

# Early Research Progress on Quanta Image Sensors

Saleh Masoodian, Yue Song, Donald Hondongwa, Jiaju Ma, Kofi Odame and Eric R. Fossum

*Thayer School of Engineering at Dartmouth, Hanover, New Hampshire USA*

**Abstract**—Early research progress in the realization of the Quanta Image Sensor is reported. Simulation of binary data acquisition and image formation was performed. Initial analysis and simulation of a readout signal chain has been performed and bounds on power dissipation established. Photodetector device concepts have been explored using TCAD.

## I. INTRODUCTION

The Quanta Image Sensor (QIS) was first proposed in 2005 in conjunction with an algorithm to form a “digital film sensor.” [1,2] Advances in SPAD devices [e.g., 3] and binary pixel theory and imaging algorithms [e.g., 4] have been made since then, and the ensemble of work has encouraged us and our sponsor to engage in a research program to investigate methods to realize binary pixels, photoelectron counting and the QIS. Key elements of this research include sub-electron read noise, high scan rates, massive binary data handling, and low power dissipation. Investigating methods of creating high quality images from the binary data is another part of the effort.

The QIS is comprised of an array of specialized binary pixels called jots that provide sensitivity to a single photoelectron. To improve the storage capacity of the sensor (bits/um<sup>2</sup>) jots are anticipated to be sub-diffraction-limit (SDL) in size (e.g., 100-200 nm pitch). To improve SNR and dynamic range, the readout scan rate of the jot array may be 16x or higher than that of conventional image sensors (e.g., 480 or 960 fields per sec.). Thus a QIS may consist of 0.1-100Gjots with data rates from 0.1-100Tbits/s.

## II. END-TO-END SYSTEM SIMULATION

### Generating The Jot Data Cube

Simulation of the QIS system starts with simulating the acquisition of binary jot data. The data can be artificially generated from existing images. For simplicity we started with small scale JPEG images, understanding that some JPEG artifacts could obscure artifacts from the emulation process. A standard 256x256 image was used (“Lena”). Each pixel was converted to  $s \times s$  subpixels using bicubic interpolation. We used  $s=16$  for most of this initial work. Each subpixel was then converted to a binary jot value (0,1) using Poisson statistics [5]. If  $S_H$  is the subpixel value (0-255) and  $h_o$  an illumination factor (0-10) then the quanta exposure  $H$  for the subpixel is determined as:

$$H = \frac{S_H h_o}{255} \quad (1)$$

The probability of the jot receiving at least one photoelectron,  $P[J_1]$ , is given by:

$$P[J_1] = 1 - e^{-H} \quad (2)$$

The binary jot data corresponding to the image is generated. The process is repeated  $m$  times (e.g.,  $m=16$ ) to generate  $m$  time slices of jot data.

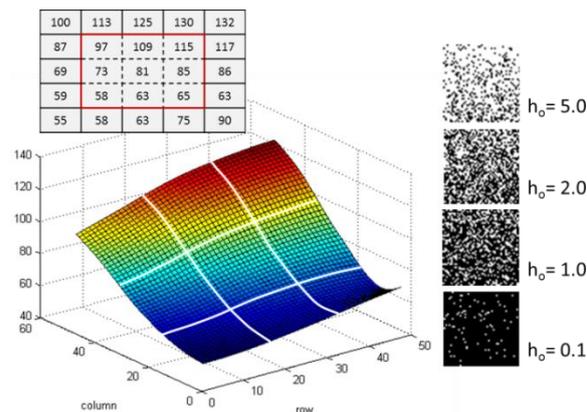


Fig. 1. Example of 3x3 pixels from original image interpolated to 48x48 subpixels and then on right, jot data corresponding to 48x48 jots for 4 different exposure settings, from sparse ( $h_o=0.1$ ) to overexposed ( $h_o=5.0$ ). Exposure  $H$  for jot determined as described in text. For example, center pixel  $S_H=81$  so with  $h_o=1.0$ ,  $H=0.32$

Starting with an original image of 256x256 pixels yields a binary jot data cube ( $x,y,t$ ) of 4096x4096x16 bits. Essentially each original pixel has been translated to 16x16x16 jots.

Read noise and dark current noise corruption can be included in the jot data cube. The bit error rate ( $BER$ ) – the probability of a bit flip - is determined from the signal chain input-referred read noise  $n_r$  according to:

$$BER = \frac{1}{2} \operatorname{erfc} \left( \frac{1}{\sqrt{8n_r}} \right) \quad (3)$$

Generally, we target QIS design with  $BER$  less than 0.001 corresponding to  $n_r = 0.15$  e- rms. Dark bits from thermally generated electrons are expected to be below this rate but can be included in the total  $BER$  for simulation. Since  $BER$  grows rapidly with  $n_r$  (e.g.,  $BER$  grows 100x if  $n_r$  doubles to 0.30 e- rms) rapid deterioration in image quality is found for  $n_r$  above the target value. Likewise, due to the rapid fall off of  $BER$  with read noise below the target value (1500x for  $n_r = 0.10$  e- rms) we found that simulation of read noise corruption is not too revealing.

From this point forward, the original pixelation is ignored and we proceed as if the jot data cube was captured by an actual QIS.

### Image Formation

The simplest method, perhaps, to form an image from the jot data cube is to sum the binary data over some region to form each pixel. Let the jot data cube consist of jot values  $J_{xyt}$ , where  $x \in (1-4096)$ ,  $y \in (1-4096)$  and  $t \in (1-16)$ . A reconstructed pixel value  $R_{ab}$ , is given by:

$$R_{ab} = \sum_{x,y,t=j(a-1)+1,j(b-1)+1,1}^{ja,jb,16} J_{x,y,t} \quad (4)$$

where  $jxj$  is the number of jots utilized in  $x$  and  $y$  to create the reconstructed pixel (e.g., 16x16, 4x4, or even 2x2). For the case of 16x16, the maximum sum is 4096 which may need to be normalized by 255/4096 for 8b rendering. For 4x4, the maximum sum is lower and the shot noise is proportionally higher. An example of this is shown below in Fig. 2.



Fig. 2. (a) Original 256x256 image. (b) Reconstructed image using simple summation, based on  $jxj=4x4$ . Image is 1024x1024 and shrunk to match the original image. Close examination shows "shot" noise in the reconstructed image as expected.

In essence, the summation approach is equivalent sampling of a box filter convolution, with filter weight of unity inside the box and zero outside. We chose a spatial sampling rate commensurate with the box filter size ( $jxj$ ). This is consistent with the current paradigm in imaging. In conventional sensors, we just accumulate the total number of photoelectrons in x-y dimensions determined by the pixel pitch, and in time according to the sensor integration period. In the QIS however, we could spatially sample the filtered output at a higher or lower rate. We can also dynamically adjust the size of the box filter in any of the 3 dimensions. This has the effect of adjusting the effective pixel size and integration time to optimize resolution v. SNR.

The filter weight need not be just binary in value. We have explored filter weights that result in a Gaussian-like distribution, but where the actual weights are weighted in powers of 2. This is because such weighting is easy to accomplish on the image sensor itself by simple shifts if so desired. An example of such a filter is shown below in Fig. 3. Results from the filter typically gave better results than simple summation and noise depends on the size of active kernel. Generally for output pixel pitch

corresponding to  $j$  jots, a kernel size of  $2j$  gave the best trade between spatial resolution and noise.

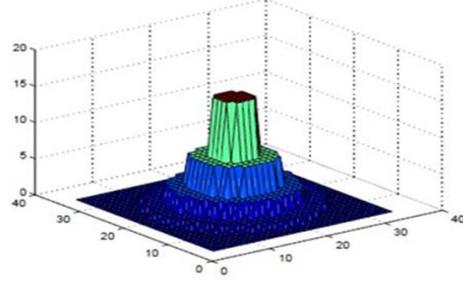


Fig. 3. Example of binary-power-weighted pseudo-Gaussian filter extending over 32x32 jots for 16 jot output pixel pitch.

Many different filters were explored with varying degrees of trade-off between modulation transfer function (MTF) and noise. Dynamic filters, where the kernel size is dynamically adjusted in accordance with spatial frequency in the output images, has been an interesting avenue, but requires more processing since it is a sort of recursive algorithm.

### Synthetic Images

To quantify the relationship between noise and MTF, we created synthetic jot images based on ideal gray patches and spatial frequency patterns. An example of a 80x420 pixel synthetic image we used is shown below in Fig. 4. The image is converted to a jot data cube using the method described above. The filter is applied and then the output image can be measured automatically. The image is collapsed vertically to measure MTF in the horizontal direction. Pixels are summed in vertical stripes to measure noise vs. grey level.



Fig. 4. Synthetic image for measuring both MTF and noise.

An example of using the synthetic image as the starting point for exploring a digital film algorithm (basically, region-growing) is shown below. First, if a "grain" of size  $jxj$  contains at least one jot with value 1, all jots in that grain are converted to a value of 1. This is effectively a gain step. Second, a dynamic kernel-size pseudo-Gaussian filter is applied. The results of each step are shown below in Fig. 5.

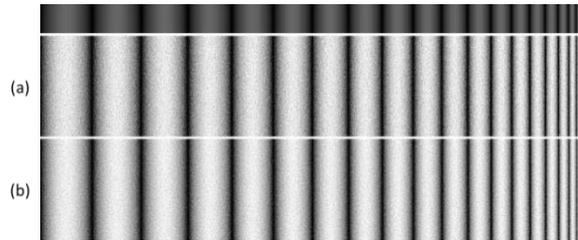


Fig. 5. (a) DFS algorithm (b) followed by dynamic-kernel-size pseudo-Gaussian filter (Original synthetic image shown at top for reference)

### III. READOUT SIGNAL CHAIN SIMULATION

#### Sense Amplifier

An early concern with the QIS concept was the required power dissipation for readout. To address this concern, design, simulation and layout of a strawman signal chain in 0.18  $\mu\text{m}$  CMOS was performed. A test chip is presently in fab.

Reduction to practice of a jot device requires use of advanced fabrication capability, not currently available to our research group. For the purpose of the design activity, a generic jot device was modeled – one with a column access transistor and with conversion gain of 1 mV/e<sup>-</sup>. The device is simulated as a conventional CMOS APS with floating diffusion (FD) reset, followed by intra-pixel transfer using a pinned photogate to the FD. In fact, all that counts is a presence or absence of a -1 mV step voltage at the output of the in-pixel source-follower. The design assumes the number of vertical jots in the column,  $J_V$ , is 10,000. The capacitance of the column line was loaded with the capacitance of  $J_V$  access transistors (nominally for a 0.1 Gjot QIS). The column bias current was determined in accordance with timing and drive requirements.

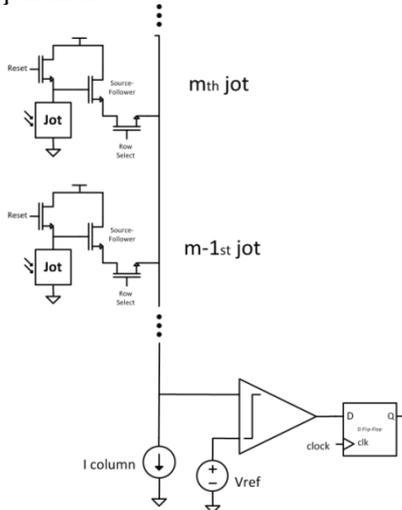


Fig. 6. Schematic illustration of readout signal chain including jots, preamplifier and latch.

To conserve power, a low power preamplifier [6] was adapted, followed by a D-latch. The preamplifier applies a gain of about 10 to the jot signal – enough to overcome threshold variations in the D-latch as determined by Monte-Carlo simulations. The circuit is shown below in Fig. 7.

The design uses  $V_{DD}=1.8\text{V}$ , a column bias current of 3.6  $\mu\text{A}$  and comparator power of 1.3  $\mu\text{W}$ . Simulation shows that for 100 fps readout (10k jots per column) or 1Mj/s, the design achieves total power of 7.7  $\mu\text{W}/\text{column}$ . For a 0.1 Gjot sensor (10k x 10k) this corresponds to 77 mW for 10Gb/s internal readout of the array. Of course any digital signal processing will add to the total power, and off-chip drive of data, if not compressed, will add significantly to the power budget.

A test circuit chip was taped out for fab to prove the principles employed in the design.

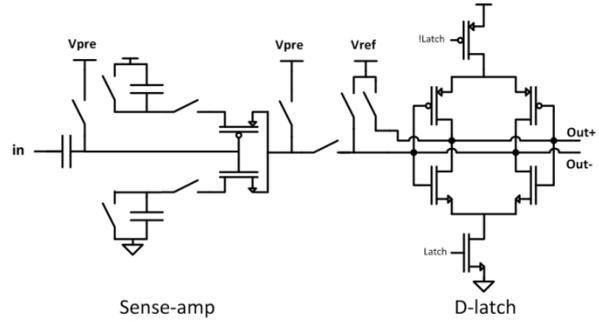


Fig. 7. Preamp and latch at bottom of column line.

#### Scaling

Scaling of the design results using normal scaling rules was performed [7-9]. Power per column  $P_{col}$  scales simply as:

$$P_{col} = V_{DD} \times I_{col} = V_{DD} \times C_{col} \times \frac{\Delta V}{\Delta t} \quad (5)$$

where  $I_{col}$  is the column bus bias current, and  $\Delta V$  is the voltage swing on the column bus.  $C_{col}$  is the total capacitance on the column bus given approximately by:

$$C_{col} = J_V \times C_{sg} \propto J_V \times \frac{W \times L_{ov}}{t_{ox}} \quad (6)$$

and scales with the overlap capacitance of the access transistor. The comparator is a non-continuous-time comparator and only consumes dynamic power. Its power  $P_{comp}$  scales as:

$$P_{comp} \propto C \times V_{DD}^2 \times f \quad (7)$$

The results of scaling are shown below in Fig. 8.

Process	$V_{DD}$	Jot array	Column Speed	Column power	Comp power	Total	Array Power
<b>CURRENT DESIGN</b>							
0.18 $\mu\text{m}$	1.8V	0.001 Gjots (1k X 1k)	1Mj/s (1000fps)	0.71 $\mu\text{W}$	1.28 $\mu\text{W}$	1.99 $\mu\text{W}$	1.99mW
0.18 $\mu\text{m}$	1.8V	0.1 Gjots (10k X 10k)	1Mj/s (100fps)	6.44 $\mu\text{W}$	1.28 $\mu\text{W}$	7.72 $\mu\text{W}$	77.2mW
<b>SCALED DESIGN</b>							
0.18 $\mu\text{m}$	1.8V	0.1 Gjots (10k X 10k)	10Mj/s (1000fps)	64.4 $\mu\text{W}$	12.8 $\mu\text{W}$	77.2 $\mu\text{W}$	772mW
45nm	1.1V	1 Gjots (24k X 42k)	24Mj/s (1000fps)	57 $\mu\text{W}$	2.9 $\mu\text{W}$	59.9 $\mu\text{W}$	2.5W
22nm	0.8V	1 Gjots (24k X 42k)	24Mj/s (1000fps)	20 $\mu\text{W}$	0.74 $\mu\text{W}$	20.74 $\mu\text{W}$	0.87W
45nm	1.1V	10 Gjots (75k X 133k)	75Mj/s (1000fps)	553 $\mu\text{W}$	9 $\mu\text{W}$	562 $\mu\text{W}$	75W
22nm	0.8V	10 Gjots (75k X 133k)	75Mj/s (1000fps)	197 $\mu\text{W}$	2.3 $\mu\text{W}$	199.3 $\mu\text{W}$	26.5W

Fig. 8. Projected power in scaled designs.

As mentioned above, array scan rates of 480 or 960 fps are desired. For scaling, we adopted 1000 fps as a nominal target value. Using the present approach with scaled designs in 16:9 aspect ratio, array power can become large at 1 Gjot (1 Tb/s) and prohibitive at 10 Gjot (10 Tb/s). Improved architectures for power reduction are being pursued and an order of magnitude reduction in power is considered possible.

#### IV. JOT DESIGN

There are many interesting avenues to explore for jot implementation [2] and many early 1T or 2T active pixel technologies can also be considered [e.g., 10,11]. Presently we are investigating two approaches congruent with current CMOS image sensor technologies. Our basic philosophy is to change as little as possible from what industry is making today. Both investigations are at an early stage and complicated by not having an actual baseline process flow at competitive technology nodes.

The first approach is scaling a 1.5T BSI CMOS APS (shared readout) to meet the size, conversion gain and noise requirements of a jot. In this concept, we store the photoelectron(s) in a well buried under the transfer gate, with the gate surface pinned with holes in the off state to suppress dark current. Activating the transfer gate transfers signal to the FD. Reduction of FD (and SF gate) capacitance to allow 1 mV/e<sup>-</sup> conversion gain is under study.

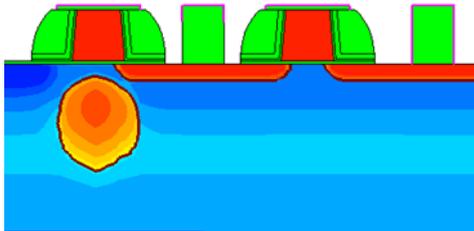


Fig. 9. Screen shot of work in progress on a BSI jot with storage under the transfer gate.

In the second approach, we are considering a bipolar transistor structure with the base fully depleted and used to store the photocarrier(s) (electron or hole). In this configuration the device also resembles a static induction transistor (SIT). The BJT is biased as an emitter follower, to follow the change in base voltage and to minimize base-emitter capacitance. The challenge in such a device is that the emitter-collector carriers want to cross at the low point in the potential barrier, whereas the photocarrier(s) wants to be at the high point, thus some sort of confinement is necessary. The lifetime of the photocarrier(s) stored in the base is also of interest. In our device the base is reset using a transfer gate to sweep out the photocarrier(s) laterally and completely.

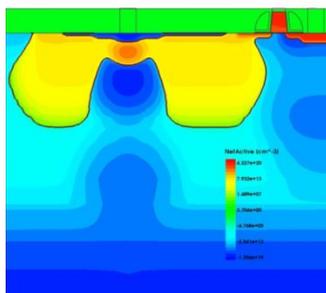


Fig. 10. BJT-type jot under investigation

#### V. CONCLUSIONS

Creating a practical realization of the Quanta Image Sensor involves a plenitude of fun problems to solve. However, initial work has revealed many of the critical challenges and is starting to yield approaches to conquer those challenges. The research is still at an early stage.

#### VI. ACKNOWLEDGMENTS

This work was supported by Rambus, Inc. and a Thayer Graduate Fellowship (DH). Discussion with Prof. Alex Hartov, Dr. Bill Richards, and Mr. Song Chen, and cooperation of Synopsys is appreciated.

#### VII. REFERENCES

- [1] E.R. Fossum, "What to do with Sub-Diffraction-Limit (SDL) Pixels? – A Proposal for a Gigapixel Digital Film Sensor (DFS)," Proc. of the 2005 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, Karuizawa, Japan, June 2005.
- [2] E.R. Fossum, "The Quanta Image Sensor (QIS): Concepts and Challenges" in Proc. 2011 Opt. Soc. Am. Topical Meeting on Computational Optical Sensing and Imaging, Toronto, Canada July 10-14, 2011.
- [3] S. Burri and E. Charbon, "SPAD Image Sensors: From Architectures to Applications," in Imaging and Applied Optics Technical Papers, OSA Technical Digest (online) (Optical Society of America, 2012), paper ITu4C.1.
- [4] F. Yang, Y. M. Lu, L. Sbaiz, M. Vetterli, "Bits From Photons: Oversampled Image Acquisition Using Binary Poisson Statistics," IEEE Trans. Image Processing, 21(4), pp. 1421-1436 (2012).
- [5] E.R. Fossum, "Application of Photon Statistics to the Quanta Image Sensor," in Proc. 2013 Int. Image Sensor Workshop (IISW), Snowbird, Utah USA, June 12-16, 2013.
- [6] K. Kotani, T. Shibata, and T. Ohmi. "CMOS charge-transfer preamplifier for offset-fluctuation cancellation in low-power A/D converters." IEEE J. Solid-State Circuits, 33(5) pp. 762-769, (1998).
- [7] Y. Taur, et al. "CMOS scaling into the nanometer regime," Proc. IEEE 85(4) pp. 486-504 (1997).
- [8] S. Borkar, "Design challenges of technology scaling," IEEE Micro 19(4) pp. 23-29 (1999).
- [9] D. Duarte, et al. "Impact of scaling on the effectiveness of dynamic power reduction schemes," in Proc. 2002 IEEE Int. Conf. on Computer Design: VLSI in Computers and Processors, 2002., pp. 382-387.
- [10] N. Tanaka, T. Ohmi, and Y. Nakamura. "A novel bipolar imaging device with self-noise-reduction capability." IEEE Trans. Electron Devices, 36(1), pp. 31-38 (1989).
- [11] A. Yusa, et al. "SIT image sensor: design considerations and characteristics." IEEE Trans. Electron Devices, 33(6) pp. 735-742 (1986).