

ViPro: Focal-Plane Spatially-Oversampling CMOS Image Compression Sensor

Ashkan Olyaei

Department of Electrical and Computer Engineering
University of Toronto
Toronto, ON M5S 3G4, Canada
Email: olyaei@eecg.toronto.edu

Roman Genov

Department of Electrical and Computer Engineering
University of Toronto
Toronto, ON M5S 3G4, Canada
Email: roman@eecg.toronto.edu

Abstract— The CMOS image sensor computes spatially-compressing convolutional transforms directly on the focal plane, yielding digital output at a rate proportional to the mere information rate of the video. A bank of column-parallel $\Delta\Sigma$ -modulated analog-to-digital converters (ADCs) performs distributed column-wise focal-plane oversampling of a set of adjacent pixels and concurrent weighted average quantization. The number of samples per pixel and switched capacitor sampling sequence order set the amplitude and sign of the respective pixel coefficient. Outputs of a set of adjacent ADCs are accumulated to realize a two-dimensional block matrix transform in parallel for all columns. The 3.1 mm \times 1.9 mm prototype, ViPro, containing a 128 \times 128 active pixel array and a bank of 128 hybrid algorithmic $\Delta\Sigma$ -modulated ADCs yields 4 GMACS (multiply-and-accumulates per second) computational throughput in real-time discrete cosine transform (DCT) video compression when scaled to HDTV 1080i resolution.

I. INTRODUCTION

Block matrix transforms such as discrete cosine transform (DCT) or discrete wavelet transform (DWT) are widely used in image and video compression algorithms. They reduce output data rate but require computationally expensive spatial filtering of images demanding vast computing resources and interface bandwidths for real-time operation.

Digital signal processors rely on high-throughput architectures to compute spatial weighted sums needed in block matrix transforms. They utilize a convenient digital data format, and sustain accuracy and configurability of computation. This often comes at the expense of significant area and power resources as well as limited input data rate and memory-processor bandwidth. An analog-to-digital converter (ADC) is also required to quantize the analog sensory input prior to processing.

Analog circuits perform area-efficient and low-power computation directly on the focal plane, eliminating the need for an external processor [1]. The intrinsic parallelism of imaging architectures yields high computational throughput, often beyond that of modern digital processors, allowing to perform complex video processing operations in real time. On-focal-plane spatial analog video compression yields the output data rate of an imager proportional to the mere information rate of the video sequence, not the imager resolution or frame rate. Capacitor bank implementations use charge sharing to compute weighted sum and difference [2], [3],

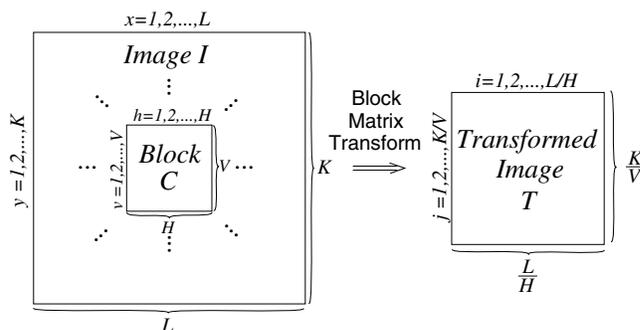


Fig. 1. An illustration of a block matrix transform computation for non-overlapping blocks.

[4] but may have limited scalability. Current-mode weighted averaging implementations [5] use zero-latency current-mode addition but employ multiple matched current mirrors at the expense of increased pixel area. Charge integration and gain-stage voltage summation [6] utilized in variable resolution imaging do not allow for weighted averaging and require additional column-parallel amplifiers. Current-mode vector-matrix multiplication [7] architectures employ floating-gate arrays for block matrix storage and achieve high power efficiency. Kernel-dependent scan-out imager architectures have been shown to reduce memory requirements in focal-plane spatial image processing [8]. A tree-based partitioning algorithm that implements adaptive compression has also been reported [9]. All of the aforementioned analog architectures require an extra analog-to-digital converter to provide the digital output.

Mixed-signal CMOS imaging and signal processing combine the benefits of both analog and digital domains [10]. We present a mixed-signal VLSI implementation of a digital CMOS imager computing block matrix transforms on the focal plane for real-time image compression. Our approach combines weighted spatial averaging and oversampling quantization in a single algorithmic $\Delta\Sigma$ -modulated analog-to-digital conversion cycle, making focal-plane computing an intrinsic part of the quantization process. The approach yields conversion time, integration area, and power dissipation comparable to those of a conventional CMOS digital imager performing no computation. The rest of this paper is organized as follows.

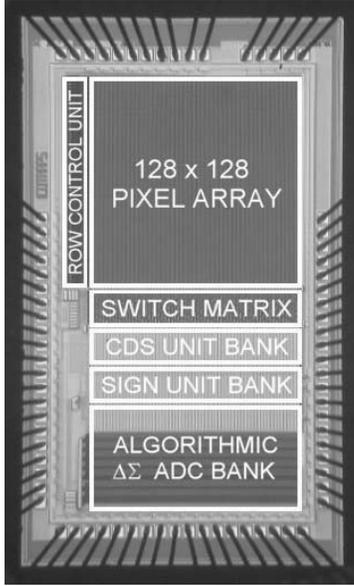


Fig. 2. Die micrograph and floor plan of the focal-plane spatially-oversampling CMOS image compression sensor. The integrated $3.1\text{mm} \times 1.9\text{mm}$ prototype was fabricated in a $0.35\ \mu\text{m}$ CMOS technology. Digital accumulation and thresholding (not shown) blocks are implemented off-chip.

Section II gives an overview of the block matrix transform method for image compression. Section III presents the architecture and circuit implementation of the image compression sensor. Section IV contains experimental results obtained from a $0.35\ \mu\text{m}$ CMOS computational imager prototype.

II. BLOCK MATRIX TRANSFORM

In image compression, block matrix transforms correlate a segment of an image with a spatial kernel in order to identify statistical redundancies. These redundancies are then eliminated by thresholding. To transform an image I into a block-transformed image T , the block matrix C is tiled vertically and horizontally across the image as illustrated in Figure 1. The block matrix is tiled in overlapping or non-overlapping fashion depending on the block matrix transform type. For the non-overlapping case shown in Figure 1, coefficients of T are obtained by computing the two-dimensional dot product of C and I at each tile location:

$$T_{ij} = \sum_{h=1}^H \sum_{v=1}^V C_{hv} I_{xy}, \quad (1)$$

$$x = h + (i-1)H, \quad i = 1, 2, \dots, \frac{L}{H}, \quad (2)$$

$$y = v + (j-1)V, \quad j = 1, 2, \dots, \frac{K}{V}, \quad (3)$$

where $C_{hv} \in \mathbb{Z}$ are the block matrix coefficients comprising a spatial kernel; L and K are the image horizontal and vertical sizes, assumed for simplicity to be multiples of the kernel dimensions H and V ; h and v are the horizontal and vertical

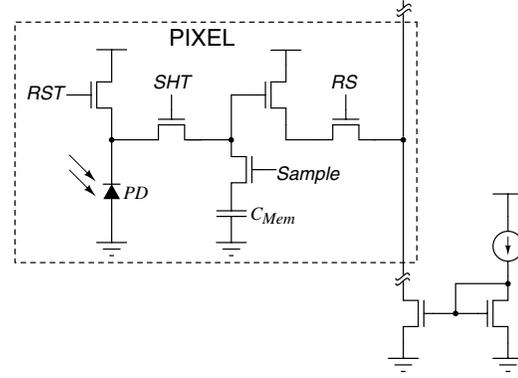


Fig. 3. Photodiode-based active pixel sensor circuit.

block matrix indices, and i and j are the indices of the block-transformed image.

The block matrix transform of the form (1) for the transformed image pixel at location $(1,1)$ can be decomposed as follows [11], [12], setting indices $i = j = 1$ and omitting them for simplicity:

$$T = \sum_{h=1}^H \sum_{v=1}^V C_{hv} I_{hv} = \sum_{h=1}^H T_h, \quad (4)$$

with partial sums

$$T_h = \sum_{v=1}^V C_{hv} I_{hv} = \sum_{v=1}^V |C_{hv}| S_{hv}, \quad (5)$$

with the sign of C_{hv} factored into the sign-transformed pixel outputs

$$S_{hv} = \text{sign}(C_{hv}) I_{hv}, \quad (6)$$

where $C_{hv} = \text{sign}(C_{hv}) |C_{hv}|$, and I_{hv} is the output of a pixel at location (h,v) . The general form of the block matrix computation remains the same for other pixel locations (i,j) as set by (2) and (3). The presented image compression sensor efficiently implements computations (6), (5), and (4) in that order, in parallel for all pixels in one row j of the transformed image T , as described next.

III. ViPRO: VLSI IMPLEMENTATION

The micrograph of the mixed-signal VLSI prototype efficiently computing the block matrix transform in (1) for $L = K = 128$ is depicted in Figure 2, with the floor plan superimposed over it.

Image acquisition is performed by a photodiode-based active pixel array with in-pixel frame buffer depicted in Figure 3, the row control circuitry, and the correlated double sampling (CDS) units, yielding the offset-compensated pixel output I_{hv} in (6).

The sign unit shown in Figure 4(a) is implemented as a switched-capacitor difference circuit. The amplifier is a single-stage common-source cascoded nMOS amplifier. The sign unit applies the sign of the coefficient C_{hv} to the input

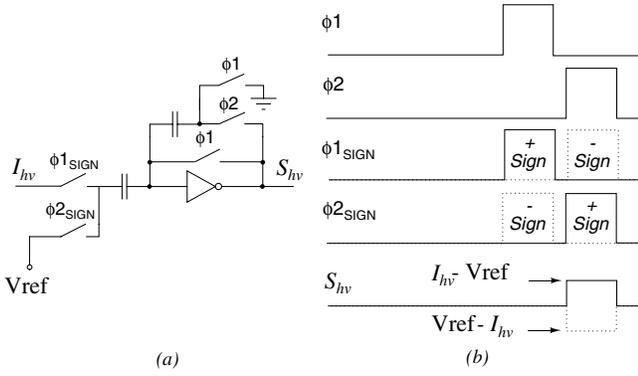


Fig. 4. (a) The sign unit circuit diagram. (b) Sign circuit timing diagram, for a continuous range of signed outputs, $V_{ref} = \min\{I\}$.

I_{hv} by selecting a switched-capacitor sampling sequence order as illustrated in the timing diagram in Figure 4(b). This directly implements equation (6).

Weighted averaging in (5) of V adjacent pixel outputs in an image column is performed by a first-order incremental $\Delta\Sigma$ -modulated multiplying ADC as depicted in Figure 5. It oversamples v -th pixel sign-transformed output, S_{hv} , $|C_{hv}|$ times for $v = 1, \dots, V$, without resetting the integrator. The coefficient $|C_{hv}|$ and the sign of C_{hv} are presented from two looped shift registers each with a period of V rows. Each coefficient is stored in a binary format and is digitally oversampled to yield its unary representation

$$|C_{hv}| = \sum_{i=0}^{N-1} c_{hv}[i], \quad (7)$$

to match the sampling mechanism of the oversampling ADC and correspondingly weight each pixel output [13]. This single quantization cycle yields \hat{T}_h , the digital representation of T_h in (5).

The computational throughput is maximized for an arbitrary block matrix transform by algorithmically resampling the modulation residue in each oversampling ADC to obtain higher resolution bits [14]. This yields a bit resolution linear in the number of conversion cycles.

The switch matrix shown in Figure 6 routes the H different time-dependent block matrix coefficients and sign signals to $\frac{L}{H}$ groups of adjacent column-parallel ADCs and sign unit circuits respectively.

A simple digital delay and adder loop performs spatial accumulation over H adjacent ADC outputs in the digital domain as they are read out to yield the digital representation of T in (4).

IV. EXPERIMENTAL RESULTS

Two-dimensional Haar wavelet transform, a block matrix transform commonly used in image compression, is chosen here as a simple example to validate the functionality of the integrated prototype. Figure 7(a) shows an image acquired by the pixel array at 30 frames per second. Image read-

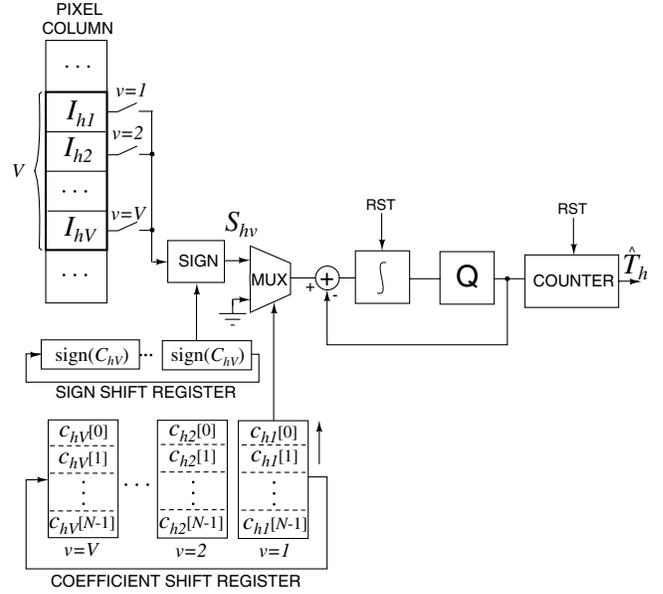


Fig. 5. Architecture of one column of the image compression sensor for $i = j = 1$.

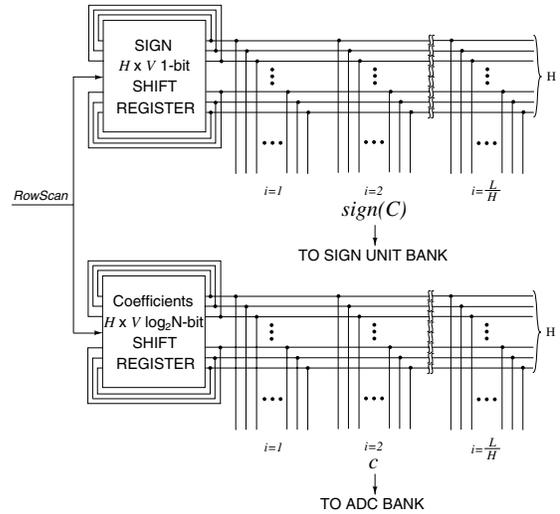


Fig. 6. Block diagram of the switch matrix.

out and computational quantization are characterized off line in two sequential steps. Figure 7(b) depicts experimentally measured two-dimensional one-, two-, and three-level Haar wavelet transforms of the original image. Figure 7(c) shows the reconstructed images of the corresponding Haar wavelet transforms. The reconstructed images of one-level Haar transform are compared in Figure 8 for various peak signal-to-noise and compression ratios.

Table I summarizes electrical and optical characteristics experimentally obtained from the 0.35 micron CMOS prototype shown in Figure 2. Based on 40 kfps measured quantizer sampling rate, when scaled to HDTV 1080i resolution, the image compression processor is projected to

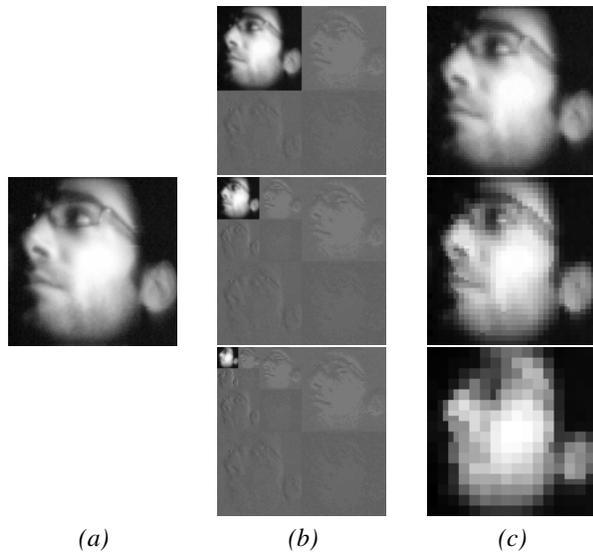


Fig. 7. (a) An image captured by the CMOS image compression sensor at 30 fps. (b) Experimentally recorded one-level (top), two-level (center), and three-level (bottom) Haar wavelet transforms of the image in (a) computed on the CMOS image compression sensor. (c) Reconstructed images for one-level (top), two-level (center), and three-level (bottom) Haar wavelet transforms for the same compression threshold. Compression ratios from top to bottom are: 5.33, 20.27, and 41.53.

TABLE I
SUMMARY OF CHARACTERISTICS

Technology	0.35 μm CMOS
Area	3.1mm \times 1.9mm
Supply Voltage	3.3 V
Array Size	128 \times 128 pixels
Pixel Size	10.45 μm \times 10.45 μm
Fill Factor	42%
Frame Rate	30 fps
Kernel Size	2 \times 2 — 8 \times 8 programmable
Throughput	4 GMACS in HDTV 1080i DCT
Optical Dynamic Range	105 dB
Dark Current	17.5 fA/pixel
ADC Power Consumption	4.3 mW
Output Resolution	8-bit

yield a computational throughput of 4 GMACS (multiply-and-accumulates per second) in discrete cosine transform with 64 \times 8 blocks at 30 frames per second.

V. CONCLUSIONS

We present a mixed-signal VLSI implementation of a digital CMOS imager computing block matrix transforms on the focal plane for real-time video compression. The approach combines weighted spatial averaging and oversampling quantization in a single algorithmic $\Delta\Sigma$ -modulated analog-to-digital conversion cycle, making focal-plane computing an intrinsic part of the quantization process. The approach yields conversion time, integration area and power dissipation comparable to those of a conventional CMOS digital imager. The experimental results obtained from a 0.35 micron 128 \times 128-pixel CMOS prototype validate the utility of the design for

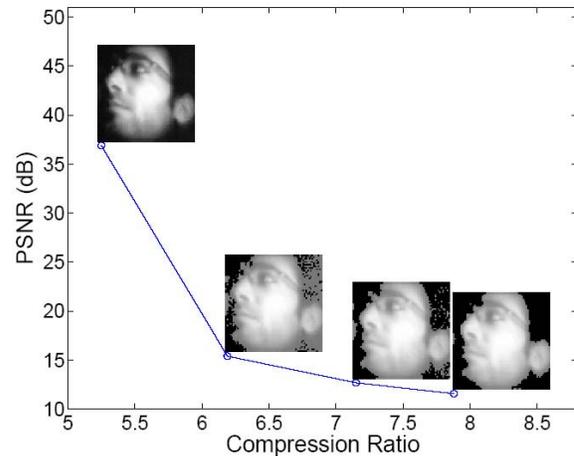


Fig. 8. Reconstructed images obtained by decompression of the experimentally computed one-level transform of the original image (top of Fig. 7(b)) for varying compression thresholds.

large-scale focal-plane signal processing.

REFERENCES

- [1] E.R. Fossum, "CMOS image sensor: Electronic camera-on-a-chip," *IEEE Trans. Electron Devices*, vol. **44**, pp 1689-1698, 1997.
- [2] S.E. Kemeny, R. Panicacci, B. Pain, L. Matthies, and E.R. Fossum, "Multiresolution image sensor," *IEEE Trans. Circuits and Systems for Video Technology*, vol. **7** (4), pp 575-583, 1997.
- [3] Q. Luo, and J.G. Harris, "A novel integration of on-sensor wavelet compression for a CMOS imager," *IEEE Int. Symp. on Circuits and Systems (ISCAS'02)*, Scottsdale, AZ, May 26-29, 2002. ISCAS'2002.
- [4] S. Kawahito, M. Yoshida, M. Sasaki, K. Umehara, D. Miyazaki, Y. Tadokoro, K. Murata, S. Doushou, and A. Matsuzawa, "A CMOS image sensor with analog two-dimensional DCT-based compression circuits for one-chip cameras," *IEEE J. Solid-State Circuits*, vol. **32** (12), pp 2030-2041, 1997.
- [5] V. Gruev, and R. Etienne-Cummings, "Implementation of steerable spatiotemporal image filters on the focal plane," *IEEE T. Circuits and Systems II*, vol. **49** (4), 2002.
- [6] Z. Zhou, B. Pain, and E.R. Fossum, "Frame-transfer CMOS active pixel sensor with pixel binning," *IEEE Trans. Electron Devices*, vol. **44** (10), pp 1764-1768, 1997.
- [7] A. Bandyopadhyay, J. Lee, R. Robucci, and P. Hasler, "A 80 μW /frame 104 \times 128 CMOS imager front end for JPEG Compression," *IEEE Int. Symp. on Circuits and Systems (ISCAS'05)*, pp 5318-5321, 2005.
- [8] E. Artyomov, Y. Rivenson, G. Levi, and O. Yadid-Pecht, "Morton (Z) scan based real-time variable resolution CMOS image sensor," *IEEE Trans. Circuits and Systems for Video Technology*, vol. **15** (7), pp 947-952, 2005.
- [9] E. Artyomov, and O. Yadid-Pecht, "Adaptive multiple resolution CMOS active pixel sensor," *IEEE Int. Symp. on Circuits and Systems (ISCAS'04)*, vol. **4**, pp 836-839, 2004.
- [10] A. Graupner, J. Schreiter, S. Getzlaff, and R. Schuffny, "CMOS image sensor with mixed-signal processor array," *IEEE J. Solid-State Circuits*, vol. **38** (6), pp 948-957, 2003.
- [11] A. Olyaei, and R. Genov, "CMOS Wavelet Compression Imager Architecture," *IEEE CAS Emerging Technologies Workshop*, St. Petersburg, Russia, June 23-24, 2005.
- [12] A. Olyaei, and R. Genov, "Mixed-Signal CMOS Haar Wavelet Compression Imager Architecture," *Midwest Symposium on Circuits and Systems (MWSCAS05)*, Cincinnati, Ohio, Aug. 7-10, 2005.
- [13] A. Olyaei, and R. Genov, "Algorithmic Delta-Sigma Modulated FIR Filter," *IEEE Int. Symp. on Circuits and Systems (ISCAS'06)*.
- [14] G. Mulliken, F. Adil, G. Cauwenberghs, and R. Genov, "Delta-Sigma Algorithmic Analog-to-Digital Conversion," *IEEE Int. Symp. on Circuits and Systems (ISCAS'02)*, Scottsdale, AZ, May 26-29, 2002. ISCAS'2002.