13’000 FPS Vision System-on-Chip with Mixed-Signal Compressed Sensing

Jens Döge, Christoph Hoppe, Peter Reichel, Nico Peter, Andreas Reichel and Christian Skubich
Fraunhofer Institute for Integrated Circuits IIS / EAS, Münchner Straße 16, 01187 Dresden, Germany
E-mail: jens.doege@eas.iis.fraunhofer.de, Phone: +49 351 45691-320

***Abstract*—This paper presents a monolithic high-speed VSoC (Vision-System-on-Chip) with three software-programmable 16-bit ASIPs (application-specific instruction-set processors), a 1024-fold column-parallel data path of charge-based convolution functionality, freely configurable A/D conversion, 8-bit processor elements with 128 bytes of RAM each, and asynchronously compressing output of sparse column data. While 3D integration allows for combining a sensor field in optimal technology with a digital processing chip, it increases chip development, manufacturing and testing costs. In this design, the classical monolithic integration approach is pursued to achieve a single-chip solution with good fill factor and competitive performance in a classical 180 nm 1P6M CIS technology. To demonstrate the advantageous compressed sensing approach for fast and low-latency image processing, an algorithm for laser sheet-of-light triangulation was implemented.**

***Index Terms*—image sensor, vision-system-on-chip, compressed sensing, sheet-of-light, laser triangulation, VSoC, PE, ASIP**

I. Introduction

Fast and low-latency image acquisition and processing are essential to use CMOS image sensors as measuring devices in industrial applications and for image-based process control. Depending on the resolution, fill factor and speed requirements, so-called Vision Systems-on-Chip (VSoC) may operate pixel-parallel [1], column-parallel [2], serially on the sensor chip or on an external basis. Complex processing close to the pixel can considerably reduce the fill factor. Column-parallel processing not only allows for the use of more complex A/D converters, but is also a good compromise between fill factor, complexity of the digital processing elements (PE) and speed.

3D integration opens up a wide range of possibilities for combining a sensor field in optimal technology with a digital processing chip [3], [4]. Disadvantages include high development effort for the two chips as well as manufacturing and testing effort. Thus, economic use of this technology usually requires high production volumes.

In this paper, the classical monolithic integration approach is pursued to achieve a single-chip solution with good fill factor and competitive performance in a classical 180 nm 1P6M CIS technology.

II. System Architecture

The heart of this VSoC is the pixel cell shown in figure 1 based on a standard 8T pinned pixel with various extensions. The most important one relates to the output of the brightness signal as charge values equivalent to the corresponding irradiation. The internal voltage representing brightness information charges the output capacitor $\mathtt{C_{PE}}$ via one of the two output transistors $\mathtt{T_{AP}}$ and $\mathtt{T_{AN}}$ as positive (`EnAp`) or negative (`EnAn`) charge, respectively. This way, multiples of charge values from multiple pixel lines can be written to the respective column lines to sum them up – either simultaneously or continuously.

Line voltages must be kept as constant as possible for accurate charge output and summation. The analog readout path (see figure 2) achieves this with two charge amplifiers `ampdp` and `ampdn`. Over-charging is prevented by regularly reducing the common mode component via DC-compensation.

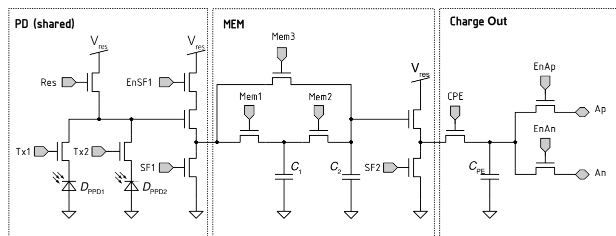

Figure 1. Schematic of the pixel cell with two photodiodes, an internal memory and switched charge-based output.

Known charges from pulsed current sources are used to A/D convert the differential charge. The specific algorithm, e.g. single- or dual-slope and the resolution between 1 and 12 bit with linear or nonlinear mapping is entirely software-defined. The results are further processed in the column-parallel single-instruction multiple-data processing element (SIMD-PE) shown in figure 3.

Each PE consists of an 8-bit ALU with 8 working registers and a 1-bit flag ALU with 8 working flags. Both support arithmetic, logical and shift operations, e.g. `ADD`/`SUB`, `SHL`/`SHR`, `AND`/`NAND`. The operands are provided according to a three-address logic. Each PE has access to the 8 registers of the left (`r0l..r7l`) and right

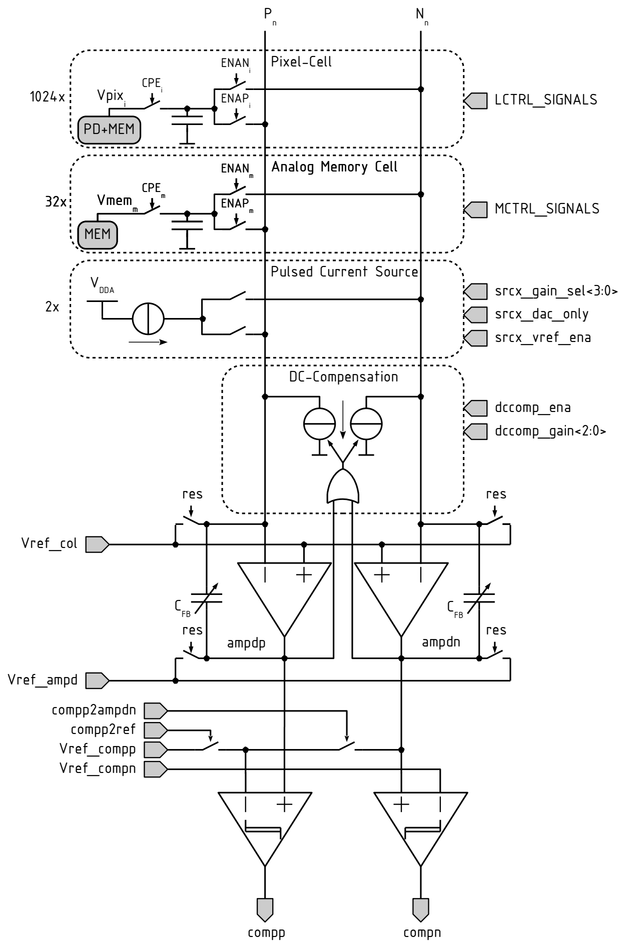

Figure 2. Analog readout path for a pixel column.

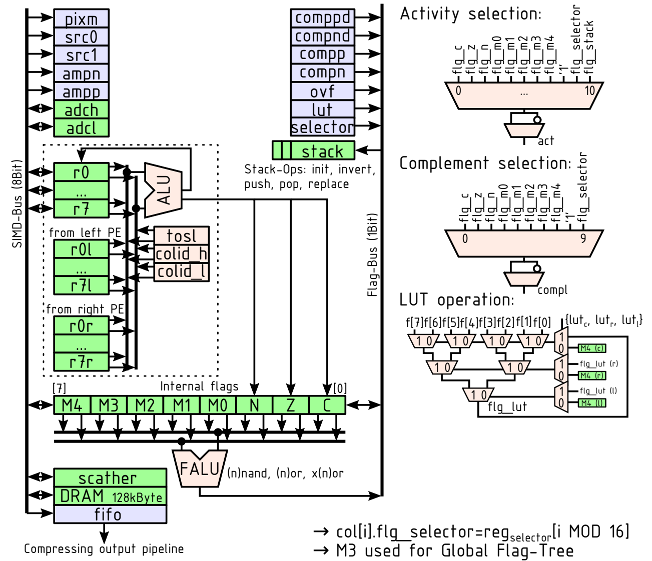

Figure 3. Single instruction multiple data processor element (SIMD PE).

(r0r..r7r) neighboring columns as well as 128x8 bit DRAM and a binary look-up table (LUT) for linking flags to neighboring columns. Besides the working registers and flags, there are further ones for analog data path calibration (src0, src1, ampp, ampn), setting the address of the analog memory (pixm), reading/writing the ADC result (adcl, adch, comp*, ovf), fixed selection of columns (selector) and retrieving the LUT result (lut). A flag stack is used to execute conditional statements in the SIMD array. Each PE's activity status can be set to any flag or its inverse using a multiplexer. Either an ALU operation or its inverse can be executed depending on a local flag (complement selection). The scather unit can turn neighboring PEs into macro PEs of size 4, 8 or 16 to exchange data over up to 16 PEs within one clock cycle. Data is output via an asynchronous pipeline with local FIFOs, which may operate in 8, 16, 24 or 32-bit mode.

The features provided by the various analog and digital functional units of the VSoC are abstracted for use in arbitrary image processing algorithms by the instruction sets of individual processors of an integrated, multi-ASIP (application specific instruction set processor) based control unit. To meet the parallelism requirements, they are grouped and distributed to three independent ASIPs: The SIMD ASIP for controlling ADC, the digital PEs and the output pipeline, the LCTRL ASIP (Line Control) for controlling the analog pixel matrix and memory and the GLB ASIP (Global Control) to communicate with the surrounding logic. They all consist of a stack-based processor core with associated program memory, methods for data input and output and for synchronization among each other, a connection to the integrated Network-on-Chip (NoC) and a scratch-pad memory. The instruction sets are optimized to maximize flexibility and minimize latency for analog component control. In addition to data exchange between the ASIPs, the integrated NoC is an effective testing and debugging option and enables an image acquisition and processing algorithm to control further peripheral components.

The overall architecture of the VSoC is shown schematically in figure 4 and as a die photo in figure 5. The different digital interfaces LVDS, SPI, JTAG and GPIO provide the connection to the sensor peripherals depending on the application.

III. Application and Compressed Sensing

To demonstrate the advantages of highly parallel compressive image processing on this VSoC, the sheet of light laser triangulation method was implemented as an example. Assuming a single visible laser line profile in a SoL set-up, the input signal in each column is highly redundant. Making use of this redundancy, the VSoC can effectively compress the amount of data per laser line profile by a factor of 1024. By scanning the image at a significantly lower, spatial sampling rate (Compressed Sensing), profile rates of 13 kHz can be achieved.

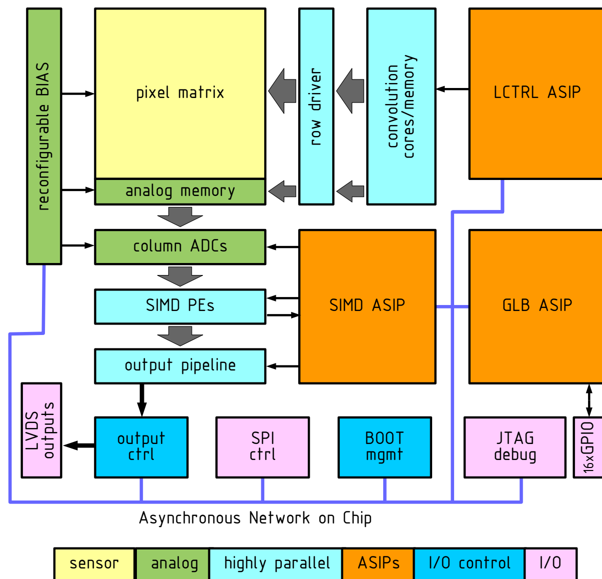

Figure 4. Overall architecture of the Vision System-on-Chip.

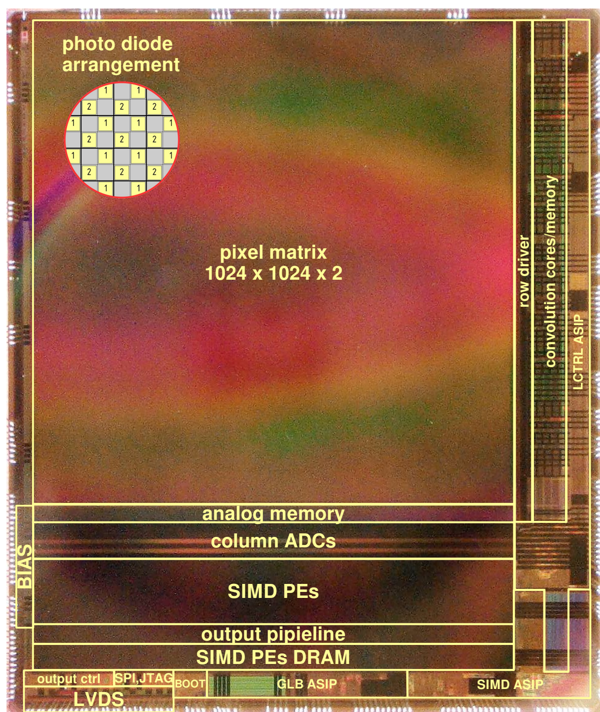

Figure 5. Chip microphotograph of the Vision System-on-Chip and checker board photo diode arrangement.

The implemented SoL algorithm, as depicted in I, works in two steps: First, the image field is scanned at a low resolution using a lowpass 1st derivative filter of size C (e.g. $C = 32$) at step size $C/2$. Minimum and maximum positions of the convolved signal determine the position of the laser line in the respective column. This value, which is accurate within $C/2$ pixels, becomes a column-specific offset for the high-resolution scan. In this step, the input signal is convolved with a Savitzky-Golay 1st derivative filter of size 7 at step size 1. The laser line peak is then determined within $1/8$ pixels by linear interpolation around the zero-crossing of the convolved signal.

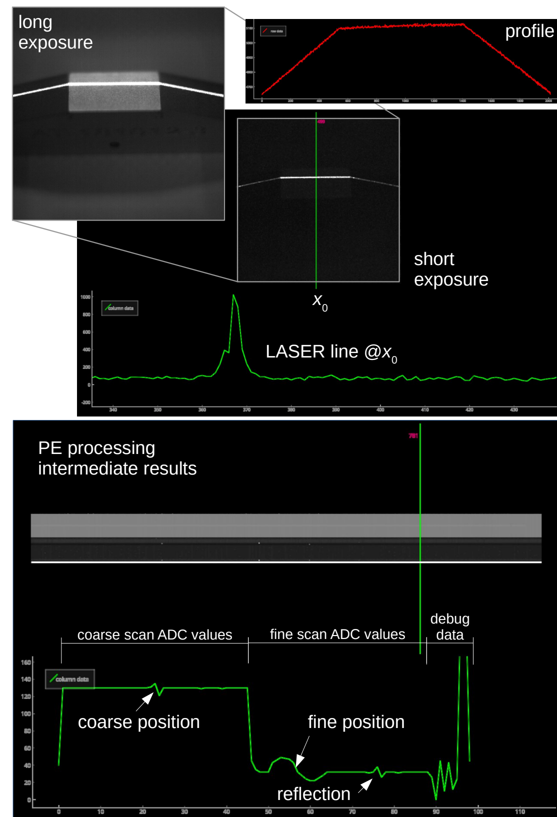

Figure 6. Determination of the position of the laser line.

IV. Conclusion

In this paper a monolithic software-programmable high-speed VSoC for compressed sensing has been presented. Using the VSoC as an example for the acquisition and processing of sheet of light profiles with a resolution of 2048 points along the line and more than 8000 (subpixel) levels, about 118 GOps (77 GOps analog MA and 41 GOps 8-bit digital) at a system clock frequency of 60 MHz and a total profile rate of 13 kHz have been achieved. Table I compares the novel architecture and the achieved specs with other designs.

This VSoC was also tested in further applications, e.g. for presence detection and the acquisition and analysis of white-light time domain interferometry data.

Acknowledgment

The authors gratefully acknowledge the cSoC3D project of the 3Dsensation alliance (https://www.3d-sensation.de/en.html), funded by the German Federal

Table I
SPECIFICATION AND COMPARISON WITH OTHER DESIGNS. ALL SPECIFICATIONS FOR THIS WORK WERE MEASURED.

Feature	***This work***	***Yamazaki et. al. [4]***	***Millet et. al. [3]***
Fabrication Process	180 nm 1P6M CIS	90 nm 1P4M CIS 40 nm 1P7M Logic	130 nm 1P6M 130 nm 1P6M
Integration	single chip FSI	stacked BSI	stacked BSI
Supply Voltages	3.3 V (analog) / 2,2 V (pixel) / 1,8 V (digital)	3.3 V / 2.9 V / 1.8 V / 1.1 V	
Chip Size	11 mm (H) x 13 mm (V) 143 mm²	5.67 mm (H) x 4.5mm (V) 25.9 mm²	113 mm²
Focal Plane Size	8.96 mm (H) x 8.86 mm (V)	4.54 mm (H) x 3.42 mm (V)	15.55 mm (H) x 9.22 mm (V)
Number of Effective Pixels	2 x 1008 (H) x 1008 (V) 2016 (H) x 1008 (V) (eff. SoL*)	1296 (H) x 976 (V)	1024 (H) x 768 (V)
Pixel Size	8.75 µm x 8.75 µm (macro pixels) 6.2 µm x 6.2 µm (eff. CB**) 4.375 µm (H) x 8.75 µm (V) (eff. SoL*)	3.5 µm x 3.5 µm	12 µm x 12 µm
Fill Factor	23% (no micro lenses)		
Frame Rate	50 Hz at 2 MPix GS 10 bit	120 Hz at 1.27 MPix 10 bit	
Processing Frame Rate	13 kHz (full frame) SoL* mode (8000 subpixel levels)	1 KHz @ 0.31 MPix 4 bit	0.34 kHz @ 0.78 MPix 11 bit 5.5 kHz @ 0.050 MPix 9 bit
Power Consumption	1 W	353 mW @ 0.31 MPix 4 bit 1 KHz with sensing	720 mW @ 0.050 MPix 9 bit 5.5 kHz
SNR	42.1 dB		
Dynamic Range	59.9 dB	80 dB	54 dB
Operating Frequency	ASIPs / PEs: 60 MHz (max. 100 MHz) LVDS: 500 MHz DDR	CPU: 108 MHz	
ASIPs	16 bit architecture Line Control SIMD Control Global Control		
PE	1024 / 8 bit architecture	1304 / 4 bit architecture	3072
Processing Power	117 GOps @ 60MHz / analog + 8b (SoL* mode)	140 GOps @ 108 MHz / 4b	61 GOps @ 80 MHz / 8b
Analog Memory	32 x 1024 cells		
ASIP / CPU Memory	3 x 256 x 16 bit stack SRAM 3 x 8 kB instruction SRAM 3 x 4 kB scratchpad DRAM	7 kB instruction 3 kB line 12 kB template	65 kB instruction
PE / Data Memory	128 kB DRAM	165 kB	73 kB + 98 kB data

***SoL:**
Sheet-of-light triangulation algorithm

****CB:**
Checkerboard pixel arrangement

Ministry of Education and Research, for the financial support.

REFERENCES

[1] S. J. Carey, D. R. Barr, B. Wang, A. Lopich, and P. Dudek, "Locating high speed multiple objects using a scamp-5 vision-chip," in *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*. IEEE, 2012, pp. 1–2.

[2] J. Döge, C. Hoppe, P. Reichel, and N. Peter, "Megapixel HDR Image Sensor SoC with Highly Parallel Mixed-Signal Processing," in *International Image Sensor Workshop (IISW)*, 2015.

[3] L. Millet, S. Chevobbe, C. Andriamisaina, L. Benaissa, E. Deschaseaux, E. Beigne, K. B. Chehida, M. Lepecq, M. Darouich, F. Guellec *et al.*, "A 5500-frames/s 85-gops/w 3-d stacked bsi vision chip based on parallel in-focal-plane acquisition and processing," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 4, pp. 1096–1105, 2019.

[4] T. Yamazaki, H. Katayama, S. Uehara, A. Nose, M. Kobayashi, S. Shida, M. Odahara, K. Takamiya, Y. Hisamatsu, S. Matsumoto *et al.*, "4.9 a 1ms high-speed vision chip with 3d-stacked 140gops column-parallel pes for spatio-temporal image processing," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2017, pp. 82–83.