Image Sensor with General Spatial Processing in a 3D Integrated Circuit Technology

Viktor Gruev†, Jan Van der Spiegel
Dept. of Electrical & Systems Engineering
University of Pennsylvania
Philadelphia, PA, USA
{vgruev, jan} @seas.upenn.edu

Ralf M. Philipp†, Ralph Etienne-Cummings
Dept. of Electrical & Computer Engineering
Johns Hopkins University
Baltimore, MD, USA
{rphilipp, retienne} @jhu.edu

***Abstract*—An architectural overview of an image sensor with general spatial processing capabilities on the focal plane is presented. The system has been fabricated on two separate tiers, implemented on silicon-on-insulator technology with vertical interconnect capabilities. One tier is dedicated to imaging, where photosensitivity and pixel fill have been optimized. The subsequent layers contain noise suppression and digitally controlled analog processing elements, where general spatial filtering is computed. The digitally controlled aspect of the processing unit allows generic receptive fields to be computed on read out. The image is convolved with four receptive fields in parallel. The chip provides parallel readout of the filtered results and the intensity image.**

I. INTRODUCTION

Biologically inspired sensors have traditionally been implemented in standard CMOS technologies. Vision has been one of the most active bio-inspired researched areas, and various systems-on-a-chip have been created. One of the main shortcomings in these systems is mapping biological functions, which are constructed in the 3D world, onto a planar 2D Si structure. For example, a silicon retina reported by [1], models the spatiotemporal processing of the five different layers of cells in the human retina. The penalty of mapping the functionality of these 3D cell layers onto a 2D integrated circuit is extremely large photo pixels. One can envision, that a more suitable approach would include direct one-to-one mapping of biological layers into stacked silicon circuits. Recent advancements in 3D integration of silicon dies [2] allow such implementation in more elegant manner. For example, a single layer can be dedicated to imaging, where noise and pixel sensitivity is optimized. The functionality of horizontal, bipolar, amacrine and ganglion cells can be mapped and optimized on subsequent layers, which have direct vertical interconnection with a single or group of photoreceptors. The power of digital programmability would then allow another important extension in such a vision sensor. Parallel analog processing on multiple layers, in combination with fast, programmable, digital circuitry, could allow the creation of image processing architectures with unprecedented capabilities.

Stacked CMOS technology or 3D integration has been implemented primarily in silicon-on-insulator (SOI) or silicon-on-sapphire (SOS) processes. The optically transparent substrates allow vertical alignment of multiple dies economically feasible. The relatively small thickness of the substrate has been beneficial in the 3D integration process, alleviating some of the power concerns. On the same account, the thin substrate does not provide an optimum medium for light absorption. The typical absorption length for visible light in silicon is many microns, far greater than the thickness of the silicon film in typical modern fully-depleted (FD) silicon-on-insulator or silicon-on-sapphire processes. Previous works have focused on achieving respectable levels of quantum efficiency by using lateral PIN photodiodes with large photosensitive intrinsic regions, while balancing the tradeoff between quantum efficiency and pixel area [3], [4].

The use of an FD-SOI process permits the use of backside illumination, where the incident light is directed towards the underside of the photosensitive chip. The incident light thereby bypasses any obstructing interconnect or silicide.

II. SYSTEM OVERVIEW

The 3D imaging system is split into two main architectural blocks. The first block, the imaging array, is composed of a 100×50 pixel block of active pixel sensors (APS), together with scanning registers and current conveyor circuitry. These circuits, excluding the current conveyor, are placed on the first tier, where maximizing photosensitivity and pixel fill factor is of primary concern. The second architectural block, residing on the second tier, is dedicated to noise suppression and spatial image processing. This tier is composed of a noise suppression correlated double sampling unit (CDS), analog current memory cells and a digitally controlled analog scaling unit. The first tier allows for sequential readout of the entire pixel array. Once a pixel is addressed, both integrated and reset photocurrents are presented to the CDS unit in a sequential manner. The CDS unit reduces the mismatch in pixel output current due to transistor threshold variations of the pixel read out transistor;

† Equally contributing authors

it also reduces the pixel *kTC* and 1/*f* noise. The noise suppressed current is then stored in a current memory array. Block parallel access of neighborhood of memory cell is possible in order to perform spatial filtering in the analog scaling unit. Four convolved images in parallel with the intensity image are presented outside the chip. Although a single analog processing unit is used in this architecture, the 3D integration can allow parallel spatiotemporal processing with multiple analog processing units residing in the subsequent tiers.

III. IMAGER

The proposed chip uses an APS pixel that occupies 10 × 10μm^2 of area and has a PIN photodiode area of 51μm^2. The fabrication process used, a 0.18μm FDSOI CMOS process, limited the area of the intrinsic region of the photodiode to approximately 15% of the pixel area. The maximum density of polysilicon limited the area of the intrinsic region. Larger intrinsic areas would be possible with the addition of a silicide block, or with less restrictive polysilicon density constraints.

The PIN photodiode was fabricated in an annular "doughnut" shape to avoid edge effects, which can potentially increase dark current. Fig. 1 shows a simplified layout of the PIN photodiode (without contacts and metallization). The central N+ region of the diode, located in the center of the doughnut, was coated with silicide (required by the fabrication process). The intrinsic region of the PIN photodiode was fabricated using polysilicon to mask the silicide from the silicon island. The P+ region of the diode, also coated with silicide, surrounds the intrinsic doughnut. The polysilicon extension is used to provide a DC bias voltage to prevent the silicide shield from floating at an undesirable voltage level.

The imager is back-illuminated through 600nm of SiO_2. No metal interconnect, silicide or contacts reside between the back surface of the imaging area and the PIN photodiode. Aluminum metal sheets were placed underneath the photodiodes ("above" the layout in Fig. 1) to reflect any light not absorbed by the photodiode, polysilicon, or silicide back to the diode.

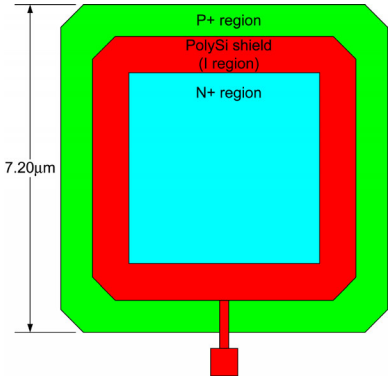

Figure 1. PIN photodiode layout (back-illuminated). The PolySi acts as a silicide block.

The APS pixel schematic is shown in Fig. 2. The PIN photodiode's photocurrent is integrated over a frame to produce the photodiode voltage V_p. The APS converts V_p into an output current I_p. Keeping the pixel's output (column) voltage V_c constant about 100 to 200mV below V_{dd} allows M2 to act as a linear transconductor. V_c is held at a fixed voltage using a first-generation current conveyor (CCI+), also shown in Fig. 2. The CCI's r_{in}, approximately 100Ω, must be sufficiently low as to not deteriorate the linearity of the pixel's transconductance; this is achieved by making r_{in} an order of magnitude lower than the pixel's row select transistor M3's series impedance (5kΩ). A more complete analysis of this pixel's operation can be found in [5] and [6].

The pixel design allows implementation of simple spatial processing on focal plane. Linear summation of pixel outputs (i.e. spatial averaging) can be performed by turning on multiple pixels; linear scaling can be performed by changing V_{ref} (approximately equal to V_c). Inactive columns are tied to V_{ref}; switch S_C (one per column) connects columns either to the CCI or to V_{ref}, as shown in Fig. 2. Multiple columns can be selected simultaneously when doing spatial summation.

A 100 × 50 pixel array was created. Latched scanners allow fully random x-y addressing of both reset and select signals, while simplifying serial scanning. The pixel's independent x and y reset transistors allow fully random x-y (per pixel) resetting.

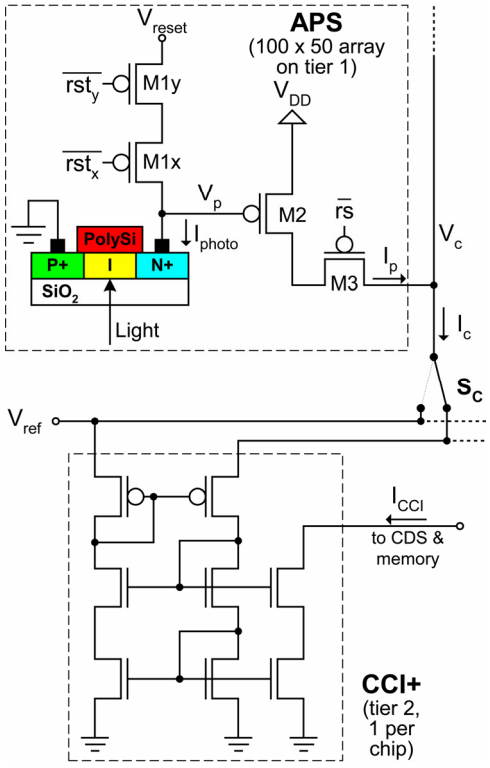

Figure 2. Pixel and CCI+ schematics

IV. PROCESSING

A block diagram of the image processing elements on tier 2 is presented in Fig. 3. The image processing circuitry on this tier is composed of a correlated double sampling (CDS) noise suppression unit, a current memory array, access registers for the memory array and a digitally controlled analog scaling unit.

The functionality of the noise suppression circuitry is to remove photo current threshold variations of the pixel read out transistor and to reduce APS 1/*f* noise. These variations are easily cancelled when dealing with a linear output (current or voltage) from the pixel. Therefore, the photo pixel read out transistor is operated in linear mode, leading to linear correlation between light intensity and output current and easy incorporation of CDS circuitry at the read out.

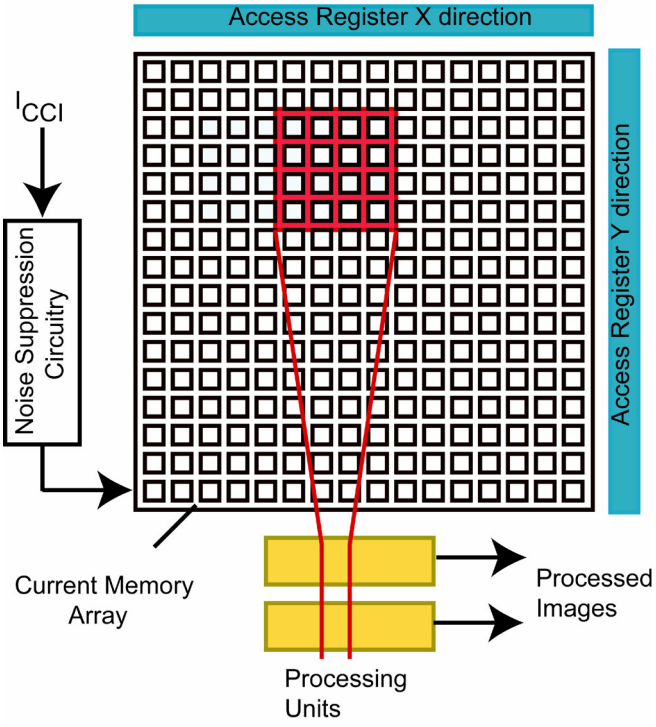

Figure 3. Tier 2 image processing circuitry

The CDS circuitry is based on the I^{2}C current memory cell described by Hughes *et al.* [7]. The memory cell is composed of a coarse sub-memory and a fine sub-memory cell (Fig. 4). The coarse sub-memory cell is composed of transistor M1, capacitor C1 and switch transistor S1, while the fine sub-memory cell is composed of transistor M2, capacitor C2 and switch transistor S2. During the memorization stage of the coarse memory cell, charge injection errors dependent on the input current level are introduced on capacitor C1. These signal dependent charge injections are memorized in the fine memory cell, capacitor C2, and subtracted from the coarse memory cell. SPICE simulations indicate that the final memorized current can replicate the original current with 10-bit accuracy.

The noise suppression of the photo pixel is performed in two steps. Initially the output current (I_{CCI} after integration) I_{Cint} is memorized in the current memory cell (2). The pixel is then reset and the output, I_{Crst} (I_{CCI} when $V_p = V_{reset}$), defined by (3), is automatically subtracted from I_{Cint}. The final current output is the difference between I_{Cint} and I_{Crst} and is given by (4).

$$V_{sd2} \approx V_{dd} - V_C \approx V_{dd} - V_{ref} \tag{1}$$

$$I_{Cint} = \beta_2 \left[\left(V_{dd} - V_p - |V_{tP}| \right) V_{sd2} - \frac{V_{sd2}^2}{2} \right] \tag{2}$$
$$\beta_2 = \mu_{eff2} C_{OX} W_2 / L_2$$

$$I_{Crst} = \beta_2 \left[\left(V_{reset} - |V_{tP}| \right) V_{sd2} - \frac{V_{sd2}^2}{2} \right] \tag{3}$$

$$I_{CDS} = I_{Cint} - I_{Crst} = \beta_2 V_{sd2} \left(V_{dd} - V_p - V_{reset} \right) \tag{4}$$

The final current output I_{CDS} does not depend on the threshold voltage variations of the read out pixel transistor and is linearly proportional to the photodiode voltage, V_p. V_{sd2} (1) should be approximately 100 to 200mV. It can be seen that I_{CDS} can be linearly scaled by adjusting V_{ref}.

The noise suppressed current I_{CDS} is memorized in one of the storage elements in the memory array (Fig. 4). The memory elements are addressed via the access registers in the X and Y direction (Fig. 3). The memory cells have one set of registers, which allow write access, and another set of registers, which allow read access. For simplicity, these registers are not shown in Fig. 3 or Fig. 4. The memory cell is composed of two parts. The first part is the standard I^{2}C current memory cell, composed of coarse and fine sub-memory units [7]. Although the chip area occupied by the I^{2}C memory cell is increased compared to a standard single transistor memory cell, the precision of the memory cell is improved to 8 bits. The decreased performance of the memory element compared to the CDS memory element is largely due to the smaller capacitance used in the earlier one. The second part of the memory cell is composed of read out transistors M8 through M10 and corresponding switch transistors S1X through S3X, which allow three copies of the memorized current to be output on common horizontal buses. These three copies allow access to three distinct memory elements per row of the processing unit. Each one of the memory currents are then scaled in the processing unit according to the convolution kernel specifications. Since three rows of the current memory array can be accessed in parallel, and each row contains three distinct memory currents, a total of nine distinct currents are available outside the memory array. The concept of block-parallel access of several memory elements is similar to the block-parallel access of multiple photo pixels for spatiotemporal image

processing at the focal plane [8]. This concept has great advantages for single-die image processing architectures; however, it does not present an optimal addressing scheme for a 3-D integrated circuit implementation. Further algorithmic optimization will be needed for efficient 3-D spatiotemporal image processing.

The processing unit is a digitally controlled analog processing unit, consisting of four sub-units. The sub-units are identical in structure and consist of a digital control memory with 45 bits per sub-unit and analog scale and add circuits. Each of the nine input currents is first mirrored four times, and then passed to the sub-processors for individual computation. The digital memory assigns a 5-bit signed-magnitude control word per current, specifying the kernel coefficient for each current [8]. The coefficient can vary between +/-3.75 in increments of 0.25 (31 possible coefficients). The appropriate weight factors will vary depending on the given mask of interest. After each current is weighted by the appropriate factor, all currents are summed together to produce the desired processed image.

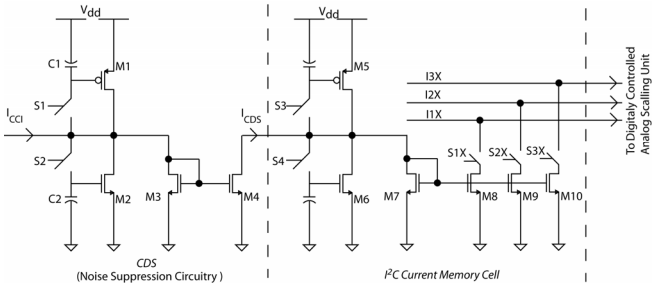

Figure 4. Detail of circuitry on second tier

V. Algorithmic Spatial Filtering

In order to fully realize the power of the parallel processing capabilities of the image sensor, the size of the pixel/memory groups are kept small. Minimizing the number of memory elements per group maximizes the number of independent kernels that can be implemented in parallel. Ideally, if every pixel/memory value for a given neighborhood is available to the processing unit, the kernels can be completely general (i.e. every pixel can be given its own coefficient). This is not possible in this architecture without using a large number of pixel current copies in the memory element. This would result in a large memory size and spacing, due to large number of current routing lines, making such a completely general implementation impractical. A trade-off between generality and memory size must be taken into an account in designing the memory array. Hence, our design allows for computation of variable sizes of kernels based on a 3 x 3 canonical model, where nine unique coefficients can be applied to the nine pixels. The distribution of these coefficients depends on the configuration of the selection and routing switches (registers).

In general, the computation performed by the processing unit is given in equation (5), where the window of convolution is *NxM* elements, *J(x,y)* is the memory current in the window of convolution and *a(x,y)* is the kernel coefficient.

$$J_{out}(i,j) = \sum_{x=i-N/2}^{i+N/2} \sum_{y=j-M/2}^{j+M/2} a(x,y) J(x,y) \qquad (5)$$

$$\text{where } a(x,y) = n/4 \quad \forall n \left| \{ n \in \mathbb{Z}, -15 \le n \le 15 \} \right.$$

VI. Conclusion

We have presented an architectural overview of an active pixel sensor with general spatial image processing at the focal plane implemented in an emerging 3D integration technology. Many of the image processing concepts are an extension of previous work on focal plane image sensors implemented in standard CMOS process. Although the true parallelism available in 3D integration system is not fully explored, this system will be used a guide to explore the imaging and processing capabilities at the focal plane in this technology.

Acknowledgment

This work has been possible by a research grant from MITLL, AFOSR grant #FA9550-05-1-0052, NSF grant #0428042, and ONR grant #140010562 NCSU provided the 3D design kit.

References

[1] K A Boahen, "A retinomorphic vision system," *IEEE Micro*, vol. 16, n. 5, pp. 30-39, Oct. 1996.

[2] V. Suntharalingam, et al., "Megapixel CMOS image sensor fabricated in three-dimensional integrated circuit technology," *ISSCC Dig. Tech. Papers*, pp. 356-357, Feb. 2005.

[3] A. Afzalian and D. Flandre, "Physical modeling and design of thin-film SOI lateral PIN photodiodes," *IEEE Trans. Electron Dev.*, vol. 52, no. 6, pp. 1116-1122, June 2005.

[4] C. Xu, C. Shen, W. Wu, and M. Chan, "Backside-illuminated lateral PIN photodiode for CMOS image sensor on SOS substrate," *IEEE Trans. Electron Dev.*, vol. 52, no. 6, pp. 1110-1115, June 2005.

[5] R.M. Philipp and R. Etienne-Cummings, "A 1V current-mode CMOS active pixel sensor," *ISCAS 2005*, pp. 4771-4774, May 2005.

[6] V. Gruev, R. Etienne-Cummings and T. Horiuchi, "Linear current mode imager with low fix pattern noise," *ISCAS 2004*, vol. 4, pp. 860-863, May 2004.

[7] J.B. Hughes and K.W. Moulding, "S^3I: the seamless S^2I switched-current cell", *ISCAS 1997*, vol. 1, pp. 113-116, June 1997.

[8] V. Gruev and R. Etienne-Cummings, "Implementation of steerable spatiotemporal image filters on the focal plane," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, pp. 435-440, 2002.