light intensity-to-frequency converter suited for monolithic integration," *IEEE J. Solid-State Circuits*, vol. 17, no. 3, pp. 588–591, 1982.
- [26] K. Tanaka *et al.*, "Novel digital photosensor cell in GaAs IC using conversion of light intensity to pulse frequency," *Jpn. J. Appl. Phys.*, vol. 32, no. 11A, pp. 5002–5007, 1993.
- [27] W. Yang, "A wide-dynamic-range, low-power photosensor array," in *IEEE ISSCC 1994 Dig. Tech. Papers*, 1994, pp. 230–231.
- [28] A. Kitchen, A. Bermak, and A. Bouzerdoum, "A digital pixel sensor array with programmable dynamic range," *IEEE Trans. Electron Devices*, vol. 52, no. 12, pp. 2591–2601, Dec. 2005.
- [29] Q. Luo and J. Harris, "A time-based CMOS image sensor," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2004)*, 2004, vol. IV, pp. 840–843.
- [30] D. Bauer *et al.*, "Embedded vehicle speed estimation system using an asynchronous temporal contrast vision sensor," *EURASIP J. Embedded Syst.*, vol. 2007, 2007, doi: 10.1155/2007/82174.
- [31] C. Posch *et al.*, "A dual-line optical transient sensor with on-chip precision time-stamp generation," in *IEEE ISSCC 2007 Dig. Tech. Papers*, Feb. 11–15, 2007, pp. 500–501.
- [32] D. Matolin, C. Posch, and R. Wohlgenannt, "True correlated double sampling and comparator design for time-based image sensors," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2009)*, May 24–27, 2009, pp. 1269–1272.
- [33] C. Posch, D. Matolin, and R. Wohlgenannt, "An asynchronous time-based image sensor," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2008)*, 2008, pp. 2130–2133.
- [34] X. Guo, X. Qi, and J. Harris, "A Time-to-first-spike CMOS image sensor," *IEEE Sensors J.*, vol. 7, pp. 1165–1175, 2007.
- [35] D. Matolin, R. Wohlgenannt, M. Litzberger, and C. Posch, "A load-balancing readout method for large event-based PWM imaging arrays," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2010)*, May 2010.
- [36] D. Matolin, C. Posch, and R. Wohlgenannt, "Area and power reduction techniques for time-based image sensor pixel design," in *21th Int. Conf. Microelectronics (ICM 2009)*, Dec. 2009.
- [37] T. Delbruck and C. A. Mead, "Adaptive photoreceptor with wide dynamic range," in *IEEE Int. Symp. Circuits and Systems (ISCAS 1994)*, 1994, vol. 4, pp. 339–342.
- [38] M. H. White *et al.*, "Characterization of surface channel CCD image arrays at low light levels," *IEEE J. Solid-State Circuits*, vol. SCC-9, no. 1, 1974.
- [39] C. Xu, S. Chao, and M. Chan, "A new correlated double sampling (CDS) technique for low voltage design environment in advanced CMOS technology," in *Proc. ESSCIRC 2002*, pp. 117–120.
- [40] C. Xu, C. Shen, A. Bermak, and M. Chan, "A new digital-pixel architecture for CMOS image sensor with pixel-level ADC and pulsewidth modulation using a $0.18 \mu\text{m}$ CMOS technology," in *Proc. IEEE Conf. Electron Devices and Solid-State Circuits*, 2003, pp. 265–268.
- [41] B. R. Chawla and H. K. Gummel, "Transition region capacitance of diffused p-n junctions," *IEEE Trans. Electron Devices*, vol. ED-44, no. 3, pp. 178–195, 1971.
- [42] K. R. Lakshmi Kumar *et al.*, "Characterization and modeling of mismatch in MOS transistors for precision analog design," *IEEE J. Solid-State Circuits*, vol. SCC-21, pp. 1057–1066, 1986.
- [43] S. Kleinfelder *et al.*, "A 10,000 frames/s CMOS digital pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, no. 12, pp. 2049–2059, Dec. 2001.
- [44] M. Hofstaetter, P. Schön, and C. Posch, "A SPARC-compatible general purpose address-event processor with 20-bit 10 ns-resolution asynchronous sensor data interface in $0.18 \mu\text{m}$ CMOS," in *IEEE Int. Symp. Circuits and Systems (ISCAS 2010)*, May 2010.
- [45] R. Wang, "Histogram Equalization," 2009 [Online]. Available: http://fourier.eng.hmc.edu/e161/lectures/contrast_transform/node3.html
- [46] T. Yamada *et al.*, "A 140 dB-dynamic-range MOS image sensor with in-pixel multiple-exposure synthesis," in *IEEE ISSCC 2008 Dig. Tech. Papers*, Feb. 3–7, 2008, pp. 50–594.
- [47] P.-F. Ruedi *et al.*, "An SoC combining a 132 dB QVGA pixel array and a 32b DSP/MCU processor for vision applications," in *IEEE ISSCC 2009 Dig. Tech. Papers*, Feb. 8–12, 2009, vol. 47a, pp. 46–47.



Christoph Posch (M'07) received the M.Sc. and Ph.D. degrees in electrical and electronics engineering and experimental physics from Vienna University of Technology, Vienna, Austria, in 1995 and 1999, respectively.

From 1996 to 1999, he worked on analog CMOS and BiCMOS IC design for particle detector readout and control at CERN, the European Laboratory for Particle Physics in Geneva, Switzerland. From 1999 onwards he was with Boston University, Boston, MA, engaging in applied research and analog/mixed-signal integrated circuit design for high-energy physics instrumentation. In 2004 he joined the newly founded Smart Sensors Group at AIT Austrian Institute of Technology (formerly Austrian Research Centers ARC) in Vienna, Austria, where he was promoted to Principal Scientist in 2007. His current research interests include neuromorphic analog VLSI, CMOS image and vision sensors, and biology-inspired signal processing.

Dr. Posch has been recipient and co-recipient of several scientific awards including the Jan van Vessel Award for Outstanding European Paper at the IEEE International Solid-State Circuits Conference (ISSCC) in 2006. He is a member of the Sensory Systems and the Neural Systems and Applications Technical Committees of the IEEE Circuits and Systems Society. He has authored more than 50 scientific publications and holds several patents in the area of vision and image sensing.



Daniel Matolin received the Dipl.-Ing. (M.Sc.) degree in electrical engineering from the Dresden University of Technology, Dresden, Germany, in 2003.

From 2003 to 2005, he was a research fellow at TU Dresden, working on neuromorphic VLSI and image processing. Currently, he is with the Neuroinformatics and Smart Sensors Group at AIT Austrian Institute of Technology in Vienna, Austria. His research interests include vision sensors and mixed-signal circuit design.

Mr. Matolin is an author and coauthor of more than 30 scientific publications, and corecipient of several IEEE publication awards including the Best Paper Award at the IEEE International Conference on Electronics, Circuits and Systems in 2006 and the Best Demonstration Award at the IEEE International Symposium on Circuits and Systems (ISCAS) in 2010.



Rainer Wohlgenannt received the M.Sc. degree in electrical engineering with honors from Vienna University of Technology, Vienna, Austria, in 2004.

In 2004 he concluded his master theses at the Telecommunications Research Center Vienna (ftw.), Vienna, Austria and was invited guest scientist at the Centre for Wireless Communications (CWC), Oulu, Finland. From 2004 to 2007 he was with Austriamicrosystems AG, Graz, Austria as a project leader and VLSI design engineer for high performance analog ASSPs. In 2007 he joined the Smart Sensors

Group at AIT Austrian Institute of Technology in Vienna, Austria. His research interests include CMOS image sensors and high performance analog VLSI design.

Mr. Wohlgenannt is author and coauthor of more than 15 scientific publications, and co-recipient of the Best Paper Award at IEEE International Conference on Electronics, Circuits and Systems, ICECS, in 2006 and the Best Demonstration Award at the IEEE International Symposium on Circuits and Systems ISCAS in 2010.