



February 2003

A biomorphic digital image sensor

Eugenio Culurciello
Johns Hopkins University

Ralph Etienne-Cummings
Johns Hopkins University

Kwabena A. Boahen
University of Pennsylvania, boahen@seas.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/be_papers

Recommended Citation

Culurciello, E., Etienne-Cummings, R., & Boahen, K. A. (2003). A biomorphic digital image sensor. Retrieved from https://repository.upenn.edu/be_papers/20

Copyright 2003 IEEE. Reprinted from *IEEE Journal of Solid-State Circuits*, Volume 38, Issue 2, February 2003, pages 281-294.

Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=26391&puNumber=4>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/be_papers/20
For more information, please contact repository@pobox.upenn.edu.

A biomorphic digital image sensor

Abstract

An arbitrated address-event imager has been designed and fabricated in a 0.6- μm CMOS process. The imager is composed of 80 x 60 pixels of 32 x 30 μm . The value of the light intensity collected by each photosensitive element is inversely proportional to the pixel's interspike time interval. The readout of each spike is initiated by the individual pixel; therefore, the available output bandwidth is allocated according to pixel output demand. This encoding of light intensities favors brighter pixels, equalizes the number of integrated photons across light intensity, and minimizes power consumption. Tests conducted on the imager showed a large output dynamic range of 180 dB (under bright local illumination) for an individual pixel. The array, on the other hand, produced a dynamic range of 120 dB (under uniform bright illumination and when no lower bound was placed on the update rate per pixel). The dynamic range is 48.9 dB value at 30-pixel updates/s. Power consumption is 3.4 mW in uniform indoor light and a mean event rate of 200 kHz, which updates each pixel 41.6 times per second. The imager is capable of updating each pixel 8.3K times per second (under bright local illumination).

Keywords

arbitrated, address event, digital image sensor, high dynamic range, low-power imager

Comments

Copyright 2003 IEEE. Reprinted from *IEEE Journal of Solid-State Circuits*, Volume 38, Issue 2, February 2003, pages 281-294.

Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=26391&puNumber=4>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

A Biomimetic Digital Image Sensor

Eugenio Culurciello, Ralph Etienne-Cummings, and Kwabena A. Boahen

Abstract—An arbitrated address-event imager has been designed and fabricated in a 0.6- μm CMOS process. The imager is composed of 80×60 pixels of $32 \times 30 \mu\text{m}$. The value of the light intensity collected by each photosensitive element is inversely proportional to the pixel's interspike time interval. The readout of each spike is initiated by the individual pixel; therefore, the available output bandwidth is allocated according to pixel output demand. This encoding of light intensities favors brighter pixels, equalizes the number of integrated photons across light intensity, and minimizes power consumption. Tests conducted on the imager showed a large output dynamic range of 180 dB (under bright local illumination) for an individual pixel. The array, on the other hand, produced a dynamic range of 120 dB (under uniform bright illumination and when no lower bound was placed on the update rate per pixel). The dynamic range is 48.9 dB value at 30-pixel updates/s. Power consumption is 3.4 mW in uniform indoor light and a mean event rate of 200 kHz, which updates each pixel 41.6 times per second. The imager is capable of updating each pixel 8.3K times per second (under bright local illumination).

Index Terms—Arbitrated, address event, digital image sensor, high dynamic range, low-power imager.

I. INTRODUCTION

CONVENTIONAL cameras produce images by scanning the photosensitive pixels in a sequential (raster) format, functionally dividing the output bandwidth equally among all pixels. The sequential scan requires that signal processing performed on the video stream be completed within one pixel readout time. This requirement can be difficult to fulfill for large ($>256 \times 256$) or fast (>100 frames per second) imaging arrays. To circumvent this sequential bottleneck, in the late 1980s researchers demonstrated a new imaging paradigm that mimicked the human retina with silicon integrated circuits [1]. The main advantage of the silicon retina was its highly parallel computational nature, which allowed high-speed pixel-parallel image processing at the focal plane. Mahowald and Mead's silicon retina provided the first glimpse of the great potential of CMOS integrated circuits technology for imaging [1]. This potential, however, has still not been fully realized today. It should be noted that CMOS imagers designed as substitutes for charge-coupled device (CCD) imagers have made significant

inroads into the commercial marketplace, yet the focal plane image-processing capabilities of the technology has not been fully exploited [2]. The early silicon retinas were doomed as an alternative imaging approach because the CMOS technology in the early 1990s was not mature enough to compete with the quality of CCD imagers. This is especially true when considering that the noise introduced by the photo detector, amplification circuits, and image processing (edge and motion detection) circuits are significantly higher than CCD imagers, although the latter do not provide any processing on the image plane. Furthermore, the silicon retina pixels were too large to realize high-resolution arrays at a reasonable yield per cost. Consequently, the idea of a silicon retina as a commercially viable imager was abandoned. Recently, the silicon retina concept has been resurrected because three-dimensional (3-D) integration techniques promise small footprints with pixel-parallel spatiotemporal image processing [3], [4]. However, we are still far from a commercial product in these technologies. The research on biologically inspired imagers and image processing chips in standard CMOS processes have continued over the past ten years [5]–[7]. The imager presented here continues the trend of “reverse engineering biology,” where the outcome is a silicon retina with focal-plane image processing/encoding, small pixel sizes, extremely high dynamic range, relatively low power consumption, and “photon-to-bits” phototransduction.

Conventional imagers integrate the photocurrent for a fixed time, usually dictated by the scanning period. Subsequently, the integrated voltage is output according to a raster scan. Here, we invert the process by integrating the photocurrent to a fixed voltage (threshold). When the threshold is crossed, a 1-b pulse (spike) is generated by the pixel. The magnitude of the photocurrent is represented as the interspike interval between two successive spikes. This interspike interval is inversely proportional to the intensity. Our system is also different from conventional methods because the readout of each spike is initiated by the pixel itself. That is, each pixel requests access to the output bus when the integration threshold has been crossed [8].

This biologically inspired readout method simultaneously favors brighter pixels, minimizes power consumption by remaining dormant until data is available, and offers pixel-parallel readout. In contrast, a serially scanned array allocates an equal portion of the bandwidth to all pixels independent of activity and continuously dissipates power because the scanner is always active. Here, brighter pixels are favored because their integration threshold is reached faster than darker pixels, i.e., the request–acknowledge–reset–integrate cycle operates at a higher frequency. Consequently, brighter pixels request the output bus more often than darker ones. Also, virtually no power is used by the pixel until an event is generated; therefore, low-intensity pixels consume little power. Furthermore,

Manuscript received January 22, 2002; revised August 15, 2002. The work of E. Culurciello was supported by the Defense Advanced Research Projects Agency under DARPA/ONR MURI N0014-95-1-0409. The work of R. Etienne-Cummings was supported by the National Science Foundation under CAREER Award 9896362. The work of K. A. Boahen was supported by a Whitaker Foundation Research Initiation Award.

E. Culurciello and R. Etienne-Cummings are with the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: euge@jhu.edu).

K. A. Boahen is with the Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA.

Digital Object Identifier 10.1109/JSSC.2002.807412

representing intensity in the temporal domain allows each pixel to represent a large dynamic range of outputs [11], [12]. The integration time is, in fact, not dictated by a regular scanning clock and, therefore, a pixel can use the whole bus bandwidth by itself or can abstain from the image forming process. This provides a simple and efficient way of obtaining dynamic range control, without the use of additional circuitry that varies the integration time of each pixel based on the light intensity [13]. Pixel-parallel automatic gain control is an inherent property of our time-domain imaging and readout scheme, which is called *address-event representation* (AER) [8]–[10], [14].

We will describe the AER architecture in Section II, the event or spike generation circuits in Section III, the spike communication circuits in Section IV, the imager operation and its analysis in Section V, and present results and discussion in Section VI and the conclusion in Section VII.

II. AER

The imager uses AER output format. The address-event (AE) communication channel is a model of the transmission of neural information in biological systems [14]. Information is presented at the output in the form of a sequence of pulses or spikes, where the interspike interval or the spike frequency encodes the analog value of the data being communicated. Encoding the data as a stream of digital pulses provides noise immunity by quantization and redundancy. The frequency-modulated signal can be reconstructed by integration or simply by counting the number of received events over a predetermined window of time. The imager presented here mimics the octopus' retina by converting the light intensity directly into a spike train [15]; most other biological retinas represent light intensity as an analog signal [16], [17].

The AER model trades the complexity in wiring of the biological systems for the processing speed of integrated circuits. Neurons in the human brain make up to 10^5 connections with their neighbors [16], [17], a prohibitive number for integrated circuits. Nevertheless, the latter are capable of handling communication cycles that are six orders of magnitude smaller than the interevent interval for a single neuron. Thus, it is possible to share this speed advantage among many cells and create a single communication channel to convey all the information between two neural populations. AER uses an asynchronous protocol for communication between different processing units [8]–[10].

As shown in Fig. 1, the information, divided into "events," is sent from a unique sender to a unique element in a receiving population. Events are generally in the form of a spike; therefore, only their address is the important data to reconstruction and the time of occurrence. The information packet is, therefore, the address of the spiking cell or transmitter. In the case of our imager, events are individual pixels reaching a threshold voltage and requesting the bus for communication with a receiver. As a result, the system represents light intensity on a pixel as a frequency-modulated sequence of addresses, where the time interval between identical addresses (pixels) is inversely proportional to the intensity. An AE system is generally composed of a multitude of basic cells or elements either transmitting, receiving, or transceiving data. Reconstruction of data necessi-

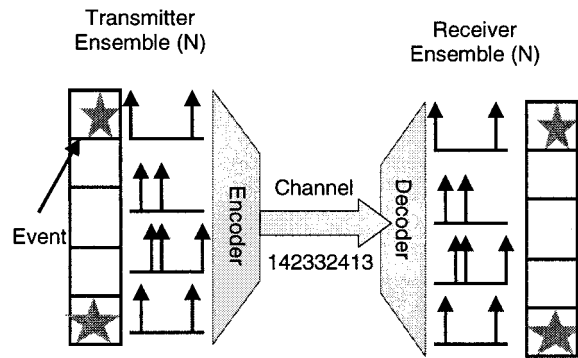


Fig. 1. AE system: A general-purpose protocol for the transmission of data from an array of senders to an array of receivers.

tates storage, since events must be counted or accumulated to reassume the form of intensity signals.

A few frequency-modulated and/or AE imaging systems have been previously reported, however, the one presented here is the first to combine a conventional active pixel sensor (APS) with a fully arbitrated AE system, to provide a high-resolution image with one of the best quality reported [2], [11], [12], [19], [20].

III. EVENT GENERATION

The key element in an address event imager is the spike generator circuit. This element, generally incorporated in the pixel cell, is responsible for requesting access to the output bus when a pixel has reached the integration threshold. Generally, a prototypical CMOS imager employs a photodiode as a photosensitive element. The relatively small photocurrent is integrated on a capacitor and subsequently read out. An AE imager will convert light into events by integrating photocurrent up to a fixed threshold. The integrated voltage changes very slowly if the light intensity is low. The event generator must convert this slow-changing voltage into a fast-changing signal in order to minimize the delay between the time when the threshold is passed and when the output bus access is requested. Furthermore, the fast transition also limits power consumption. Hence, the event generator is an important component of the AER imager and will be described in detail. After the pixel's request has been acknowledged, the pixel is reset and all accumulated charges on the integration capacitor are drained. The integration process is then immediately restarted. Notice that a natural ordering of the pixels' readout occurs that minimizes pixel request collisions. Collisions translate into temporal jitter, which degrade the image quality. Jitter due to arbitration will also be discussed in Section V-C.

A. Simple Inverter as Event Generator

The simplest event generator is a solitary inverter. The high inversion gain of a CMOS inverter is an immediate solution for implementing a threshold circuit with a binary output. Its gain is capable of amplifying the tiny slew rate of the input signal. On the other hand, its power consumption is proportional to the switching time, which, in turn, is proportional to the input signal slew rate.

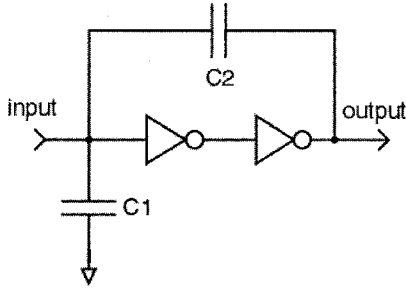


Fig. 2. Capacitive feedback in integrate and fire neurons.

In ambient lighting, the photosensor input slew rate is six orders of magnitude slower than typical digital signals (or 1 V/ms). This means that the input voltage remains in the high power consumption region of the inverter for a long time, creating a direct current path between the supplies. A simple inverter used as an event generator, in a $0.5\text{-}\mu\text{m}$ process and 3.3-V supply, consumes about 3.9 nJ ($15\text{ }\mu\text{W} \times 0.26\text{ ms}$). A typical digital inverter using minimum size transistors, in a $0.5\text{-}\mu\text{m}$ process and 3.3-V supply, consumes only about 0.06 pJ ($40\text{ }\mu\text{W} \times 3\text{ ns} \times 0.5$) per off-transition (rising input, falling output) and about 0.18 pJ ($120\text{ }\mu\text{W} \times 3\text{ ns} \times 0.5$) per on-transition (falling input, rising output). Therefore, the power consumption of the inverter as the event generator is about four to five orders of magnitude greater than that of a minimum-size inverter in a digital circuit. Clearly, a simple inverter is not a good candidate as an event generator for low-power imaging applications. To limit power consumption, a starved inverter can be used, where the output current is limited by a current source to a few nanoamperes. However, there is a severe impact on switching speed when this approach is taken, as will be evident in Section III-D.

B. Capacitive-Feedback Inverters as Event Generator

In order to decrease the power consumption of the event generator, it is necessary to increase its gain, at least in the vicinity of the threshold. A voltage feedback circuit employing capacitive feedback can speed up the transition and, therefore, limit the time spent in the high power consumption region (Fig. 2). The capacitive feedback multiplies the inverter ac gain by the feedback ratio $(C_1 + C_2)/C_2$ [23].

A further improvement is obtained by operating the capacitive feedback inverters with the MOSFETs in weak inversion. This improves power consumption significantly in ambient light conditions of 1 W/m^2 . The second inverter uses about $7\text{ }\mu\text{A}$ for only 7 ns to generate an output spike, but the first inverter remains for $4\text{ }\mu\text{s}$ in the high power consumption region because of the slow rising input. The pixel readout rate is, however, severely reduced when the event generator operates in subthreshold. While we receive some power consumption benefits from the capacitive-feedback circuit, those benefits are shadowed by the increased size (a large feedback capacitor is required) and lower readout rate of the pixel.

C. Current-Feedback Event Generator

The event generator used in the imager solves both the transition speed and power consumption problems with an

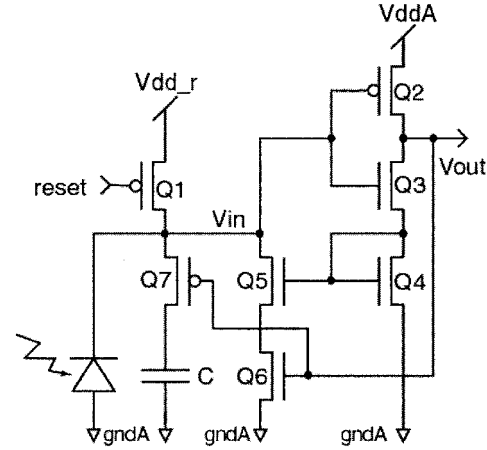


Fig. 3. Current-feedback event generator pixel.

elegant current positive feedback circuit. Power consumption and transition speed are closely related because CMOS digital circuits only consume power during switching. Hence, reducing the transition time will also reduce the power consumption. Our event generator has simultaneously a large gain, large bandwidth, and minute power consumption. This circuit can be used for various other applications where high speed and low power consumption are required. Fig. 3 shows the schematic of the pixel and the event generator. Photons collected by an n-type photodiode are integrated on a 0.1-pF capacitor to give a slew rate of 0.1 V/ms in typical indoor light (0.1 mW/cm^2). In dimmer conditions, the input slew rate can be much lower.

Event generation occurs as follows. Initially, the inverter input voltage V_{in} is high (after the reset pulse). Transistor $Q2$ is off and so is the feedback switch $Q6$. In addition, the inverter output voltage V_{out} is low. As the capacitor C is discharged by the photocurrent, V_{in} decreases and transistor $Q2$ begins conducting. Slightly before V_{in} reaches the threshold of $Q2$, a subthreshold current flows through the inverter and is fed back to the input, through transistors $Q4$ – $Q6$. Notice that V_{out} starts to rise before the feedback circuit is activated, which subsequently switches $Q6$ on and starts the current feedback. The mirror pair $Q4$ – $Q5$ is sized for current gain. The feedback current mirror operates in subthreshold initially, but increases exponentially as V_{in} decreases further. We approximate the start of the switching process as the value of V_{in} where the feedback current equals and surpasses the photocurrent. At this point, the V_{in} accelerates toward ground, V_{out} accelerates toward V_{dd} , and the switch transistor $Q7$ turns off, which disconnects the integration capacitor from V_{in} and causes V_{in} to accelerate further. Furthermore, as V_{in} plunges below the threshold voltage of $Q3$, it shuts off the feedback mirror, which cuts off the current in the $Q2$ – $Q4$ branch and causes V_{out} to accelerate further toward V_{dd} . As can be seen, the transition takes place just before the threshold voltage of $Q2$ is reached. The capacitance at the V_{in} node is suddenly decreased, and $Q3$ and $Q4$ cut off for a low-current yet high-speed circuit. This circuit is unique in this respect. Fig. 4 shows a SPICE simulation of the circuit operation. The upper traces plot the input and output voltage versus time. Note first the slow rise in the voltage, due to the photocurrent, then the sudden switch

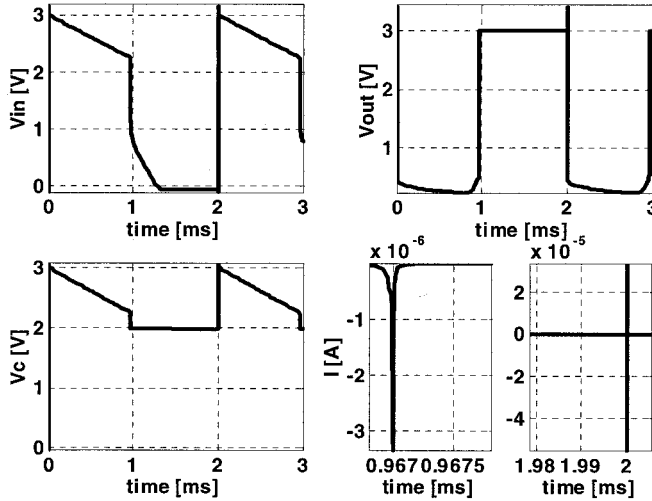


Fig. 4. SPICE simulation of the pixel's spike generator, V_{in} , V_{out} , V_c plots and current consumption during spike and reset.

as the feedback circuit comes into action. The lower traces show the voltage on the integrating capacitor and the current consumption during an event and reset.

Using the proposed circuit with positive current feedback, as shown in Fig. 3, we obtained a switch time of 8 ns (0.6- μ m CMOS process and input slew rate of 1 V/ms) while using only 0.043 pJ (SPICE simulation). In addition, for an APS photo-sensor, the majority of the pixel's power consumption occurs during reset. To reduce reset power, the integration capacitor is disconnected from the comparator when a request is generated. This is a very important feature because the capacitor is then reset from $\sim(V_{dd} - V_{tp})$ to V_{dd} instead of Gnd to V_{dd} (considering $V_{dd}A = V_{dd}x = V_{dd}$ from Fig. 3). During reset, a simulation of the pixel operation computed 3.88 pJ as power consumption.

D. Comparison Between Event Generators

To demonstrate the strength of the current-feedback event generator, we compared it to a simple inverter, a simple starved inverter, and a capacitive-feedback inverter. We used SPICE for the comparison, using AMI 0.5- μ m CMOS parameters from MOSIS. Tests were conducted on all four circuits to measure the total energy consumption and slew-rate gain by applying an input current to decrease V_{in} at different slew rates. Slew-rate gain is defined as the output slew rate divided by the input slew rate. The tests were conducted with a common power supply of 3 V and the input slew rate varied over the expected range of ambient lighting conditions for which the imager will be used. Other than the additional devices required to implement the four circuits, we kept the transistor sizes consistent. The capacitive-feedback inverters circuit used capacitors C_1 of 100 fF and C_2 of 5 fF, thus, the capacitive gain was 21. The output current in the starved inverter was limited to 1 nA so that its energy consumption approaches that of the current-feedback event generator.

As can be observed in Fig. 5, the event generator with current feedback greatly surpasses the performance of all the inverter-based event generators. In fact, its energy usage remains

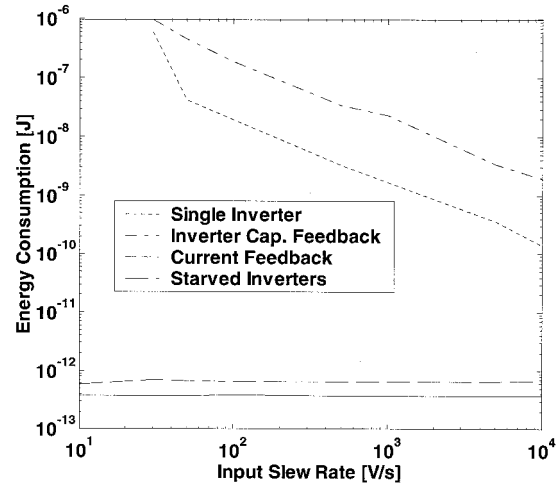


Fig. 5. Energy consumption per event versus input slew rate.

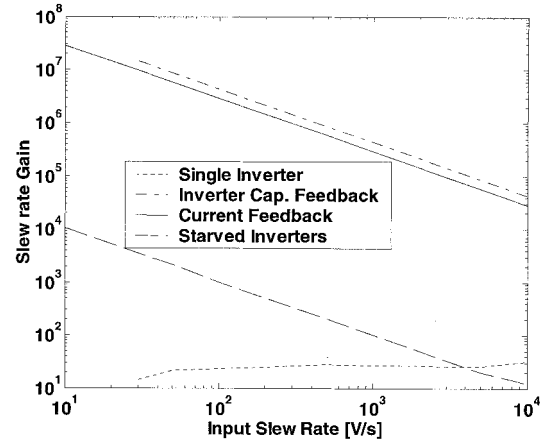


Fig. 6. Slew-rate gain versus input slew rate.

several orders of magnitude smaller than the competition, except for the starved inverter, whose design approaches the energy consumption of the current-feedback event generator. However, it will be soon proven that the starved inverter cannot match the proposed circuit in switching speed. Because the energy consumption is independent of the input slew rate in our event generator, the current-feedback circuit guarantees constant power consumption per cycle. For an array, the power consumption will be a linear function of light intensity, depending on only the integrate-request-acknowledge-reset cycle frequency of each pixel. The other circuits, in presence of low light or in the dark, with low input slew rates, would instead consume an even larger amount of energy.

Fig. 6 presents data on the slew-rate gain versus input slew rate. Again, observe that the current-feedback event generator is much faster than the starved inverter and the inverter circuits. On the other hand, it is slightly slower than the feedback inverters. We also observe that its switching speed is independent of the input slew rate because of the positive feedback. Once the switch begins, the feedback takes over and accelerates the discharge of the input node. In the other inverter circuits without feedback, the input slew rate is unchanged. The capacitive-feedback inverter also presents higher input slew rates; however,

it is still dependent on the input slew rate. The current-feedback event generator has a constant output slew rate of approximately 10^7 V/s, independently of the input slew rate. Being limited by the input signal, the inverter-based circuits are kept longer in the high power consumption region of the inverters and, therefore, consume more power per event. Note also that the performance of the current-feedback circuit is comparable to that of a minimum-size inverter with digital input, one of the most efficient and optimized switching circuits in today's microelectronics. The good performance in power consumption for the current-feedback event generator, shown in Fig. 5, is also a direct result of its fast switching characteristics.

Short-circuit current at the event generator's input inverter is the main source of power consumption because the input slew rate is low. Assuming a triangular pulse with peak I_{sc} and width Δt , the quantity $(1/2)V_{dd}I_{sc}\Delta t$ will be dissipated. Δt is the time the output voltage V_o takes to transition from Gnd to V_{dd} , which equals the time the input voltage V_{in} takes to change by V_{dd}/A_{inv} (A_{inv} is the inverter gain), assuming the inverter is not slew-rate limited or the short-circuit current will be negligible. Hence, with $A_{inv} \approx 10$, and the input slew rate, $dV_i/dt = I_i/C_i$ (I_i is input current; C_i is input capacitance), we can obtain $\Delta t = C_i V_{dd}/(A_{inv} I_i)$. Consequently, the energy E_{sc} dissipated by the short circuit is

$$E_{sc} = \frac{1}{2} C_i V_{dd}^2 \frac{I_{sc}}{A_{inv} I_i}.$$

Notice that E_{sc} exceeds the switching energy $E_d = 1/2 C_i V_{dd}^2$ when $I_i \ll I_{sc}$. As $I_i \approx 100$ pA while $I_{sc} \approx 100$ μ A in this imager, the short circuit dissipation could be a million times larger.

The only way to reduce short-circuit power consumption is to increase the input current I_i as by using positive feedback. In the capacitive-feedback event generator design, a fraction $C_{1,2}/(C_{1,2} + C_o)$ of the output current I_o is fed back ($C_{1,2}$ is the series capacitance of C_1 , C_2 , and C_o is the load capacitance). As $I_o = (C_o + C_{1,2}) A_{inv}^2 dV_i/dt$, assuming again that the inverters are not slew-rate limited, we obtain

$$I_{fb} = \frac{C_1 C_2}{(C_1 + C_2)^2} A_{inv}^2 I_i$$

once we express the input slew rate in terms of the input current I_i and the input capacitance C_i , and we substitute $C_{1,2} = C_1 C_2 / (C_1 + C_2)$ and $C_i = C_1 + C_2$. The capacitance terms attain a maximum of $1/4$ when $C_1 = C_2$. Hence, this design cannot reduce short-circuit power dissipation by more than a factor of $A_{inv}^2/4$, or about 25. In contrast, $I_{fb} = I_{sc}$ for the current-feedback event generator design, making its short-circuit dissipation comparable to the switching energy, thus, achieving a millionfold reduction in power.

IV. EVENT COMMUNICATION

After an event has been generated (see Section III-C), an additional AER infrastructure in the pixel is required to communicate the event to the output bus by means of the boundary arbitration circuitry. Fig. 7 shows a schematic caption of the pixel, where the right portion is the digital circuitry responsible for

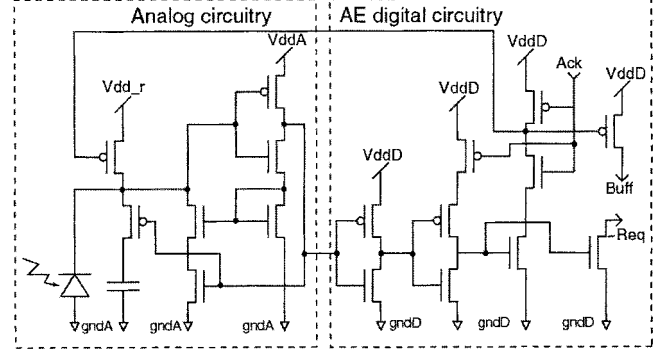


Fig. 7. Imager pixel schematic.

communicating the event to the outer array circuitry. This digital portion of the pixel generates a row request \sim Req. To provide robust noise immunity between the analog and digital portions of the pixel, the output of the event generator is buffered before passing it to a row-wise wired OR. The wired OR indicates that a pixel in that row has requested access to the output bus.

The second inverter in the buffer has an additional pMOS transistor controlled by the returning acknowledge Ack signal. The additional transistor blocks any other request that might arise if the Ack signal has not been previously reset (i.e., a communication cycle has been completed). Analogously, an additional nMOS in the Ack signal path prevents racing conditions by only acknowledging a pixel whose request has been allowed to reach the boundary circuits. Hence, a handshaking protocol is initiated by the pixel which requests the output bus, provided it has previously been acknowledged; also, the pixel acknowledges provided it has previously issued a request and gained access to the bus. This forms a four-phase handshaking sequence, which is also repeated at the row and column level. Fig. 8 illustrates the boundary arbitration circuitry for the communication of the event.

The boundary circuits are used to arbitrate between active pixels (i.e., pixels that have generated events). This arbitration is executed in two steps. First, a row arbitration tree selects one row from which at least one request has been generated. Next, the column arbitration tree selects and outputs the individual pixels within the row. When a row is selected, the entire row is copied into a buffer located above the array (Row Latch). This buffering step provides a pixel access speedup and improved parallelism by realizing a pipelined readout scheme. Simultaneously, the address of the row is also decoded and placed on the output bus Y . When a row request, i.e., the wired OR signal, is asserted, many active pixels may exist within the row. The signal Buff indicates which pixel in the row has issued a request. Once copied, the entire row is acknowledged/reset (signal Ack), and photon integration starts anew. Column arbitration is performed on the buffered row. The arbitration tree selects the active elements in the buffer and computes and outputs their X addresses before clearing the buffer. A new active row is obtained when the buffer is clear. Performing column arbitration on the buffered row also improves readout speed by eliminating the large capacitance associated with the column lines. This capacitance is encountered when arbitration is performed within the whole array. Fig. 8 shows the architecture of the row and

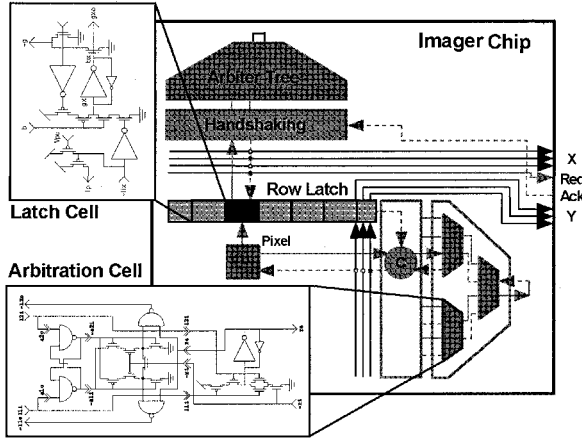


Fig. 8. Row and column arbitration architecture.

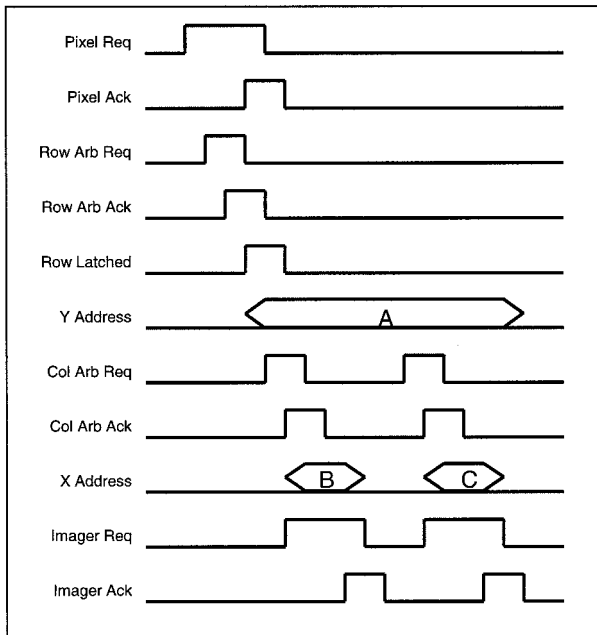


Fig. 9. Array arbitration timing diagram.

column arbitration circuits. Fig. 9 illustrates the signaling and the handshaking generated by the boundary arbitration circuitry of the imager array.

As a final note, the imager power consumption can be reduced even further by using more elaborated circuits that eliminate the wired OR $\sim\text{Req}$ and Buff lines. The AE architecture employs pseudo-CMOS logic, which can be substituted with fully static or dynamic logic for larger power savings. On the other hand, the use of pseudo-CMOS logic greatly simplified the design of the large number of input OR gates required per each row and column.

V. IMAGER OPERATION AND ANALYSIS

A. Pixel Operation

Because the proposed imager measures the time to integrate photon-generated charges to a threshold voltage, the consistency of this threshold voltage, which is set by the event generator in each pixel, plays an important role in the image quality.

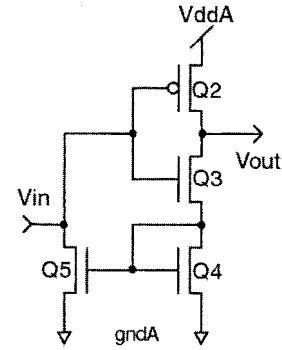


Fig. 10. Simplified view of the pixel intended for analysis.

From an analysis of the circuit in Fig. 3, we can define that the switching transition begins when the feedback current becomes comparable to the photocurrent. This definition is justified by the fact that at the switching point the input slew rate doubles because of feedback. As this happens, the positive feedback quickly switches the output. The input voltage at the start of switching $V_{\text{in, switch}}^-$ is given by

$$V_{\text{in, switch}}^- = \frac{nKT}{q} \ln \left(\frac{I_{\text{ph}}}{\left(\frac{W}{L}\right)_{Q4} \left(\frac{L}{W}\right)_{Q5} \left(\frac{W}{L}\right)_{Q2} I_{Q2,0}} \right) \quad (1)$$

where I_{ph} is the input photocurrent and $I_{Q2,0}$ is the weak inversion transistor $Q2$ current for zero bias. Before the switching event, the time-domain representation of the input voltage is given by

$$V_{\text{in}} = V_{dd,r} - \frac{t \cdot I_{\text{ph}}}{C}. \quad (2)$$

The subthreshold current through transistor $Q4$ causes the current feedback to start operating, and the inverter's output voltage also starts increasing. At the same time, transistor $Q7$ disconnects the integrating capacitors from the input of the inverter, thus, reducing its load. The fast increasing positive current feedback can then quickly drain the inverter's input capacitance. The magnitude of this positive feedback is at all times directly related to the current generated by $Q2$ and the gain of the feedback-current mirror. Once the input of the inverter reaches ground, the inverter current goes to zero and so does the feedback, because the nMOS transistor $Q3$ turns the diode-connected transistor $Q4$ off. Thus, at initial and final state there is no power-supply current. Consequently, the entire array of 80×60 , including the event generator and excluding the boundary circuits, dissipates $100 \mu\text{W}$, where $V_{dd,\text{analog}} = 2.75 \text{ V}$, and running at 200 kHz (events per second) in uniform room light of about $100 \mu\text{W}/\text{cm}^2$. When imaging a typical indoor scene, the analog power consumption drops to below $10 \mu\text{W}$, since the mean firing rate decreases.

B. Analysis of the Photosensor

To get an intuitive understanding of the operation of the current-feedback event generator, it is necessary to impose a few simplifications of the circuit and operational hypothesis. With the input voltage high and starting to decrease, transistors $Q3$ and $Q4$ in Fig. 10 sink the current sourced by transistor

TABLE I
TRANSISTOR OPERATION MODE DURING THE OCCURRENCE OF A TRANSITION

Transistor Operation mode	START	MID	END
Q 2	Saturation	Saturation	Triode
Q 3	Linear	Saturation	Saturation
Q 4	Saturation	Saturation	Saturation
Q 5	Saturation	Saturation	Triode
V _{in}	V _{dd}	V _{in} = V _{out}	0
V _{out}	~V _{thn}	V _{in} = V _{out}	V _{dd}

$Q2$ because, for similar size devices, nMOS transistors have larger transconductances and slightly lower threshold voltages than pMOS transistors. Furthermore, since V_{gs} of transistor $Q6$ (from Fig. 3) is given by the sum of the V_{gs} and V_{ds} of transistors $Q4$ and $Q5$, respectively, it is reasonable to expect $Q6$ to be on when the feedback mirror starts to operate. Hence, $Q6$ can be left out of the circuit. Transistor $Q7$ (from Fig. 3), which disconnects the capacitor from the input node, is on before the switching, and can also be neglected in the analysis.

Detailed analysis of the spike generator produces complicated mathematical relationships that provide no intuitive insights into the operation of the circuit. This results from the fact that the MOSFETs operate in all the modes—cutoff, saturation, and triode—in both weak and strong inversion. Hence, considerable abstractions must be made to obtain a simple and useful model for the switching characteristics of this circuit. To capture the modes of operation of the transistors, Table I has been compiled. By identifying the critical points from the table, we can develop approximate relationships for the currents in the output branch of the event generator, from which the switching speed and power consumption can be calculated.

The analysis of the onset of the transition has been already provided [(1)]. To determine the power consumption of the circuit, we must determine the peak current in the output branch. This occurs when $V_{in} = V_{out}$, and all transistors are operating above threshold in the saturation region. From Fig. 10, we determine that the peak current is given by

$$I_o^{\max} = \frac{1}{2} \frac{\beta_2 \beta_3 \beta_4}{\beta_2 \beta_3 + \beta_2 \beta_4 + \beta_3 \beta_4} \cdot (V_{dd} - V_{TON} - |V_{TOP}| - V_{TN})^{1/2} \quad (3a)$$

$$V_{out} = V_{in} = \sqrt{\frac{2I_o^{\max}}{\beta_3}} + \sqrt{\frac{2I_o^{\max}}{\beta_4}} + V_{TON} + V_{TN}. \quad (3b)$$

In (3), $\beta = \mu_o C_{ox}(W/L)$, V_{TO} is threshold voltage without body effect, and V_T is threshold voltage with body effect. From the process parameters, we compute $I_o^{\max} = 1.88 \mu A$ and $V_{in} = V_{out} = 2 V$. To calculate the energy consumption, the switching time of the circuit is required.

The rise time of the circuit is determined by the current in the output branch and the capacitance at the output node. The current that charges the output capacitor is the difference between the current sourced by $Q2$ and that sunk by the $Q3$ – $Q4$ pair.

This difference is initially negative when the $Q3$ – $Q4$ pair demands more current than $Q2$ can provide. As $Q2$ is turned on more, the capacitor current changes sign and eventually goes to zero when the $Q3$ – $Q4$ pair turns off as the input voltage V_{in} goes to zero. In the latter case, $Q2$ tries to provide a large current, i.e., its V_{sg} is maximum at V_{dd} , but goes into triode mode to match the sinking capability of the $Q3$ – $Q4$ pair. Hence, it is fair to approximate the largest current in the output branch, given by (3), to be equal to the current that charges the output node capacitance since the actual capacitor current will be both smaller and larger than I_o^{\max} . Using this approximation, we obtain (4) for the rise time of the event generator.

$$t_r = \frac{C_{out}}{I_o^{\max}} V_{swing}. \quad (4)$$

With the output swing running from V_{TON} to V_{dd} , the output transition was estimated at 6.75 ns. The energy consumption during the output ON transition is 0.021 pJ. These approximations are compatible with the simulations (8-ns rise time and 0.043-pJ energy consumption); measured data cannot be directly compared because additional circuits are included in the output path of the event generator.

C. Pixel Noise

The noise sources present at output of the proposed imager can be combined into two main categories. One, spatial noise, is caused by mismatch in circuit components, similar to that found in standard CMOS imager. The second category presents temporal jitter introduced by the phototransduction process and electrical circuit noise, by arbitration circuitry and by digital switching crosstalk. The former sources introduce fixed pattern noise (FPN), while the latter introduce temporal noise to the imager.

The imager has an FPN of approximately 4%, where FPN is given by the ratio of standard deviation to mean pixel value, under uniform ambient illumination. This value is worse than other CMOS imagers, primarily because FPN reduction steps, such as correlated double sampling (CDS), cannot be easily performed on time-domain phototransduction. CDS compensates for component mismatch by subtracting the output of the pixel during reset from the output after integration. This operation cannot be easily adapted to the presented time-domain imager, because the output is a spike and also because of the pixel-initiated readout method. A future version of the imager will include

a tentative emulation of the CDS by using an in-pixel analog memory and switched-capacitor circuit.

Blooming is another form of image degradation that commonly plagues CCD and some CMOS imagers. It occurs when the integrated charges overflow their holding wells, in-pixel capacitors, and spill into the neighboring pixels or the output line. Blooming occurs when the integration time is too long under bright lighting conditions. In our case, blooming effects are eliminated by allowing each pixel to self-regulate its integration time, based on the local brightness. Only the arbitration of output request collisions can momentarily lengthen the integration cycle; after the imager has been operating for some time, and provided the scene does not change significantly, collisions are reduced by the natural ordering of the pixels' integration cycle imposed by the arbitration circuit. Hence, if arbitration and readout happen sufficiently fast, the pixel has no opportunity to overflow.

First, we attempt to identify and quantify the spatial noise (i.e., FPN) sources in the imager. FPN has two sources: 1) mismatch in the photosensitive element, the photodiode, and 2) mismatch in the event-generator circuit that varies the value of the switching threshold voltage. The first component is a strict function of process variation and photodiode size. Typically, the larger the photodiode, the better matched they are across the chip. Unfortunately, constrained by the pixel size, the photodiode must be designed small enough so that the desired pixel count can be realized in the available die area. The second source, however, is dependent on the event-generator circuit. We can determine the sensitivity ($S_x^y = (x/y)(dy/dx)$) of the onset of the output switching point [provided by (1)] with respect to the mirror gain (ratio of Q_5/Q_4) and transistor Q_2 size in Fig. 3, as given in (5). Here, the onset of the switching process was used in place of the switching point $V_{in} = V_{out}$ since at this point the switching of the output has already reached the highest slew rate. Hence, a change in the voltage as $V_{in} = V_{out}$ has little impact on its temporal dispersion. On the other hand, the temporal dispersion of the onset of the switching process is strongly influenced by its voltage value.

$$S_{V_{in}, \text{switch}} = S_{M}^{V_{in}, \text{switch}} + S_{(W/L)_p}^{V_{in}, \text{switch}} = -2/V_{in, \text{switch}}. \quad (5)$$

The value for $S_{V_{in}, \text{switch}}$ was estimated to be -4.25 , where the mirror gain, $M = (W/L)_5/(W/L)_4 = 2.3$, $(W/L)_2 = 1.034$, and $I_{ph} = 0.1$ nA (typical room light), and the other parameters in (1) are typical values and/or determined by the fabrication process. This means that 1.5% error [27] due to size mismatch will produce an FPN of 6.38%, which is close to the measured data of Table II. Additional variation in the threshold voltages of the transistors, which will also vary the value of the switching threshold, will contribute to additional FPN.

For assessing temporal noise, we must consider integration, reset, arbitration, and crosstalk noise. For a typical APS imager, integration and reset noise, respectively, are expressed in (6) [19] by the first and the second term.

$$\overline{V_n^2} = \overline{V_n^2(t_{\text{reset}})} + \overline{V_n^2(t_{\text{int}})} = \frac{1}{2} \frac{kT}{C} + q \frac{I_{ph} + I_{dc}}{C^2} t_{\text{int}}. \quad (6)$$

In (6), I_{ph} is input photocurrent, while I_{dc} is the dark current, and t_{int} is the light integration time. We can estimate the dark

TABLE II
SUMMARY OF CHIP CHARACTERISTICS

Technology	0.6 μ m 3M CMOS
Array Size	80 (H) x 60(V)
Pixel Size	32 μ m x 30 μ m
Fill Factor	14%
Dynamic Range (with no minimum update rate per pixel)	180dB (Pix.) 120dB (Array)
Dynamic Range (>30 update per second per pixel)	108dB (Pix.) 48.9 dB (Array)
Bandwidth (dark current - arbiter speed limited)	8mHz – 8MHz (Pix.) 40Hz-40MHz (Array)
Pixel Inter-Event Time Jitter (STD/Mean)	~40% (illumination: 0.1mW/cm ²)
Sensitivity [Hz/mW/cm ²]	2x10 ⁶ (Array) 42 (Pix.)
FPN (STD/Mean pixel-pixel) @ 0.1 mW/cm ²	4%
Max. Update Rate Per Pixel	8.3K
Digital Power @ 2.9V Supply	(1.7F/MHz)+3.1) mW 3.4mW @ 0.1mW/cm ²
Analog Power @ 2.7V Supply	< 10 μ W for scene @ 0.1mW/cm ²

current the response of the imager in the dark. An event rate of 40 Hz in the dark for the whole array translates to 8 mHz per pixel on average. This means a spike every 120 s due to dark current at 20 °C temperature. This gives us $I_{dc} = 100$ fF \times 0.7/120 = 0.58 fA. Since the imager presented in this paper does not integrate a fixed amount of time but instead integrates to a fixed voltage threshold, (6) can be converted into

$$\overline{V_n^2} = \frac{1}{2} \frac{kT}{C} + q \frac{V_{in, \text{switch}}^{\text{threshold}}}{C}. \quad (7)$$

We define the switching threshold as the V_{in} that produces a feedback current which causes the input to slew faster than 100 V/ms; typical room light produces an input slew rate of ~ 1 V/ms. The value of $V_{in, \text{switch}}^{\text{threshold}}$ for our event generator is ~ 0.7 V. The interesting outcome of our approach, in contrast to typical APSs, is that the integration noise is independent of the light intensity. Here, integration noise turns into FPN through threshold voltage mismatch of each pixel's transistor Q_2 . The reset noise arises from the interaction between the reset transistor and the integrating capacitor. It is inversely proportional to the capacitor size. Since both noise sources are minimized by the use of a larger integrating capacitor, for the design of this image an explicit capacitor of much higher value than the intrinsic photodiode capacitance was used. The root-mean-square (rms) voltage noise was calculated to be 0.142 and 1.058 mV for the reset and integration noise terms, respectively. This adds to a variation of 1.067 mV at the input of the spike generator circuit. The noise voltage triggers the spike generator either earlier or later than the nominal noiseless value. Given the enormous gain of the circuit, caused by the positive feedback, the small noise variation at input can alter the position of the switch point. The resulting rms time skew error t_e in the output interevent interval can be calculated by

(8). Again we see that the percentage error is independent of light intensity.

$$\begin{aligned} t_{\varepsilon} &= \overline{V_{n, \text{rms}}} C / I_{\text{ph}} \rightarrow t_{\varepsilon} [\% \text{ of interspike interval}] \\ &= \overline{V_{n, \text{rms}}} / V_{\text{in, switch}}^{\text{threshold}}. \end{aligned} \quad (8)$$

The skew t_{ε} is estimated to be $1.067 \mu\text{s}$, which corresponds to an interspike interval error of 0.152% the integration time.

The last temporal noise components can be divided into three related causes. The first term is the arbitration jitter. The second term is digital crosstalk from the power supply. The third is readout temporal noise, which occurs when there is massive collisions of events or when the bandwidth of the channel is reached. A detailed discussion of the effects of digital crosstalk and arbiter noise can be found in [29]. Arbiter noise is 5.09% with an FPN of about 5% . This data was calculated by assuming $T_{\text{brt}} = 25 \text{ ns}$ and $T_{\text{cyc}} = 125 \text{ ns}$ and assuming that the imager operates at 90% of the channel capacity. It should be noted, however, that this magnitude of arbiter noise is not likely to be reached in normal operation since the channel capacity is not usually approached and the number of collisions are usually low.

The worst case readout noise is here presented *vis-à-vis* the measured results. When free-running (i.e., with no additional circuits in the request-acknowledge path), the request-acknowledge cycle $\tau_{\text{R-A}}$ takes 25 ns . In the worst case scenario, all pixels in the array request access simultaneously. The worst mean queueing time is $60 \mu\text{s}$ and the standard deviation is $34.6 \mu\text{s}$. The worst case variation in the interspike interval due to readout is given by (9a), where $N \cdot M$ is the size of the imager. In normal room light, $I_{\text{ph}} = 0.1 \text{ nA}$, the worst case interspike interval variation due to readout is 5% . This intrinsic limit can only be reduced by increasing the speed of handshaking and/or increasing the integration time to threshold CV .

$$\begin{aligned} t_{\text{arbitration}}^{\text{array_collision}} [\% \text{ of interspike interval}] \\ = \left(\frac{NM \tau_{\text{R-A}}}{\sqrt{12}} \right) \left(\frac{I_{\text{ph}}}{CV_{\text{in, switch}}^{\text{threshold}}} \right) \end{aligned} \quad (9a)$$

$$\begin{aligned} t_{\text{arbitration}}^{\text{row_collision}} [\% \text{ of interspike interval}] \\ = \left(\frac{N \tau_{\text{R-A}}}{\sqrt{12}} \right) \left(\frac{I_{\text{ph}}}{CV_{\text{in, switch}}^{\text{threshold}}} \right). \end{aligned} \quad (9b)$$

With the data collection system in the request-acknowledge path, we measured a minimum cycle time of 125 ns . This predicts an interspike interval jitter of 25% . As for the case of arbiter noise, this upper limit is not likely to be reached since the simultaneous request of all pixels rarely happens.

The additional measured jitter is due to digital crosstalk in the array. Crosstalk was measured on the chip analog power supplies' pins. Crosstalk noise was measured to be an average of 21.8 mV rms with a mean interevent timing of 1500 ns , up to 26.1 mV rms at 80 ns . Estimating a mean of 25 mV rms of noise on the power supply pin due to crosstalk, we can translate this differential voltage error into a timing error of 3.5% using (8) and assuming that the crosstalk noise bandwidth is lower than the cutoff frequency of the process' MOSFETS. In this case, we assume that the changes in the power supply reflect entirely on the threshold of the event generator's inverter. Locally, the

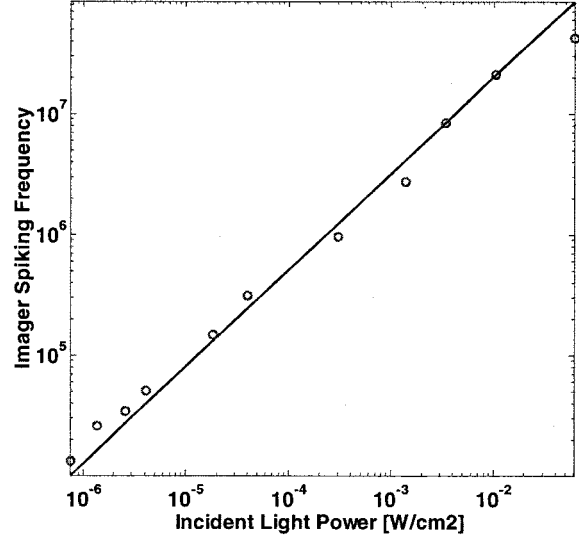


Fig. 11. Imager spiking frequency versus incident light power.

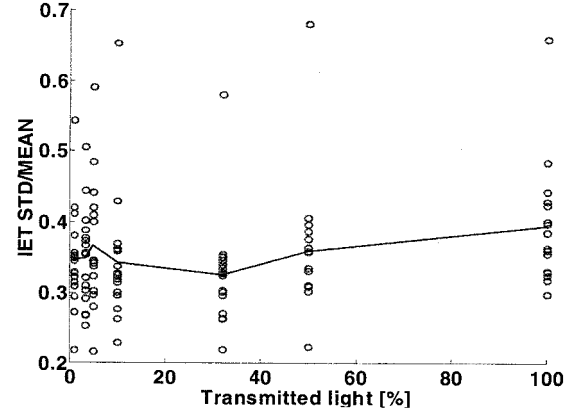


Fig. 12. STD/mean of interevent timing in different lighting conditions.

drag on the power supply is likely to be much larger (as much as 10 times larger from simulations), which will further exacerbate the problem, resulting in larger temporal jitter. Fortunately, for imaging purposes, the temporal jitter can be considerably reduced by averaging the interspike interval.

On the other hand, after the row pipelined architecture has grouped the integration cycle for each pixel in a row, and has distributed the request, i.e., the completion of integration, for each row, the arbitration error is due to pixel access within the row. In this case, the variation is given by (9b), which predicts 0.4% variation, using the 125-ns cycle time. Unfortunately, this cannot be obtained because FPN and digital crosstalk will prevent perfect pixel (in a row) grouping and row distribution. Nonetheless, it indicates that 8-b instantaneous digital imaging is possible with better matching and digital isolation, even at this slow arbitration rate.

Fig. 11 shows a plot of measured variations (standard deviation divided by the mean array value) in the interspike interval versus uniform light intensity. The figure is generated by computing the temporal statistics of a high number of pixels. We expect a linear relationship between intensity and σ/μ for the imager. The linear relationship is not strongly visible in the plot.

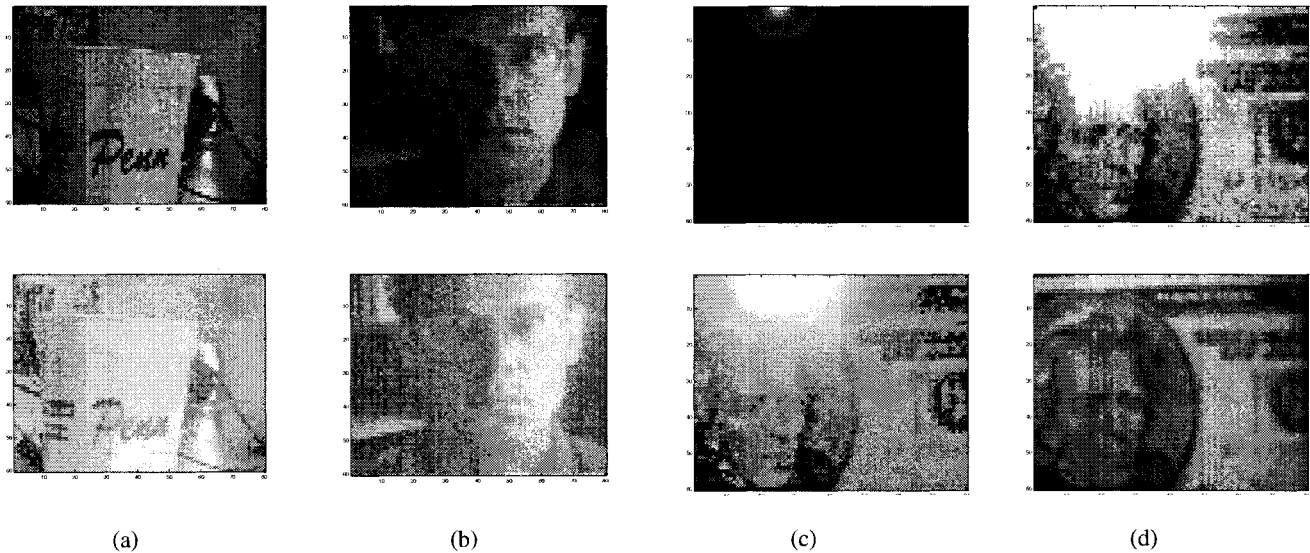


Fig. 13. Example images. (a)–(c) Linear intensity (top) and log (bottom) scales. (d) (top) Linear intensity of first 256 gray levels and (bottom) linear intensity without bright source.

Furthermore, we observe a number of pixels with large variations. The mean of the variation agrees with expectation, but the variation across pixels is unexpected. We believe that this spread is due to excessive noise on the power supplies in the pixels. Power-supply noise can strongly influence the switching voltages for the individual event generators and will be observed as jitter in the interspike interval.

In summary, the noise sources are 6.38% size mismatch, 0.15% electronic noise, <5% arbitration, 3.5% to >35% (from simulations) crosstalk, and <25% readout noise. The measured standard deviation (STD) to mean ratio of $\sim 40\%$ (from Fig. 12) can be easily explained by noting that the crosstalk induced noise can be >35%.

D. Image Reconstruction

To obtain a pixel intensity image, the interspike interval must be converted into light intensity. The photocurrent is inversely proportional to the interspike interval or directly proportional to the spike frequency. To perform these transformations, each spike is time indexed relative to a global clock and the time between successive spikes computed (instantaneous interspike interval), or the number of spike over a fixed interval is counted (average interspike interval or pixel update rate). In either case, the AER data must be stored or accumulated in a memory array. This can be in the form of analog storage (capacitive storage, for example) or in the digital domain. A workstation computer was used to accumulate events and generate the images presented here. An interface program was responsible for collecting up to one million samples, and then reconstructing an image histogram in memory. Real-time medium-quality images can be displayed every 10K–20K samples. We also associated each event with a time index to analyze the temporal characteristics of the imager. The timing circuitry was a programmable Altera field-programmable gate array acting as a 28-b counter.

The main drawback of this approach is the complexity of the digital frame grabber required to count all the spikes produced

by the array. A high-resolution timer (up to 24 bits for hundreds of picosecond resolution) and a large frame buffer are required (up to 15 MB for a full VGA array) would be required to obtain an instantaneous image for every spike. The timer indexes each event and compares it with the last time an event at that pixel was recorded. The difference is inversely proportional to the light intensity. The buffer must hold the latest pixel time index and the intensity value.

Fig. 13 shows example images recorded with the array. The figure shows the spike frequency per pixels after collecting about one million events. Conventional imagers produce linear results similar to the top row of Fig. 13. There is no information in the dark portions of the image because there, pixel intensities are below the least significant bit of the analog-to-digital converter (typically, 8 bits) used to digitize the image. However, in this imager, information in the dark portions of the image is available. After integrating for longer intervals, the low-intensity portions of the image can be constructed, while the bright portions of the image can be immediately rendered. This methodology of wide-dynamic-range imaging is commonly performed in biological visual systems. To emphasize the wide dynamic range of the proposed imager, an additional high-intensity light source is included in the scene. To demonstrate the presence of image information in the dark regions, the bottom images in Fig. 13(a)–(c) show the log of the image intensity. Pictures were taken at a uniform background lighting of 0.1 mW/cm^2 . Notice that the features in the shadows can now be observed. In Fig. 13(d) (top), we display the first 256 gray levels of the top image in Fig. 13(c). Pixel values above 256 are saturated to 256. Again, the information in the dark parts of the image is visible. Finally, in Fig. 13(d) (bottom), the high-intensity light source is turned off and the regular image is constructed. Notice that the visible parts of Fig. 13(d) top and bottom are similar. The variations in the images are primarily due to FPN. Temporal noise is mostly averaged out of the bright parts of the image due to the spike frequency representation. For the dark portions of the image, temporal

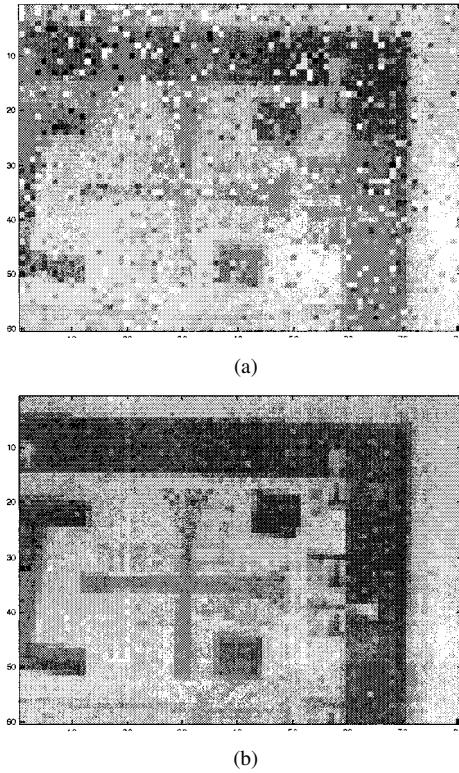


Fig. 14. Instantaneous interspike interval image compared to the spike frequency image (after $\sim 100K$ events). (a) Instantaneous image obtained by computing the temporal difference between two spikes. (b) Computation of spike frequency per pixel.

noise is further amplified because fewer events are collected. Furthermore, arbitration jitter is prevalent here, due to the high number of bright pixels competing for the readout bus.

Fig. 14 shows the effects of temporal jitter on the collected images. Fig. 14(a) shows the instantaneous image obtained by computing the temporal difference between two spikes. Conventional imagers scan a number of pixels equal to the imager pixel count before updating the image. Here, an equal number of pixels are sampled. After only 4800 events, the image is updated. In our reconstruction, the brighter regions will be updated more often than the darker regions, according to the statistics of the scene.

In this picture, the temporal noise is quite evident, however the pixel update rate is extremely high ($\sim 1.67K$ per second with the measurement system in the loop); continuous image updates are possible with each event received. In Fig. 14(b), similar to Fig. 13, spikes are collected for some time ($\sim 1M$ events) and the spike frequency per pixel (~ 208 spikes per pixel) is computed. Here, the temporal jitter is mostly eliminated and the spatial FPN of the array is visible. The two approaches trade off pixel update rate versus image signal-to-noise ratio (SNR), and the desired characteristic can be selected according to the applications.

VI. DISCUSSION

A. Imager Statistics and Light Sensitivity

Because of the output-on-demand nature of the proposed imager, the integration, readout, and reset cycles are executed

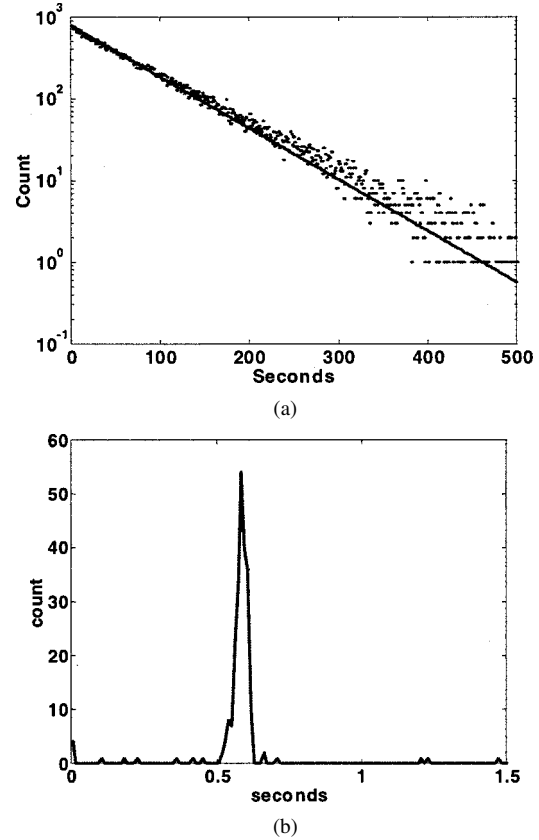


Fig. 15. Poisson distribution of events. (a) Interspike interval and (b) variation.

mainly asynchronously. The row-pipelining algorithm can impose some synchrony between pixels in the same row, provided they are exposed to the same light intensity. Because pixels act independently, the readout sequence queues and outputs individual spikes according to a Poisson process. Consequently, the probability of appearance of an address from a certain region is proportional to the light intensity in that neighborhood. This is the first reported example of a probabilistic APS, where the output activity reflects the statistics of the scene. Fig. 15 shows an example of the distribution of events for a typical lab scene. Fig. 14(a) suggests a clear exponential behavior for interevent timing for the array. The parameters extracted are 790 as intercept and -0.01446 as exponent multiplier. Fig. 13(a) shows the interevent timings of a single pixel, while viewing a scene of a room. Since the intensity distribution is scene dependent, the interspike interval distribution is influenced by both the scene statistics and the arbitration process. The cycle time statistics for events generated by pixel (30, 30) of Fig. 13(a) show a mean of 0.54527 events per second and standard deviation of 0.0580.

In addition, we provide in Fig. 11 a plot of the imager spiking frequency versus incident light power. This data was obtained by measuring the light intensity with a Newport photometer model 1830-C. A fit for the graph in Fig. 11 is represented by

$$8 \cdot 10^8 * \text{light}^{0.8}.$$

The temporal effects of the arbitration circuit are visible in high-intensity light as the curve in Fig. 11 becomes less linear with the increasing events rate. The photometer used prevents us from measuring lower light intensities accurately.

B. Imager Limitations and Scaling

Under uniform bright illumination, the array of 80×60 pixels shows a dynamic range of 120 dB (40 Hz–40 MHz). This dynamic range is possible when no lower bound is placed on the pixel update rate. That is, if a lower bound of 30 updates per second is imposed, the array rate covers 144 kHz–40 MHz, which implies an array dynamic range of 48.9 dB. The 40-MHz event rate is only observed with our data collection system out of the loop. Depending on the application and the light intensity falling onto the sensor, imaging can always trade dynamic range for pixel update rate. Depending on which is desired, one tradeoff can be made. For example, while tracking a laser spot on a target, the pixel update rate can be as high as 8.3K per second, since resolution is not important for this application. On the other hand, if video imaging is required, the spatial resolution has to be higher and the pixel update rate decreases. Still the imager presents different in-frame dynamic range variation due to light intensity. Highly illuminated areas will present fast update rates and high dynamic range, while lower light areas will suffer from motion artifacts and limited dynamic range.

Similarly, the dynamic range for an individual pixel is 180 dB (from 0.008 Hz = 40 Hz per number of pixels, to 8 MHz, provided that one pixel could access the boundary circuit by itself). To do this experiment, the reset transistor Q_1 in Fig. 3 is left slightly on to cancel the dark current in the photodiode. If this precaution is not observed, then spontaneous activations in some of the dark pixels will occur, and the pixel under observation will have to share the bandwidth with others. The 8-MHz limitation derives from the fact that the same pixel has to undergo column and row arbitration for each event, thus, increasing the interevent cycle to $125 \text{ ns} = T_{\text{cyc}}$. A pixel on the same row can, on the other hand, benefit from AE pipelining and, thus, be transmitted at the maximum speed of 40 MHz. Table II summarizes the characteristics of the array. The power consumption is 3.4 mW in uniform indoor light (0.1 mW/cm^2), which produces a mean event rate of 200 kHz (41.6 updates per second per pixel). The imager is capable of operating at a maximum speed of 8.3K updates per second per pixel (under bright local illumination). This maximum speed is obtained by sampling a number of pixels equal to the pixel count of the array (exactly like a scanned imager) and using interevent timing information with the previously sampled events to render a frame.

The relationship between event (output) frequency and power consumption is given by (10) (empirical fit)

$$\begin{aligned} \text{DR}(N \cdot M) &= \frac{\text{EF}}{\text{PF}^{\min} \cdot N \cdot M} \leq \frac{\text{SBW}}{\text{PF}^{\min} \cdot N \cdot M} \\ &= \frac{(40 \text{ MHz})}{(0.008 \text{ Hz}) \cdot N \cdot M} \end{aligned} \quad (10a)$$

$$\text{EF}(N \cdot M, \text{DR}) = \text{PF}^{\min} \cdot \text{DR} \cdot N \cdot M \leq \text{SBW} \quad (10b)$$

$$\text{FR}(N \cdot M) = \frac{\text{EF}}{N \cdot M} \leq \frac{\text{SBW}}{N \cdot M} = \frac{(40 \text{ MHz})}{N \cdot M} \quad (10c)$$

$$\text{Power}[\text{mW}](\text{EF}) = 1.7(\text{EF}[\text{MHz}]) + 3.1 \quad (10d)$$

where FR is update per second per pixel, SBW is the maximum system bandwidth (40 MHz), DR is the dynamic range (ratio of

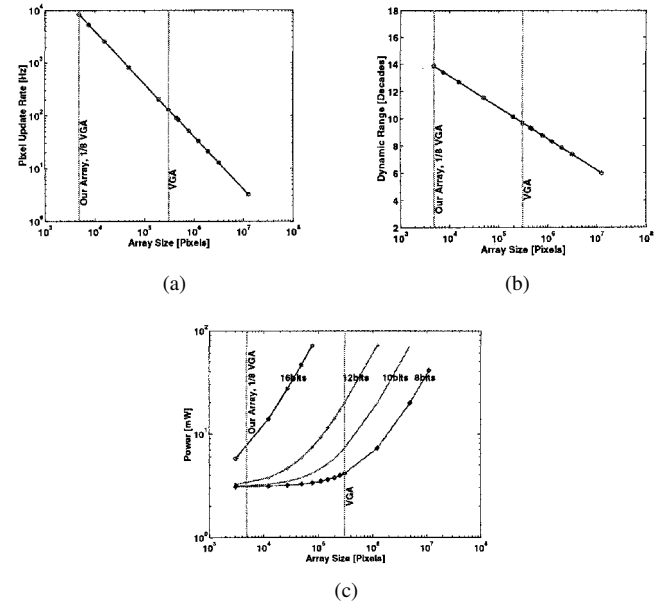


Fig. 16. Scaling properties of the array. (a) Update rate per pixel. (b) Dynamic range. (c) Power consumption.

current update rate per pixel to the smallest), PF^{\min} is the minimum pixel frequency (0.008 Hz), EF is the event frequency, and N , M the array size. The static dissipation is produced by the pseudo-CMOS logic used in this design. At full speed (40 MHz), and maximum array dynamic range (6 decades), the power consumption will be 71 mW. Normal operation produces events at a maximum of 4 MHz (0.8K updates per second per pixel), for a dynamic range of 5 decades, while consuming 10 mW. Fig. 16 shows how the update rate per pixel, dynamic range, and power consumption vary as the array size scales. Since the output bandwidth is shared between the pixels of the array, as the number of pixel increases, the dynamic range and update rate per pixel decrease. Equivalently, given the significant increase in spike rate with the number of pixels, power consumption increases at a rate proportional to the desired output precision. This is understandable, since the output precision is affected by the number of events collected for each pixel.

Since the imager produces statistical images depending on the local intensity of light, some areas of the image will be updated more frequently than others. This in turn will produce some heterogeneous motion artifacts: portions with high illumination will not suffer from motion artifacts, but darker portions will suffer in proportion to the light intensity. In this regards motion artifacts are inversely proportional to the dynamic range desired. High dynamic range requires slow integration times, which give high motion artifacts. Fig. 16(a) and (b) shows how dynamic range and update rate per pixel are related and can be traded. Another alternative is to trade resolution for update rate per pixel, combining more neighboring pixels together. In order to compute update rate per pixel at which motion blurring will occur, we can proceed as follows. First, measure the light intensity of the target environment, using the data in Fig. 11, on calibrated spike rate versus light intensity, to get an estimate of the target spike rate. Second, divide the average spike rate by the number of pixels of the array. The following equation gives

a numerical estimate of the update rate per pixel at a given light intensity:

$$FR = 1.66 \cdot 10^5 (\text{light} [\text{mW}/\text{cm}^2])^{0.8}. \quad (11)$$

This will give the number of updates per second per pixel at the target light intensity. If the update rate is lower than the desired one, then blurring will occur. That is, if 30 updates per second per pixel are required, then the mean spike rate for the array cannot drop below 144 kHz. In that case, the dynamic range of the array will be 48.9 dB. However, it should be obvious that this number is the case under uniform illumination, where all pixels are trying to access the output bus at the same rate. For real scenes, some pixels will spike at much lower rates than others. By simply treating pixels whose spike rates are below 30 Hz as black, a larger dynamic range can be achieved.

Depending on the application and the light intensity falling onto the sensor, imaging can always trade dynamic range for update rate. Depending on which is desired one tradeoff can be made. For example, in tracking a laser spot on a target, the update rate per pixel can be as high as 8.3K, since resolution is not important for this application. The imager presents different in-frame dynamic range variation due to light intensity. Highly illuminated areas will present fast update rates and high dynamic range, while lower light areas will suffer from motion artifacts and limited dynamic range.

VII. SUMMARY

An 80×60 pixels fully arbitrated AE light-to-bits imager is fabricated and tested. The imager provides a very large dynamic range of 120 dB in uniform bright illumination and when no lower bound is placed on the update rate per pixel, a low power consumption of 3.4 mW in normal indoor lighting and is capable of a maximum of 8.3K updates per second per pixel under local bright illumination. At 30 frames per second, the dynamic range for imaging ambient light scenes is 48.9 dB. The power consumption can be further reduced by removing all pseudo-CMOS logic devices. This imager compares favorably to traditional CMOS imagers (in a $0.5\text{-}\mu\text{m}$ process) in terms of speed and power, but needs additional optimization to match their image quality [21], [22]. We find that the main sources of image noise are FPN due to component and parameter mismatch and temporal jitter due to digital crosstalk-induced power-supply noise. The former can be reduced by using emulation of correlated double sampling, which must be implemented in each pixel, while the latter is a function of the image statistics. Temporal jitter can be reduced by employing layout practices that reduce digital crosstalk. Furthermore, by increasing the bandwidth of the arbitration and/or reducing the nominal spike rate per light intensity, temporal jitter due to arbitration and collisions during readout can be reduced. In addition, reducing FPN will also decrease temporal jitter since the arbitration process minimizes collisions by synchronizing pixels in a row and distributes the row access.

ACKNOWLEDGMENT

The authors would like to thank C. Wilson, A. van Schaik, A. Andreou, and P. Poliquen for their insights, fruitful discussions, and support.

REFERENCES

- [1] C. Mead and M. A. Mahowald, *A Silicon Model of Early Visual Processing*. New York: Pergamon, 1988.
- [2] E. R. Fossum, "CMOS image sensors: Electronic camera-on-a-chip," *IEEE Trans. Electron Devices*, vol. 44, pp. 1689–1698, Oct. 1997.
- [3] J. Burns *et al.*, "Three-dimensional integrated circuits for low-power high-bandwidth systems on a chip," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 44, 2001, pp. 268–269.
- [4] M. Koyanagi *et al.*, "Neuromorphic vision chip fabricated using three-dimensional integration technology," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 44, 2001, pp. 270–271.
- [5] A. Moini, *Vision Chips*. Norwell, MA: Kluwer, 1999.
- [6] R. Etienne-Cummings, M.-Z. Zhang, P. Mueller, and J. Van der Spiegel, "A foveated silicon retina for two-dimensional tracking," *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 504–517, June 2000.
- [7] X. Arreguit, F. A. van Schaik, F. Bauduin, M. Bidiville, and E. Raeber, "A CMOS motion detector system for pointing devices," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1916–1921, Dec. 1996.
- [8] K. Boahen, "A throughput-on-demand address-event transmitter for neuromorphic chips," in *Proc. ARVLSI'99*, Atlanta, GA, pp. 72–86.
- [9] A. Mortara and E. A. Vittoz, "A communication architecture tailored for analog VLSI artificial neural networks: Intrinsic performance and limitations," *IEEE Trans. Neural Networks*, vol. 5, pp. 459–466, May 1994.
- [10] K. A. Boahen, "Communicating neuronal ensembles between neuromorphic chips," in *Neuromorphic Systems Engineering*, T. S. Lande, Ed. Boston, MA: Kluwer, 1998, ch. 11, pp. 229–261.
- [11] V. Brajovic *et al.*, "Sensor computing," *Proc. SPIE*, vol. 4109, 2000.
- [12] W. Yang, "A wide-dynamic-range low-power photosensor array," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, 1994, pp. 230–231.
- [13] S. Decker, D. McGrath, K. Brehmer, and C. Sodini, "A 256×256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE J. Solid-State Circuits*, vol. 33, pp. 2081–2091, Dec. 1998.
- [14] J. Lazzaro and J. Wawrzynek, "A multi-sender asynchronous extension to the AER protocol," in *Proc. Conf. Advanced Research in VLSI*, 1995, pp. 158–169.
- [15] I. G. Gleadall, K. Ohtsu, E. Gleadall, and Y. Tsukahara, "Screening-pigment migration in the octopus retina includes control by dopaminergic efferents," *J. Exp. Biol.*, vol. 185, no. 1, pp. 1–16, 1993.
- [16] B. A. Wandell, *Foundations of Vision*. Sunderland, MA: Sinauer, 1995.
- [17] D. H. Hubel, *Eye, Brain, and Vision*. San Francisco, CA: Freeman, 1988.
- [18] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison Wesley, 1989.
- [19] Z. K. Kalayjian and A. G. Andreou, "Asynchronous communication of 2-D motion information using winner-takes-all arbitration," *Analog Integrat. Circuits Signal Process.*, vol. 13, no. 1–2, pp. 103–109, 1997.
- [20] T. Serrano-Gotarredona, A. G. Andreou, and B. Linares-Barranco, "AER image filtering architecture for vision-processing systems," *IEEE Trans. Circuits Syst. I*, pp. 1064–1071, Sept. 1999.
- [21] E. R. Fossum, A. Krymski, and K.-B. Cho, "A 1.2-V micropower CMOS active pixel image sensor for portable applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, 2000, pp. 114–115.
- [22] M. Hillebrand, B. J. Hosticka, N. Stevanovic, and A. Teuner, "A CMOS image sensor for high-speed imaging," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, 2000, pp. 104–105.
- [23] A. G. Andreou and K. A. Boahen, "A 48 000 pixel, 590 000 transistor silicon retina in current-mode subthreshold CMOS," in *Proc. 37th Midwest Symp. Circuits and Systems*, vol. 1, 1994, pp. 97–102.
- [24] M. Mahowald, *An Analog VLSI Stereoscopic Vision System*. Boston, MA: Kluwer, 1993.
- [25] R. J. Baker, H. W. Li, and D. E. Boyce, *CMOS, Circuit Design, Layout, and Simulation*. New York: IEEE Press, 1998.

- [26] B. Fowler, A. El Gamal, and H. Tian, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, pp. 92–101, Jan. 2001.
- [27] T. Serrano-Gotarredona and B. Linares-Barranco, *Analog Integrated Circuits and Signal Processing*. Boston, MA: Kluwer, 1999, vol. 21, pp. 271–296.
- [28] M. Arias-Estrada, D. Poussart, and M. Tremblay, "Motion vision sensor architecture with asynchronous self-signaling pixels," in *Proc. 4th IEEE Int. Workshop Computer Architecture for Machine Perception (CAMP)*, 1997, pp. 75–83.
- [29] K. A. Boahen, "A burst-mode word-serial address-event link: Transmitter design," *IEEE Trans. Circuits Syst. II*, to be published.



Eugenio Culurciello received the M.S. degree from the University of Trieste, Trieste, Italy, in 1997 and the M.S. degree from The Johns Hopkins University, Baltimore, MD, in 1999. He is currently working toward the Ph.D. degree in electrical and computer engineering at Johns Hopkins.

He is an Assistant Researcher at The Johns Hopkins University. His research interests are artificial vision and neural-morphology, efficient biomimetic communication channels, and wireless high-frequency communication devices.



Ralph Etienne-Cummings received the B.Sc. degree in physics from Lincoln University, Lincoln University, PA, in 1988 and the M.S.E.E. and Ph.D. degrees in electrical engineering from the University of Pennsylvania, Philadelphia, in 1991 and 1994, respectively.

He is currently an Associate Professor of electrical and computer engineering at the University of Maryland, College Park, on leave from The Johns Hopkins University, Baltimore, MD. His research interests include mixed-signal VLSI systems, computational sensors, computer vision, neuromorphic engineering, smart structures, mobile robotics, and robotics-assisted surgery.

Dr. Etienne-Cummings is a recipient of the National Science Foundation Career Development Award and the Office of Naval Research Young Investigator Program Award.



Kwabena A. Boahen received the B.S. and M.S.E. degrees in electrical and computer engineering from The Johns Hopkins University, Baltimore, MD, in 1989, where he made Tau Beta Kappa. He received the Ph.D. degree in computation and neural systems from the California Institute of Technology, Pasadena, in 1997, where he held a Sloan Fellowship for Theoretical Neurobiology.

He is currently an Assistant Professor in the Department of Bioengineering, University of Pennsylvania, Philadelphia, where he holds a Skirkanich Term Junior Chair and a secondary appointment in Electrical Engineering. His current research interests include mixed-mode multichip VLSI models of biological sensory and perceptual systems and their epigenetic development, and asynchronous digital interfaces for interchip connectivity.

Dr. Boahen was awarded a Packard Fellowship in 1999, a National Science Foundation CAREER Award in 2001, and an Office of Naval Research Young Investigator Program Award in 2002.