High-DR Frame-Free PWM Imaging with asynchronous AER Intensity Encoding and Focal-Plane Temporal Redundancy Suppression

Christoph Posch, *Member IEEE*, Daniel Matolin, Rainer Wohlgenannt
AIT – Austrian Institute of Technology GmbH, Vienna, Austria
{christoph.posch | daniel.matolin | rainer.wohlgenannt}@arcs.ac.at

***Abstract*—The presented asynchronous, time-based CMOS dynamic vision and image sensor is based on a QVGA (304×240) array of fully autonomous pixels containing event-based change detection and PWM imaging circuitry. Exposure measurements are initiated and carried out locally by the individual pixel that has detected a brightness change in its field-of-view. Thus pixels do not rely on external timing signals and independently and asynchronously request access to an (asynchronous arbitrated) output channel when they have new illumination values to communicate. Communication is address-event based (AER) – gray-levels are encoded in inter-event intervals. Pixels that are not stimulated visually do not produce output. This pixel-autonomous and massively parallel operation ideally results in optimal lossless video compression through complete temporal redundancy suppression at the focal-plane. Compression factors depend on scene activity. Due to the time-based encoding of the illumination information, very high dynamic range – intra-scene DR of 143dB static and 125dB at 30fps equivalent temporal resolution – is achieved. A novel time-domain correlated double sampling (TCDS) method yields array FPN of <0.25%. SNR is >56dB (9.3bit) for >10Lx.**

I. Introduction

A fast-growing number of high-volume consumer imaging products integrate CMOS image sensors. Due to the continuous advances in deep-submicron CMOS processes it becomes increasingly feasible to build high-performance single-chip cameras, integrating image capture and advanced on-chip processing circuitry into the focal plane. By many involved in the field, CMOS is considered the image sensing technology of the future. Despite all progress, several issues with solid-state imaging remain unresolved.

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, usually not long ago. 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, increased memory size and post-processing power, up to the point where they become prohibitive e.g. in battery-powered portable or wireless applications. Image compression partly relaxes these requirements but at the cost of additional complex processing.

According to notable experts in the field, it is clear that high dynamic range imaging will dominate the market in the near future [1]. Dynamic range (DR) is defined as the ratio of the maximum processible signal and the noise floor under dark conditions. The DR is typically limited by finite pixel capacity for integrating photo charges, the noise floor and pixel-identical integration times. Imaging wide DR natural scenes remains a major challenge for solid-state image sensors. Most voltage and current mode image sensors exhibit a saturating linear response with a DR typically limited to 60–70dB. However, light from natural scenes can span up to 140dB of DR, ranging from 1mlx up to 10klx and more. Dynamic range is not only a future issue for professional still photography and movies but will play an increasingly important role in all forms of machine vision and optical sensing for automotive, industrial and scientific applications.

The presented asynchronous time-based image sensor (ATIS) draws on the combination of PWM imaging, bio-inspired temporal contrast dynamic vision and event-based information encoding and data communication. The sensor tackles the above outlined limitations and, through its unique architecture and target-oriented implementation, achieves exceptional performance in terms of dynamic range, FPN, temporal resolution, gray-level resolution and data reduction. This paper briefly describes the idea behind the sensor concept, discusses selected design and implementation issues, and summarizes the most important results from chip laboratory tests and application-oriented characterization.

II. Imager Concept

The adverse effects of data redundancy, common to all frame-based image acquisition techniques, can be tackled in several different ways. The biggest conceivable gain however is achieved by simply not recording the redundant data in the first place, such saving power in the acquisition and reducing energy, bandwidth/memory requirements, and computing power in transmission and post-processing. For doing so, and achieving complete temporal redundancy suppression, a fundamental solution is an array of fully autonomous pixels, containing a change detector and a photo-measurement device (Fig. 1a). The change detector initiates the measurement of a new exposure/gray-level value only after a 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 and asynchronously requests access to an (ideally asynchronous and arbitrated) output channel

978-1-4244-5309-2/10/$26.00 ©2010 IEEE

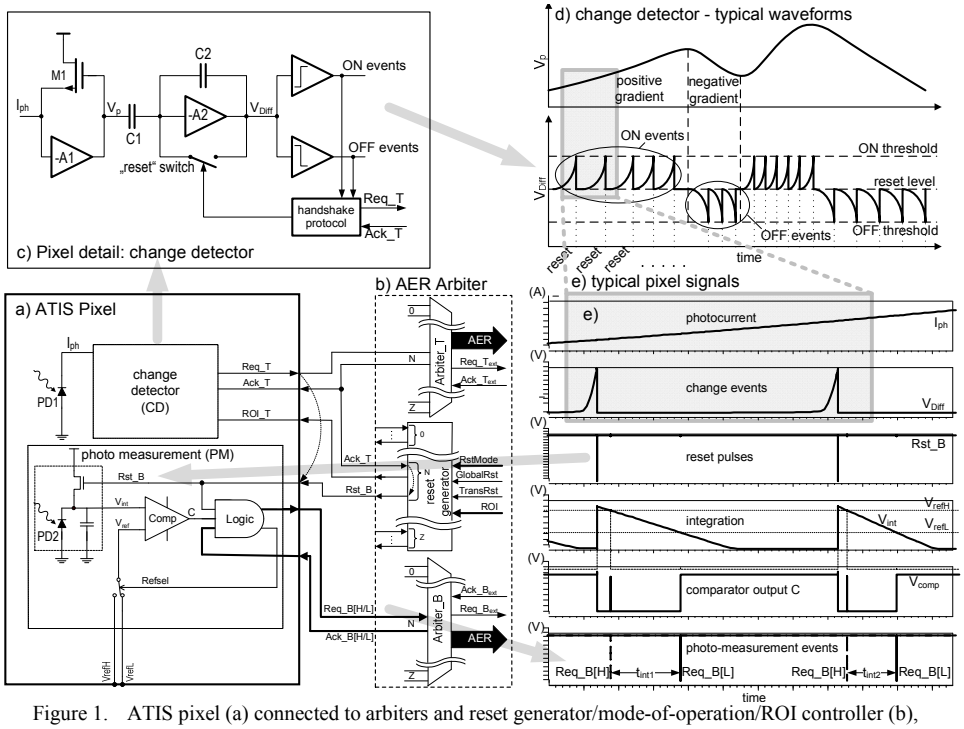

Figure 1. ATIS pixel (a) connected to arbiters and reset generator/mode-of-operation/ROI controller (b), change detector details (c) with typical signals (d), and exemplary pixel signals (e).

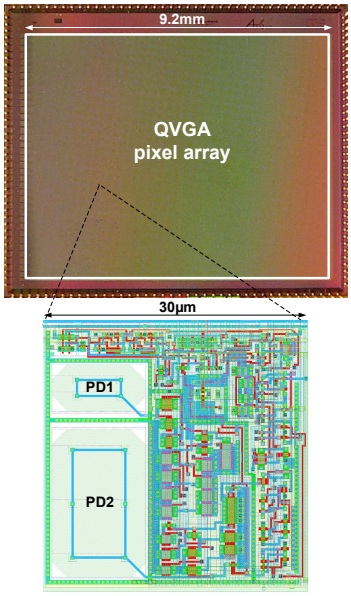

Figure 2. ATIS chip micrograph and pixel layout. Two separate photodiodes – for continuous time operation of the change detector and integrating photo-measurement – are used in each pixel.

(Fig. 1b) only when it has a new illumination value to communicate; 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.

As the 'change detector' (CD) a fast continuous-time logarithmic photoreceptor with asynchronous event-driven signal processing is used (Fig. 1c). The circuit has originally been developed for temporal contrast dynamic vision sensors [2][3][4] and combines an active, continuous-time, logarithmic photo-frontend (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 (Fig. 1d). 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.

For the 'photo-measurement' (PM) device, a time-based PWM imaging technique has been chosen. In time-based or pulse modulation imaging the incident light intensity is not encoded in amounts of charge, voltage, or current but in the timing of pulses or pulse edges. This approach automatically optimizes the integration time separately for each pixel instead of imposing a fixed integration time for the entire array, allowing for exceptionally high dynamic range (DR) and improved signal-to-noise-ratio (SNR). Particularly DR is no longer limited by the power supply rails as in conventional CMOS APS, such providing a relative immunity to the supply voltage scaling of modern CMOS technologies. The photodiode ('PD2') is reset and subsequently discharged by the photo-current. An in-pixel comparator ('Comp') triggers a digital pulse signal when a reference voltage is reached. The resulting integration time t_{int} is inversely proportional to the photo-current (Fig. 1e). Time-domain true correlated double sampling (TCDS) [5] based on two adjustable integration thresholds ($V_{refH/L}$) and intra-pixel state logic ('Logic') completes the pixel circuit (Fig. 1a).

The asynchronous change detector and the time-based photo-measurement approach harmonize remarkably well, mainly for two reasons: On the one hand because both reach DR of >120dB - the first is able to detect small relative change over the full range, the latter is able to resolve the associated gray-levels independently of the initial light intensity. Secondly, because both circuits operate event-based, namely the events of detecting brightness changes and the events of pixel integration voltages reaching thresholds. Consequently an asynchronous, event-based communication scheme (AER [6]) was used in order to provide efficient allocation of the transmission channel bandwidth to the active pixels.

The presented 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 gray-level encoding events independently and in parallel. 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 [7], or column-parallel relay readout. 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. The pixel comparator is equipped with idle-mode static current suppression to minimize pixel power consumption [8]. The chip has been implemented in a standard 0.18µm 1P6M mixed-mode/RF CMOS process. The pixel pitch is 30µm, the chip size 9.9mm × 8.2mm (Fig. 2).

III. SENSOR CHARACTERIZATION

A. Change Detector – Event Latency

The latency of the pixel 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. This variable delay time is, at low illumination, limited by the photoreceptor bandwidth and is proportional to reciprocal illumination. At high light levels, the photoreceptor is fast and the latency is dominated by the speed of A2 and the comparators, approaching a constant level with regard to illumination. For speed-optimized bias settings, the pixel latency goes below 10µs for illumination above ~30Lx and approaches 3µs for bright scenes (Fig. 3).

B. Change Detector – Contrast Sensitivity

Fig. 4 shows 'S'-shaped curves illustrating the pixel`s contrast sensitivity. Plotted is the spike probability as a function of temporal contrast given in % of $(I_{ph2} - I_{ph1})/I_{ph1}$ where I_{ph1} and I_{ph2} are the photocurrents before and after a change in pixel illumination. The 50%-'spike probability' point in the shown region of highest contrast sensitivity – between 1Lx and 100Lx – is constant at around 13% contrast for the chosen operating point settings. The slope of the S-curve is proportional to the initial ('background') illumination. For increasing illumination, the slope steepens further but the 50%-point starts drifting towards higher stimulus contrasts.

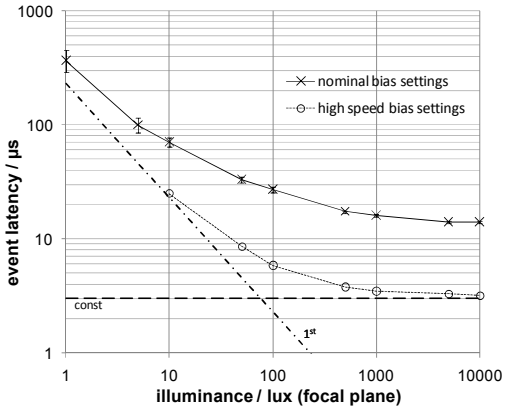

Figure 3. Pixel latency vs. illumination for two change detector circuit operating points.

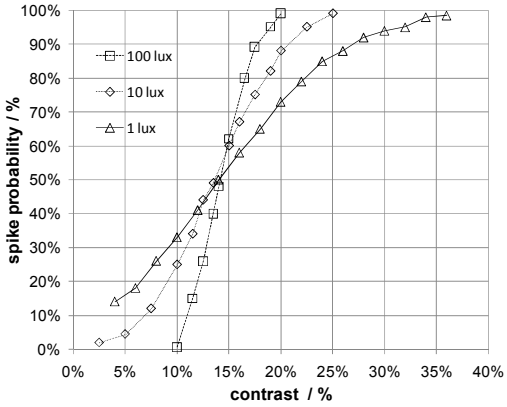

Figure 4. Contrast sensitivity: Spike probability vs. stimulus contrast for three magnitudes of illumination. The 50% point is constant, the slope of the S-curve depends on 'background' illumination.

C. Photomeasurement Basic Parameters

The charge capacity of the integration node depends on operation voltages and approaches ~450000 electrons at maximum (2V) integration swing ΔV_{th}. Sense node capacitance is 36fF, yielding a conversion gain of 4.4µV/e^-. Photodiode dark current has been measured at ~3fA, dark current shot noise is 470e^- corresponding to 2.1mV r.m.s. for an integration swing of 1V.

D. PWM Imaging – Transfer Function

Fig. 5 plots measured integration times for integration swings ΔV_{th} between 0.5V and 2V as a function of pixel illumination. The theoretically asserted 1/x-relation is accurately satisfied. Integration times range from e.g. 10ms @ 10Lx to 10µs @ 10kLx for an integration swing of 500mV.

E. Image Sensor Signal-to-Noise Ratio

In Fig. 6 measured imager SNR as a function of light intensity is shown. SNR is >56dB at an integration swing ΔV_{th} of 2V and light levels above 10Lx. Standard 8bit gray-level resolution (48dB) is achieved for very low illuminations and small integration swings. For ΔV_{th} = 100mV and 10Lx, SNR is still at 42dB allowing 7bit-resolution imaging at very short integration times (<2ms @ 10Lx) allowing for 500fps equivalent temporal resolution imaging and video at low-light conditions.

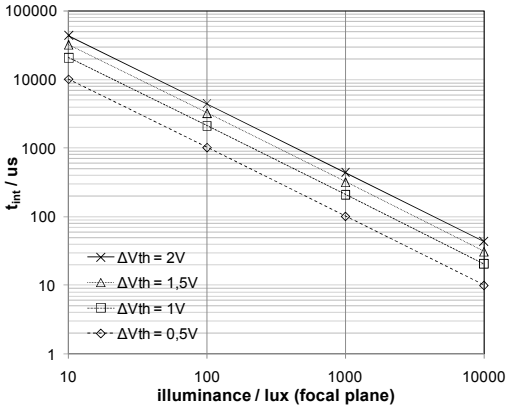

Figure 5. PWM integration time vs. lux transfer function for different values of integration swing ΔV_{th}.

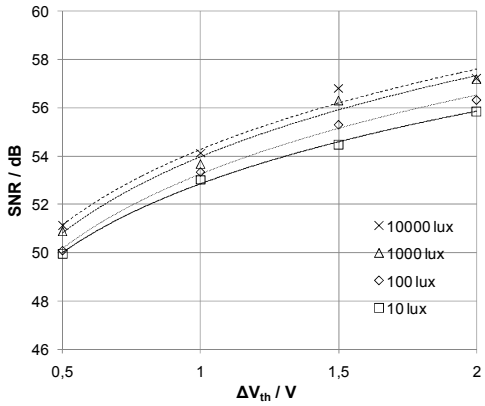

Figure 6. Measured SNR as functions of integration swing ΔV_{th} and light intensity. SNR is >56dB for an integration swing of 2V for light levels above 10Lx. For ΔV_{th} = 100mV and 10Lx SNR is still at 42.3dB

Figure 7. High DR imaging - three scalings of the same exposure image data and composite image.

IV. Results and Examples

The usable sensor DR shows a trade-off with the temporal resolution desired to adequately 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. A static scene DR of 143dB with a maximum integration time of 4.5s has been measured. (high-DR example scene Fig. 7). Keeping DR constant, the long integration times can be reduced by a factor of about 20 (to ~4.4 fps equivalent) by using the time between change event and upper TCDS threshold for determining the gray-levels in the dark parts of the scene, effectively implementing a multiple-threshold scheme. Trading-off SNR (reduces to ~42dB) and sacrificing TCDS is limited to the dark pixels in the image. For a temporal resolution corresponding to 30fps, DR is 125dB.

Fig. 8 shows a typical surveillance scene generating a 2.5k to 50k events/s @ 18bit/event continuous-time video stream. The actual event rate depends on instantaneous scene activity. Comparing corresponding bit rates – 45k to 900k bit/s – to the raw data rate of a QVGA 8bit gray-level sensor at 30fps of 18Mbit/s demonstrates lossless video compression factors between 20 and 400 for this example scene.

Figure 8. Traffic scene generating between 2.5k and 50k events/s, depending on scene activity. The ideally lossless video compression factor w.r.t. raw data from a QVGA 30fps 8bit gray-level sensor is 20 – 400.

TABLE I: Summary table of sensor characteristics.

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	30µm × 30µm
Pixel complexity	7T, 3C, 2PD
Fill factor	30% (20% photo measurement, 10% change detector)
Integration swing ΔV_{th}	100mV to 2.3 V (adjustable)
SNR typ.	>56dB (9.3bit) @ ΔV_{th} = 2V, >10Lx
SNR low	42.3dB (7bit) @ $\Delta V_{th\ min}$ (100mV), 10Lx
t_{int} @ $\Delta V_{th\ min}$ (100mV)	2ms @ 10Lx (500 fps equivalent temporal resolution)
Temporal resolution PM	500 fps equ. (@ 10Lx), 50kfps equ. (@ 1000Lx)
Temporal resolution CD	100kfps equ. (@ > 100Lx)
DR (static)	143dB
DR (30fps equivalent)	125dB
PRNU / FPN	<0.25% @ 10Lx (with TCDS)
Sense Node Cap	36fF
Conversion gain	$4.4\mu V/e^-$
Dark current	1.6nA/cm^2 (@ 25°C)
Power consumption	50mW (static), 175mW (high activity)
Readout format	Asynchronous address-events (AER), 2 × 18bit-parallel

V. Conclusion

A novel frame-free, high DR, low-power vision and image sensor with pixel-level data 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 optimal lossless pixel-level video compression through temporal redundancy suppression. Familiar deficiencies of time-based imagers have been remedied (a) using a novel time-domain correlated double sampling technique and (b) by realizing illumination-dependent readout load spreading. Intra-scene DRs of 143dB static and 125dB @ t_{int} < 30ms have been achieved, in line with or better than recent high-DR developments reported e.g. in [9][10]. Application areas are e.g. high-speed/high temporal-resolution dynamic machine vision, low-data rate video for wireless or TCP based applications and high DR, high quality, high-temporal resolution imaging and video for e.g. scientific applications.

References

[1] Ward, G., "The Hopeful Future of High Dynamic Range Imaging", 2007 SID International Symposium, 22-25 May 2007.

[2] Lichtsteiner, P.; Posch, C.; Delbruck, T., "A 128 × 128 120dB 30mW asynchronous vision sensor that responds to relative intensity change," *ISSCC, 2006, Dig. of Tech. Papers*, pp. 2060- 2069, Feb. 6-9, 2006.

[3] Lichtsteiner, P.; Posch, C.; Delbruck, T., "A 128× 128 120 dB 15 µs Latency Asynchronous Temporal Contrast Vision Sensor," *Solid-State Circuits, IEEE Journal of* , vol.43, no.2, pp.566-576, Feb. 2008.

[4] Posch, C.; Hofstätter, M.; Matolin, D.; et. al., "A dual-line optical transient sensor with on-chip precision time-stamp generation," *ISSCC, 2007, Dig. of Tech. Papers*, pp. 500-501, Feb. 11-15, 2007.

[5] Matolin, D.; Posch, C.; Wohlgenannt, R., "True correlated double sampling and comparator design for time-based image sensors," *Circuits and Systems, ISCAS 2009*, pp.1269-1272, 24-27 May 2009.

[6] K. Boahen, “A burst-mode word-serial address-event link-I transmitter design”, *Trans. Circuits and Systems I*, vol. 51, pp. 1269-1280, 2004.

[7] Xiaochuan Guo; Xin Qi; Harris, J.G., "A Time-to-First-Spike CMOS Image Sensor," *Sensors Journal*, vol.7, no.8, pp.1165-1175, 2007.

[8] Matolin, D.; Posch, C.; Wohlgenannt, R., "Area and Power Reduction Techniques for Time-based Image Sensor Pixel Design", 21th International Conference on Microelectronics (ICM 2009). Dec. 2009.

[9] Yamada, T., et al., "A 140dB-Dynamic-Range MOS Image Sensor with In-Pixel Multiple-Exposure Synthesis," *ISSCC, 2008, Digest of Technical Papers*, pp.50-594, 3-7 Feb. 2008

[10] Ruedi, P.-F., et al., "An SoC combining a 132dB QVGA pixel array and a 32b DSP/MCU processor for vision applications," *ISSCC 2009, Dig. of Tech. Papers*, pp.46-47,47a, 8-12 Feb. 2009.