CMOS Image Compression Sensor with Algorithmically-Multiplying ADCs

Alireza Nilchi, Joseph Aziz, and Roman Genov
Department of Electrical and Computer Engineering
University of Toronto, Toronto, Ontario, M5S 3G4, Canada
Email: {alireza, joseph, roman}@eecg.utoronto.ca

Abstract— A 128×128 CMOS image compression sensor fabricated in a 0.35μm CMOS process is reported. It computes block-matrix and convolutional image transforms with digital kernels of up to 8×8 pixels directly on the focal plane. A pixel output is sampled only when the corresponding bit of the kernel coefficient is one. Bit-wise accumulation of adjacent pixel outputs in a column is performed by the switched-capacitor accumulator circuit. A column-parallel algorithmic multiplying ADC performs binary-weighted summation by adding the accumulator circuit outputs with cyclic residues of the same binary weight. The signal range is maintained by generating two bits per cycle. The imager performs three computations per pixel readout. Image compression experimental results at 30fps and 8-bit output resolution are presented.

I. INTRODUCTION

Image compression algorithms are computationally expensive as they employ spatial image transforms. Spatial image transforms such as block-matrix and convolutional transforms compute localized spatial weighted averages of pixel outputs. Examples of such transforms commonly employed in image compression are discrete wavelet transform (DWT) and discrete cosine transform (DCT). Conventional dedicated digital signal processors require significant power and silicon area to compute spatial image transforms in streaming video. The intrinsic parallelism of analog VLSI architectures yields high computational throughput, often beyond that of digital processors, allowing to perform spatial image processing in real-time with small power and area overhead. Direct on-focal-plane image compression yields the output data rate proportional to the mere information rate of the video frame, independent of the pixel array size.

Analog VLSI implementations of on-focal-plane image compression include capacitor banks, current-mode weighted averaging, and floating gate circuits [1]-[3]. Capacitor bank implementations use charge sharing to compute weighted sum and difference [1] but may have limited scalability. Current-mode weighted averaging implementations [2] use zero-latency current-mode addition. Current-mode vector-matrix multiplication [3] architectures employ floating-gate arrays for block matrix storage to achieve high power efficiency. Purely analog implementations often suffer from limited accuracy and require an analog-to-digital converter.

On-focal-plane mixed-signal VLSI image compression combines the benefits of both analog and digital domains [4], [5]. Analog circuits perform area-efficient and low-power

$$\sum_{k=1}^{N} 2^{-(k+1)} (2\,D_{OUT1}[k] + D_{OUT2}[k]) = V_{IN}\cdot C + Q$$
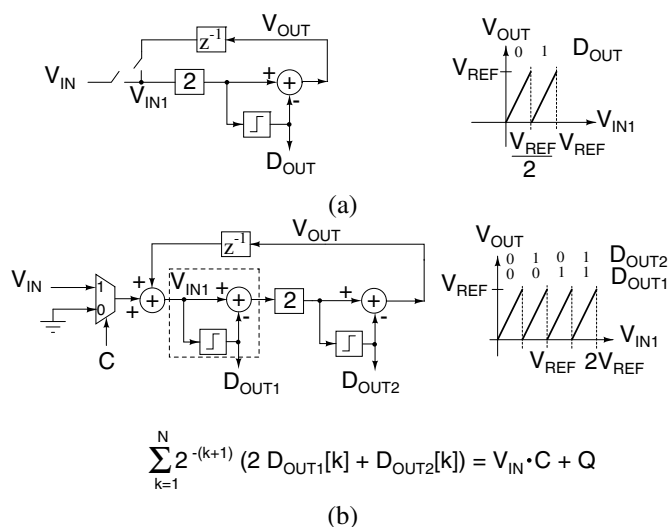
(b)

Fig. 1. (a) Algorithmic ADC architecture. (b) Algorithmic multiplying ADC architecture.

computation directly on the focal plane, eliminating the need for an external processor. Digital components maintain high accuracy, sustain configurability and provide the output in a convenient digital format.

The presented mixed-signal VLSI CMOS image sensor computes block-matrix and convolutional image transforms of each video frame with programmable digital kernels of up to 8×8 pixels, directly on the focal plane. It employs algorithmic multiplying analog-to-digital converters (MADCs) that perform spatial computation during pixel array readout, concurrently with quantization. The architecture is suitable for real-time large scale focal-plane video processing.

II. ARCHITECTURE AND VLSI CIRCUIT IMPLEMENTATION

The imager employs column-parallel algorithmic MADCs to multiply pixel outputs with their respective digital coefficients, as required in block-matrix and convolutional transforms. The architecture of the conventional algorithmic analog-to-digital converter (ADC) shown in Fig. 1 (a) is extended such that the cyclic residue is combined with a new input as shown in Fig. 1 (b) [6]. The input signal, V_{IN}, is multiplied by the bits of the kernel coefficient, C, sequentially from MSB to LSB. The outputs of the multiplexer are analog and binary-weighted, with their weight decreasing

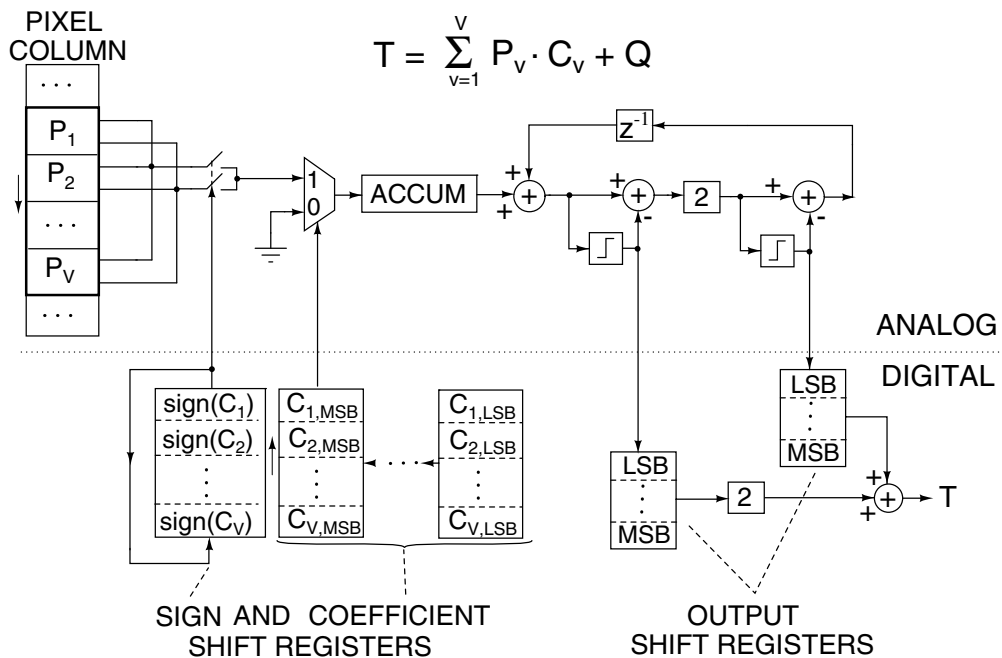$$T = \sum_{v=1}^{V} P_v \cdot C_v + Q$$

Fig. 2. Architecture of one column of the image compression sensor.

by a factor of two in each subsequent cycle, in order to match the weight of the corresponding residue signal. As the range of the accumulated signal, V_{IN1}, is twice the residue range of the conventional algorithmic ADC, an additional residue modulator is introduced as highlighted by a dashed outline in Fig. 1 (b), and two bits, D_{OUT1} and D_{OUT2}, are extracted during each kth conversion cycle, where $k = 1, 2, ..., N$ and N is the resolution of the ADC. These bits are shifted out MSB first. A weighted sum of these outputs, as shown in the equation in Fig. 1 (b), yields a digital representation of the product $V_{IN}.C$ with the corresponding quantization noise Q.

Each column-parallel MADC is further extended to include cross-pixel accumulation in order to compute not just an analog-digital multiplication but an analog-digital inner product, T, as shown in Fig. 2. The outputs, P_v, $v = 1, ..., V$, of a set of up to $V= 8$ adjacent pixels in one column are multiplied by the respective V coefficients, C_v, of one column of a programmable spatial kernel stored in a shift register. First, the pixel outputs are multiplied by the MSB bits of the respective kernel coefficients, with the results added across all V pixels by the accumulator. This procedure is repeated for all the remaining bits from MSB-1 to LSB. Signed multiplication is achieved by selecting the switched-capacitor sampling order of the pixel signal and reset levels controlled by a respective coefficient sign bit. The discrete outputs of the accumulator have a decrementing binary weight and are added to the MADC residue with a matching binary weight. The two output bit streams are combined in the digital domain as shown in Fig. 2.

Fig. 3 depicts the circuit implementation of the column signal path from a pixel to the accumulator. The pixel has two frame memories to allow for multiple sampling of the same

pixel as needed in the MADC architecture. It is also utilized for frame differencing [7]. The gain stage following the pixel combines four functionalities in a single amplifier. Double sampling of the pixel outputs suppresses fixed-pattern noise (FPN) and implements sign transformation. Binary-analog multiplication is performed by the two-switch multiplexer controlled by C and $\bar{C}$. Accumulation over V adjacent pixels is performed on the programmable feedback capacitor bank. The output dynamic range is maintained by programming the gain to match the size of the block matrix. The inverting amplifier in Fig. 3 is a single-ended cascode common-source amplifier.

Fig. 4 explains how the architecture of the MADC is optimized to reduce the number of gain stages for minimum power dissipation. As shown with the dashed outlines in Fig. 4 (a), the adders and multipliers are combined in two groups. The active stages within each group are combined and implemented with a single-amplifier three-input switched-capacitor adder, which is reused twice in the architecture as shown in Fig. 4 (b). Adding three inputs is implemented in four clock phases. Two of the eight phases needed for the two adder stages are shared which results in a total of six clock phases, as illustrated in Fig. 4 (b).

III. COMPARATIVE ANALYSIS

This section compares the presented image compression sensor architecture with a conventional digital imager where column-parallel algorithmic ADCs are utilized for raw-image quantization and no computation is performed. Comparison is made for the case when the accumulator and MADC circuits perform focal-plane 8×8 block-matrix transform computation.

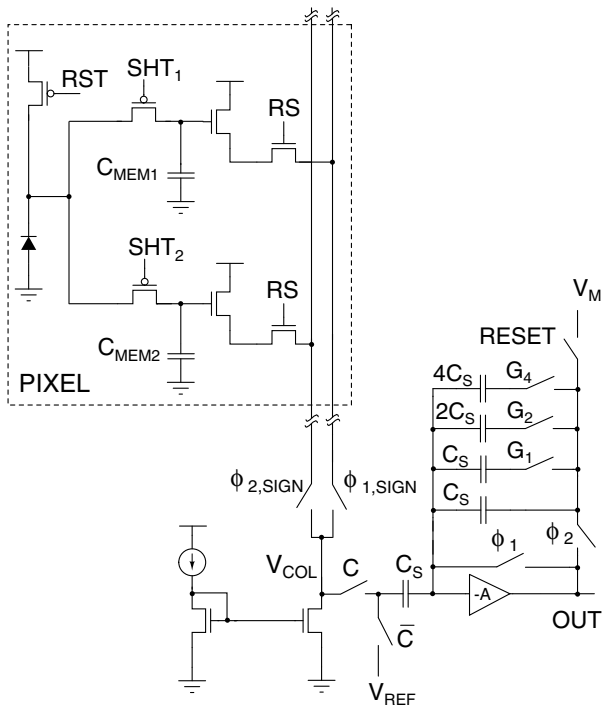In the conventional approach, ADCs directly quantize the pixel outputs. The column-parallel ADC sampling rate is

Fig. 3. The pixel and Difference/sign/multiplier/accumulator circuit.

Fig. 4. (a) Amplifier merging in the multiplying algorithmic ADC architecture. (b) The resulting two-stage multiplying ADC architecture.

proportional to the imager vertical resolution. For the MADC architecture performing 8×8 block-matrix transform computation, the sampling rate of the MADCs is effectively reduced by a factor of 8, as there is one ADC sample per 8 pixels in any column. Assuming amplifier static power consumption is dominant in both ADCs power, and that the amplifiers in the two architectures drive equal load capacitances, this translates to a factor of 8 savings in the MADC power. Since the MADC amplifiers operate off a six-phase clock, there is a factor of 3 overhead in power compared to the conventional algorithmic ADC which utilizes a two-phase clock (assuming accuracy is not limited by capacitor matching). Therefore, if the conventional algorithmic ADC power is denoted as P, the total MADC static power dissipation is equal to $3P/8$, divided equally between the two three-input adders.

The power consumption of the accumulator amplifier can also be estimated under the assumption of having equal capacitive loads for the accumulator and the MADC amplifiers. In the case of 8×8 kernels, 8 pixels have to be sampled during each conversion cycle. This operation requires 8 clock cycles or 16 clock phases. The MADC has 6 clock phases. Therefore, the accumulator amplifier has to be faster than each MADC amplifier by a factor of 16/6. As there are two MADC amplifiers, this results in a power consumption of $P/2$ for the accumulator. Therefore, the total power consumption of one MADC channel in the image compression sensor is $7P/8$. Considering the small overhead power required to bias the in-pixel source followers during multiple pixel sampling, we can conservatively assume equal power dissipation for the two cases being compared.
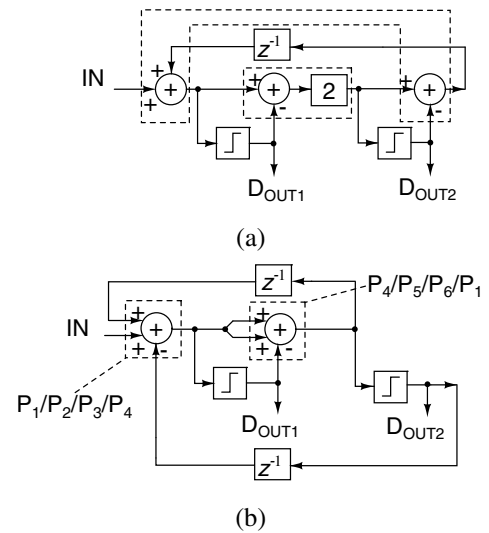
IV. EXPERIMENTAL RESULTS

This section presents experimental results measured from a 128×128-pixel integrated prototype fabricated in a 0.35μm standard CMOS process. The micrograph of the image sensor die is shown in Fig. 5.

Fig. 6 (a) depicts the original test image (portrait of Audrey) utilized to characterize the imager. Fig. 6 (b) shows the image acquired and digitized by the fabricated prototype when it was configured in the raw-image readout mode.

Two-dimensional Haar DWT was computed on the chip. To achieve image compression, the transformed image pixels are compared to a threshold value. Transformed image details which have a magnitude below the threshold are filtered out. Fig. 7 demonstrates the experimentally measured trade-off between the peak signal-to-noise ratio (PSNR) and the compression ratio for the first-level Haar transform obtained by varying the compression threshold. The inset images correspond to the experimentally recorded images that were compressed on the chip and subsequently decompressed off-chip.

Accounting for frame differencing, multiplication and accumulation, the imager performs three operations per pixel readout corresponding to 90MOPS/MPixel at 30fps frame rate. As described in section III, in 8×8-pixel block-matrix transforms with single readout this computational throughput comes at the cost of approximately the same power as dissipated by a digital imager with conventional column-parallel algorithmic ADCs. The experimentally measured power dissipation of all 128 accumulator and MADC circuits is 24.4mW. The maximum computational throughput is 160MOPS for 128×128 pixel array size and scales to 20GOPS at HDTV 1080i imager resolution. A summary of characteristics of the CMOS compression imager is given in Table I.
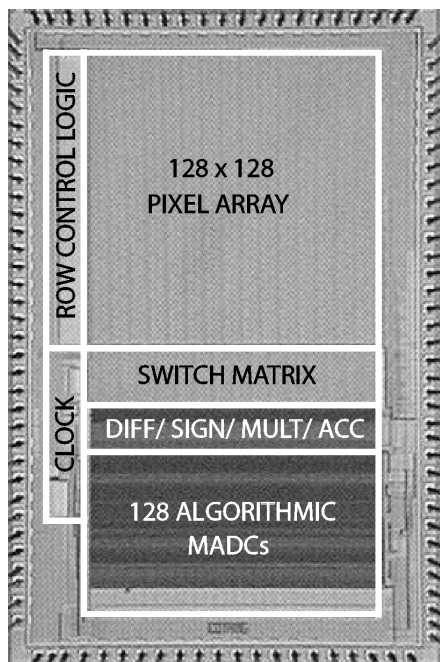
Fig. 5. Die micrograph of the focal-plane algorithmically-multiplying CMOS image compression sensor. The integrated 4.4mm×2.9mm prototype was fabricated in a 0.35μm standard CMOS technology.
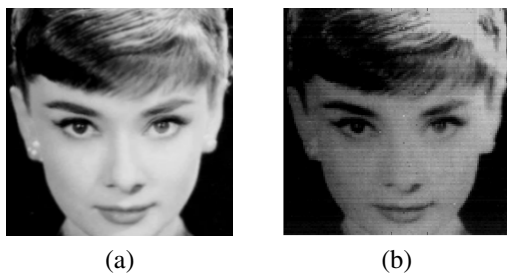

(a) (b)

Fig. 6. (a) Test image (portrait of Audrey) projected onto the pixel array. (b) Experimentally measured digital output of the CMOS imager obtained with on-chip algorithmic ADCs.
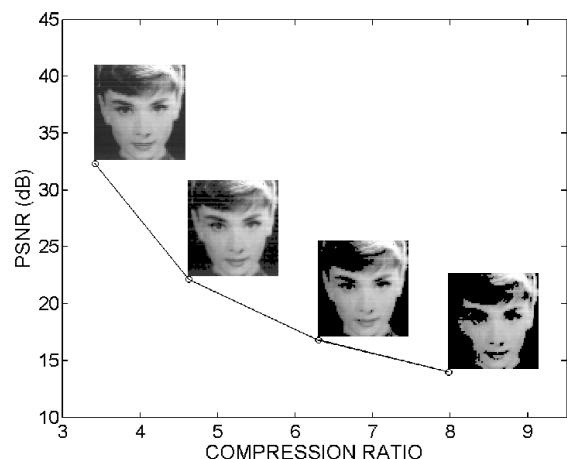

Fig. 7. Reconstructed images obtained by decompression of the experimentally computed one-level Haar discrete wavelet transform of the original image for varying compression thresholds.

TABLE I
SUMMARY OF CHARACTERISTICS

Technology	0.35μm CMOS
Supply Voltage	3.3V
Die Area	4.4mm×2.9mm
Array Dimensions	128×128 pixels
Pixel Size	15.4μm×15.4μm
Fill Factor	28%
Dark Current	36fA/pixel
Frame Rate	30fps
Output Resolution	8-bit
Kernel Size	2×2 to 8×8
Single-scan Throughput	90MOPS/Mpixel
Max Throughput	160MOPS @ 128×128
	20GOPS @ HDTV 1080i
Pixel Array	0.6mW
Accumulator and MADC	24.4mW
Digital	1.2mW
Total Power (8×8 Kernel)	26.2mW

V. CONCLUSIONS

A mixed-signal VLSI implementation of a digital CMOS imager performing block-matrix and convolutional transforms on the focal plane for real-time image compression is presented. The approach combines spatial image transform computation and quantization in a single algorithmic analog-to-digital conversion cycle. In an 8×8 block-matrix transform computation, the approach yields power dissipation almost equal to that of a conventional digital imager, while the need for a peripheral DSP is eliminated. The imager has been experimentally validated in DWT image compression. It yields 8-bit digital output at a rate proportional to the mere information rate of the video, not the dimensions of the pixel array.

REFERENCES

[1] 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.
[2] 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.
[3] A. Bandyopadhyay, J. Lee, R.W. Robucci, and P. Hasler, "MATIA: A programmable 80 μW/frame CMOS block matrix transform imager architecture," *IEEE J. Solid-State Circuits,* vol **41** (3), pp 663-672, 2006.
[4] 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.
[5] A. Olyaei, and R. Genov, "Focal-Plane spatially oversampling CMOS image compression sensor," *IEEE T. Circuits and Systems I*, vol. **49** (1), pp.26-34, 2007.
[6] R. Genov, and G. Cauwenberghs, "Algorithmic partial analog-to-digital conversion in mixed-signal array processors," *IEEE Int. Symp. on Circuits and Systems,* 2003.
[7] A. Nilchi, J. Aziz, and R. Genov "Focal-Plane Algorithmically-Multiplying CMOS Computational Image Sensor," *to appear in IEEE J. Solid-State Circuits,* 2009.