

A 132 by 104 10 μ m-Pixel 250 μ W 1kefps Dynamic Vision Sensor with Pixel-Parallel Noise and Spatial Redundancy Suppression

Chenghan Li¹, Luca Longinotti¹, Federico Corradi², Tobi Delbruck³

¹iniVation AG, ²iniLabs GmbH, ³INI UZHÐ, Zurich, Switzerland, email: chenghan.li@inivation.com

Abstract

This paper reports a 132 by 104 dynamic vision sensor (DVS) with 10 μ m pixel in a 65nm logic process and a synchronous address-event representation (SAER) readout capable of 180Meps throughput. The SAER architecture allows adjustable event frame rate control and supports pre-readout pixel-parallel noise and spatial redundancy suppression. The chip consumes 250 μ W with 100keps running at 1k event frames per second (efps), 3-5 times more power efficient than the prior art using normalized power metrics. The chip is aimed for low power IoT and real-time high-speed smart vision applications.

Keywords: dynamic vision sensor, event-based, neuromorphic, IoT

Introduction

Recent dynamic vision sensor (DVS) research focuses on increasing spatial resolution, which desires smaller pixels and higher information throughput. Earlier up-to-QVGA DVS with down to 18.5 μ m pixel achieved up to 50M events per second (eps) readout speed [1-4]. A recent VGA DVS with 9 μ m pixel achieved 300Meps peak readout speed [5]. However, all prior DVS relied on arbitration in at least 1 dimension of the pixel array, hence were susceptible to constantly-requesting “hot” pixels. Furthermore, as DVS pixel size shrinks, it becomes more susceptible to shot noise and junction leakage, which causes information-less noise events. No prior DVS could remove noise events before readout, which renders the effective information throughput lower than the claimed peak readout speed.

This paper reports a 132 by 104 DVS with 10 μ m pixel and a synchronous address-event representation (SAER) readout capable of 180Meps throughput in a 65nm logic process. The SAER architecture allows adjustable event frame rate control, eliminating “hot” pixels, and supports pre-readout pixel-parallel noise and spatial redundancy suppression. Furthermore, by optimizing with a lower 1.2V power supply voltage, the chip consumes 250 μ W with 100keps running at 1k event frames per second (efps), 3-5 times more power efficient than the prior art using normalized power metrics.

Design

Figure 1 shows the schematic of a pixel and a group of 2 by 2 (2 rows by 2 columns) pixels. In comparison to prior DVS works [1-6], the handshake logic is replaced by the Digital Domain circuitry. This circuitry allows the whole pixel array to be controlled synchronously by *SAMPLE* and *RESTART*. A *SAMPLE* pulse captures an event frame into each pixel Event Memory. Following the *SAMPLE* pulse, the event frame readout starts, and a *RESTART* pulse puts the pixels with non-empty Event Memory (*ME=1*) back to operation. The *SAMPLE* pulse frequency determines the event frame rate.

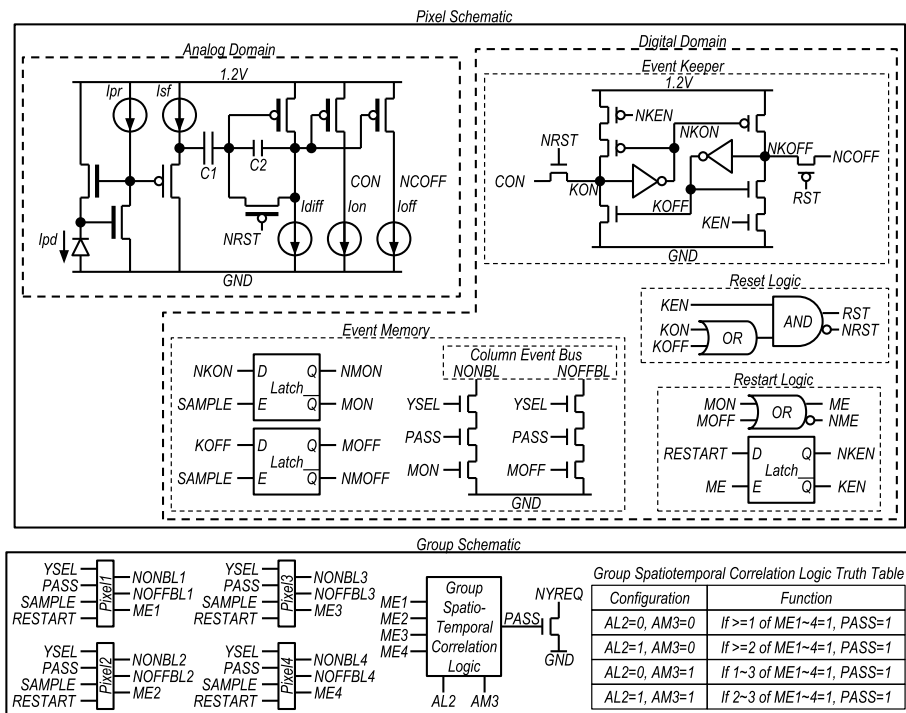


Figure 1 Schematic of a pixel and a group of 2 by 2 pixels.

Every group of 2 by 2 pixels shares a Group Spatiotemporal Correlation Logic (GSCL). After a *SAMPLE* pulse, the GSCL reads the Event Memory of all pixels in its group and decides via *PASS* whether to send the stored events using a programmable combination of 2 criteria: AL2 (at-least-2) - if there are less than 2 events stored in the group, assert *PASS*=0 to suppress noise events because noise events are usually spatiotemporally uncorrelated; and AM3 (at-most-3) - if there are more than 3 events stored in the group, assert *PASS*=0 to suppress spatially redundant events because a contiguous area of highly spatiotemporally correlated events contains little spatial information. To save space, the GSCL does not differentiate between ON and OFF events.

Figure 2 shows the readout architecture. Each column of groups shares a Column Spatiotemporal Correlation Logic (CSCL). The CSCL further filters the events during readout using the same principle as the GSCL, but with differentiation between ON and OFF events.

Similar to a token-ring scheme [6], each X/Y Scan Chain services columns/rows with $XREQ/YREQ=1$ systematically, while skipping columns/rows with $XREQ/YREQ=0$. Different from [6], each scan chain propagates a rising edge signal as the authorization signal and is clocked by a signal $XCLK/YCLK$ generated by a host independent of the propagation status of the authorization signal, eliminating synchronization delays.

Figure 3 shows the schematic of a scan chain segment. If $REQ=1$, the authorization signal propagates via the clocked Service Path. If $REQ=0$, the authorization signal propagates via the unclocked Skip Path. Depending on the numbers of scan chain segments to skip, the authorization signal may propagate on the Skip Path for longer than a CLK period and may violate CLK timing constraint when it transits to the Service Path. To prevent failure when timing constraint violation occurs, switching threshold disparities are implemented in the scan chain segment circuitry using transistors with different thresholds. Specifically, to prevent omission of a row/column that should be serviced, the *PRESENT* input to the *Rising Edge Detector* has a lower (closer to GND) switching threshold than the *PRESENT* input to the *Past DFF*; to prevent more than 1 row/column being serviced at the same time, the *PREV* input to the *Present DFF* and the *Skip Path* have a higher (further from GND) switching threshold than the *PRESENT* input to the *Past DFF*.

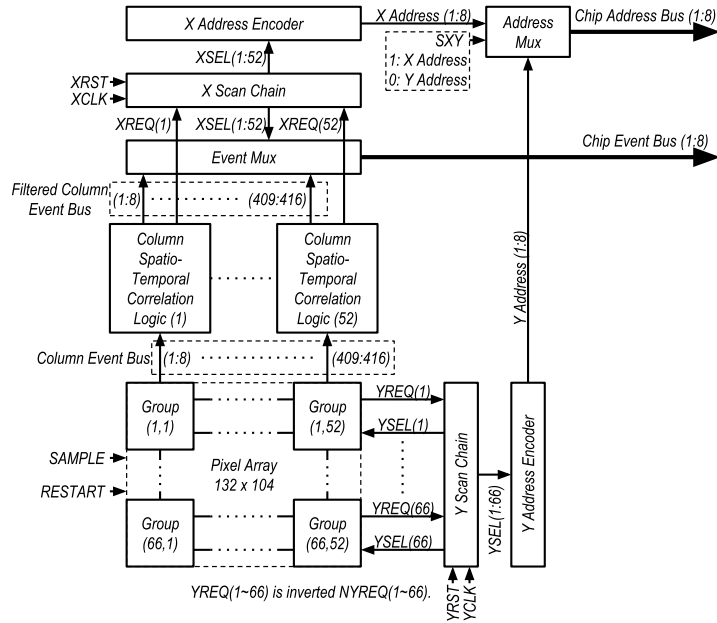
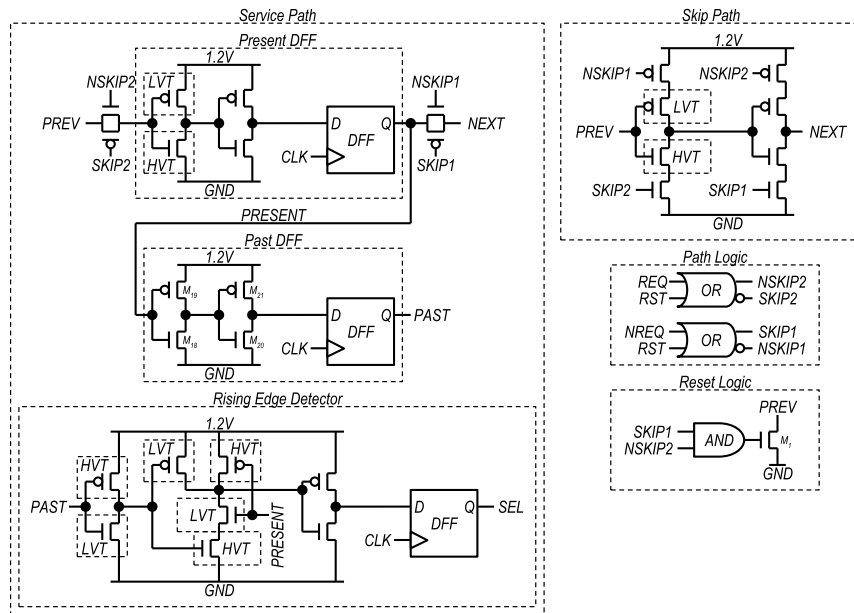


Figure 2 Block diagram of the readout architecture.



$REQ=X/YREQ$, $CLK=X/YCLK$, $SEL=X/YSEL$, $NREQ$ is the inverted REQ , HVT =high threshold, LVT =low threshold, no label=standard. When $RST=0$: If $REQ=1$, $NREQ=0$, then $SKIP1/2=0$, $NSKIP1/2=1$, service path; if $REQ=0$, $NREQ=1$, then $SKIP1/2=1$, $NSKIP1/2=0$, skip path. When $RST=1$: $SKIP1=1$, $NSKIP1=0$, $SKIP2=0$, $NSKIP2=1$, $PREV$ is reset to 0, 1 CLK is needed to reset $PRESENT$ to 0.

Figure 3 Schematic of a scan chain segment.

to prevent more than 1 row/column being serviced at the same time, the *PREV* input to the *Present DFF* and the *Skip Path* have a higher (further from GND) switching threshold than the *PRESENT* input to the *Past DFF*.

Figure 4 shows a timing diagram of the control signals and the chip output. During readout, a 0 value on the Chip Address Bus and the Chip Event Bus indicates that the authorization signal of the X/Y Scan Chain is propagating on the Skip Path. Each X/Y Scan Chain contains an additional end segment, whose end address X/YEND indicates the complete readout of a row or the whole of an event frame. Therefore, the size of an event frame depends on the number of events to readout.

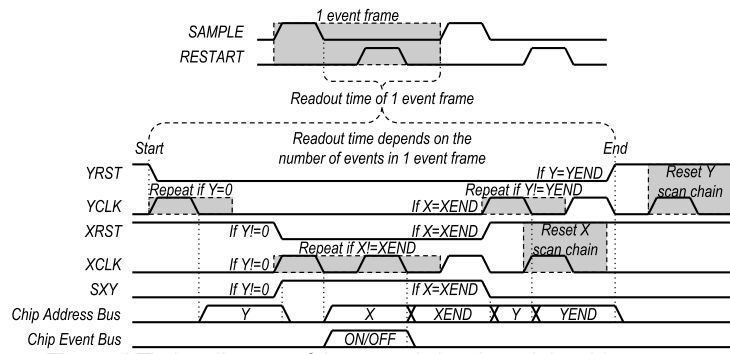


Figure 4 Timing diagram of the control signals and the chip output.

Results

The chip was implemented in a 65nm 1P9M logic process using n-well photodiodes. **Table 1** compares specifications and measured performance with prior work. The chip consumes 250μW with 100keps running at 1kefps and is 3-5x more power efficient than the prior art using the normalized power metrics. The maximum event rate of 180Meps is limited by the host-generated 50MHz clock (the chip supports up to 100MHz). The worst-case readout efficiency (events/clock) is 3x higher than the prior art. **Figure 5** demonstrates the effects of the on-chip noise and spatial redundancy suppression. The AL2 noise suppression reduced 40% of events in the example event frame, achieving similar results as post-readout noise suppression algorithms. The AL2+AM3 noise and spatial redundancy suppression reduced 87% of events in the example event frames, eliminating most of the redundant events caused by the flicker. **Figure 6** shows the die photo and a close-up of 2x2 pixels. The GSCL/CSCL event processing architecture could also support rudimentary edge and orientation detection. The chip is aimed for low power IoT and real-time high-speed smart vision applications.

References

- [1] P. Lichtsteiner, et al., ISSCC, pp. 2060-2069, 2006.
- [2] T. Serrano-Gotarredona, et al., JSSC, vol. 48, no. 3, pp. 827-838, 2013.
- [3] C. Brandli, et al., JSSC, vol. 49, no. 10, pp. 2333-2341, 2014.
- [4] C. Posch, et al., ISSCC, pp. 400-401, 2010.
- [5] B. Son, et al., ISSCC, pp. 66-67, 2017.
- [6] N. Imam, R. Manohar, ASYNC, pp. 99-108, 2011.

Table 1 Specifications and performance comparison.

	This Work	[5]	[3]	[2]	[1]	
Technology	65nm 1P9M	90nm 1P9M BSI	0.18μm 1P6M	0.35μm 2P4M	0.35μm 2P4M	
Resolution	132x104	640x480	240x180	128x128	128x128	
Chip Size (mm ²)	2x2	8x5.8	5x5	4.9x4.9	6.3x6	
Pixel Size (μm ²)	10x10	9x9	18.5x18.5	31x30	40x40	
Fill Factor (%)	20	-	22	10.5	8.1	
Power Supply (V)	1.2	2.8 & 1.2	3.3 & 1.8	3.3	3.3	
Power (mW)	High Activity	4.9@180Meps a	50@300Meps	14	-	24
	Low Activity	0.25@100keps a	27@100keps	5	4	-
Normalized	Dynamic (pJ/event)	26	77	-	-	-
	Static (nW/pixel)	18	88	-	-	-
Max Event Rate (Meps)	180	300	12	20	2	
Readout Efficiency (event/clock)	best: 4, worst: 0.25 c	best: 6.7, worst: 0.077 d	-	-	-	

a The power includes bias generator and IO power and was measured using identical bias configuration.

b The normalized power is calculated as:

Dynamic Energy = $(P_H - P_L) / (R_H - R_L)$, Static Power = $(P_L - R_L \cdot \text{Dynamic Energy}) / N_p$, where P_H is power at high activity, P_L is power at low activity, R_H is event rate at high activity, R_L is event rate at low activity, N_p is total number of pixels.

c The best case is when all events are in the same row of groups, the worst case is when all events are in different rows of groups, where a minimum of 4 clocks per row is needed.

d The best case is when all events are in the same column, the worst case is when all events are in different columns, where a minimum of 13 clocks per column is needed.

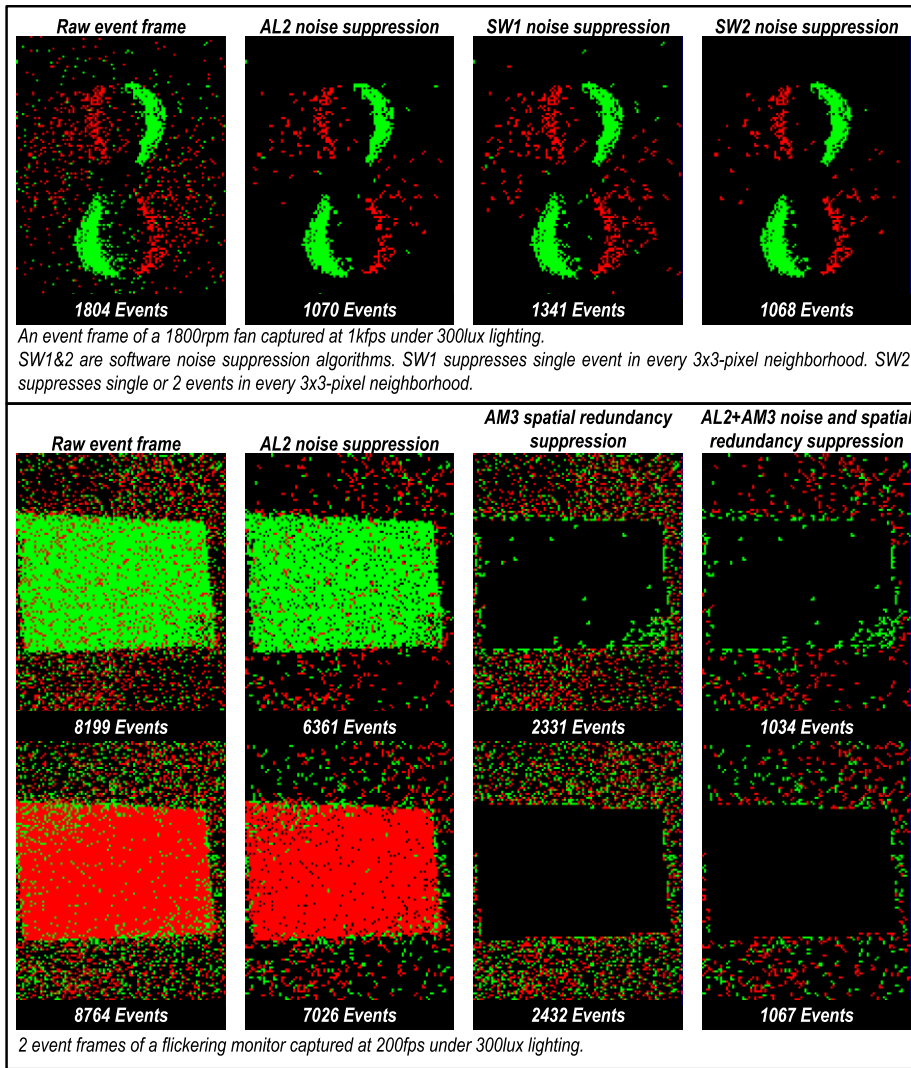


Figure 5 Demonstration of on-chip noise spatial redundancy suppression. Green dots are ON events. Red dots are OFF events.

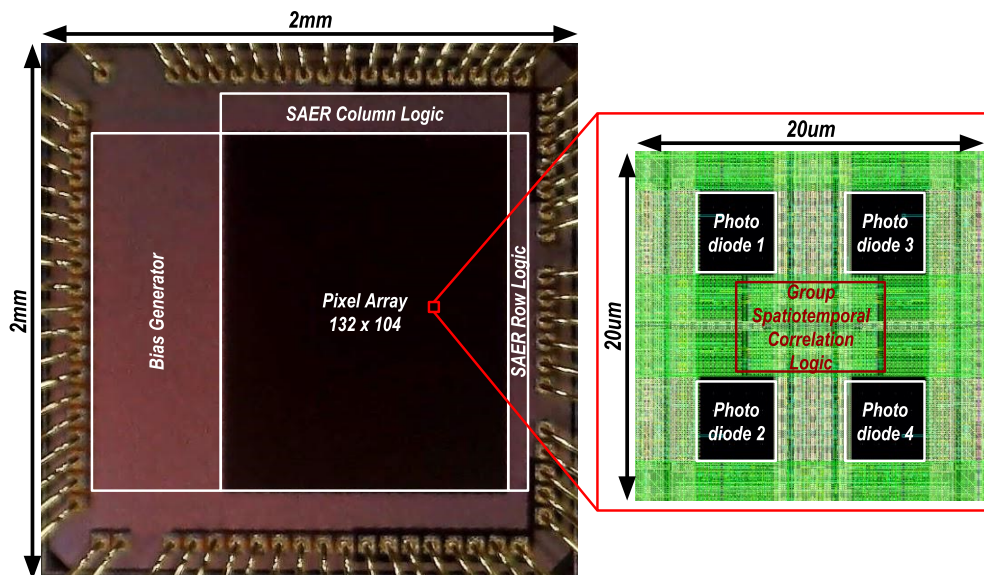


Figure 6 Die photo and a close-up of 2 by 2 pixels.