

Focal-Plane Spatially Oversampling CMOS Image Compression Sensor

Ashkan Olyaei, *Student Member, IEEE*, and Roman Genov, *Member, IEEE*

Abstract—Image compression algorithms employ computationally expensive spatial convolutional transforms. The CMOS image sensor performs spatially compressing image quantization on the focal plane yielding digital output at a rate proportional to the mere information rate of the video. A bank of column-parallel first-order incremental $\Delta\Sigma$ -modulated analog-to-digital converters (ADCs) performs column-wise distributed focal-plane oversampling of up to eight adjacent pixels and concurrent weighted average quantization. Number of samples per pixel and switched-capacitor sampling sequence order set the amplitude and sign of the pixel coefficient, respectively. A simple digital delay and adder loop performs spatial accumulation over up to eight adjacent ADC outputs during readout. This amounts to computing a two-dimensional block matrix transform with up to 8×8 -pixel programmable kernel in parallel for all columns. Noise shaping reduces power dissipation below that of a conventional digital imager while the need for a peripheral DSP is eliminated. A 128×128 active pixel array integrated with a bank of 128 $\Delta\Sigma$ -modulated ADCs was fabricated in a $0.35\text{-}\mu\text{m}$ CMOS technology. The $3.1\text{ mm} \times 1.9\text{-mm}$ prototype captures 8-bit digital video at 30 frames/s and yields 4 GMACS projected computational throughput when scaled to HDTV 1080i resolution in discrete cosine transform (DCT) compression.

Index Terms— $\Delta\Sigma$ -modulated analog-to-digital converter (ADC), block matrix transform, CMOS imager, focal-plane image compression.

I. INTRODUCTION

HIGH-RESOLUTION high-frame-rate image sensors yield high video quality at the expense of increased output data rate, placing stringent requirements on the imager read-out communication channel. In the case of portable wireless video sensors, this translates into wider channel bandwidth, higher transmitter power dissipation, and increased memory size. Image compression relaxes these requirements at the cost of additional signal processing. Block matrix transforms such as discrete cosine transform (DCT) or discrete wavelet transform (DWT) are widely used in image and video compression algorithm standards such as JPEG, JPEG2000, H.261, and MPEG. They require computationally intensive spatial filtering of images demanding vast computing resources for real-time operation [1].

Various techniques for realizing block matrix transforms in sensory systems have been developed. Dedicated digital signal

processors (DSPs) rely on high-throughput architectures to compute spatial weighted sums needed in block transforms, but require an analog-to-digital converter (ADC) to quantize the analog sensory input prior to processing. At high imager resolutions, the input data rate or memory-processor bandwidth of such processors may limit their sustained throughput [2], [3]. To overcome this problem and eliminate the need for a DSP, block matrix transforms have also been implemented in the analog domain directly on the focal plane.

On-focal-plane analog video compression yields the analog output data rate of an imager proportional to the mere information rate of the video stream, not the imager resolution or frame rate. Capacitor bank implementations use charge sharing to compute weighted sum and difference [4]–[6] but may have limited scalability. Current-mode weighted averaging implementations [7] 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 [8] utilized in variable resolution imaging do not allow for weighted averaging and require additional column-parallel amplifiers. Current-mode vector-matrix multiplication [9] 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 [10]. A tree-based partitioning algorithm that implements adaptive compression has also been reported [11]. Circuit implementations based on video compression algorithms utilize in-pixel temporal prediction [12]–[14] and array-based spatial prediction [15] to reduce the amount of transmitted data. All of the aforementioned architectures perform computation in analog VLSI domain and require an extra ADC to provide the digital output.

Mixed-signal CMOS imaging and signal processing combine the benefits of both analog and digital domains [16]. Analog circuits perform area-efficient and low-power computation directly on the focal plane, eliminating the need for an external processor [17]. 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. Digital components provide the output in a convenient digital format, and sustain the accuracy and configurability of such systems. 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 $\Delta\Sigma$ -modulated analog-to-digital conversion cycle, making focal-plane

Manuscript received February 24, 2006; revised October 4, 2006. This work was supported by NSERC, CITO, and CFI. This paper was recommended by Guest Editor O. Yadid-Pecht.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: roman@eecg.utoronto.ca).

Digital Object Identifier 10.1109/TCSL.2006.887976

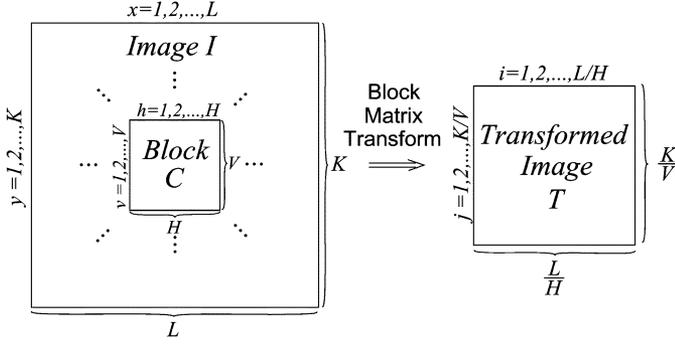


Fig. 1. Illustration of a block matrix transform computation for nonoverlapping blocks.

computing an intrinsic part of the quantization process. The approach yields power dissipation below that of a conventional digital imager while the need for a peripheral DSP is eliminated.

The rest of this paper is organized as follows. Section II gives an overview of the block matrix transform method for image compression. Section III presents the overall VLSI architecture of the focal-plane spatially oversampling CMOS image compression sensor. Sections IV and V discuss the circuit implementations of the image acquisition and image compression tasks, respectively. Section VI demonstrates the benefits of the presented approach over conventional digital imagers. Section VII presents experimental results obtained from a $0.35\text{-}\mu\text{m}$ CMOS prototype of the sensory image compression processor.

II. BLOCK MATRIX TRANSFORMS FOR 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 in the process of image compression. To transform an image I into block-transformed image T , the block matrix C is tiled vertically and horizontally across the image as illustrated in Fig. 1. The block matrix is tiled in overlapping or nonoverlapping fashion depending on the block matrix transform type. For the nonoverlapping case shown in Fig. 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 block matrix indices, and i and j are the indexes of the block-transformed image.

To achieve selective compression of the image, redundant and localized gradient values are filtered out according to a threshold bias, which is based on the required compression ratio and the reconstructed image quality specifications. Another thresholding technique, mainly employed in JPEG image

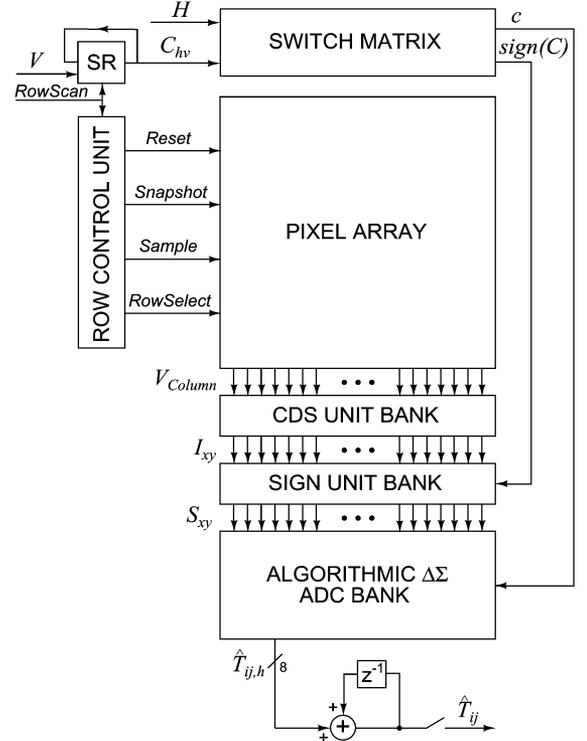


Fig. 2. Top-level architecture of the focal-plane spatially oversampling CMOS image compression sensor. Digital accumulation and thresholding (not shown) blocks are implemented off-chip.

compression, is nonuniform quantization of the block matrix transformed image, where the more significant low-frequency spatial information components are quantized with a higher resolution compared to the less important high-frequency ones.

III. ARCHITECTURE

The block matrix transform of the form (1) can be decomposed as follows:

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

with partial sums

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

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

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

where the notation is consistent with that of (2)–(3), $C_{hv} = \text{sign}(C_{hv})|C_{hv}|$, and I_{xy} is the output of a pixel at location (x,y) .

The proposed mixed-signal VLSI architecture efficiently implements computations (6), (5), and (4), in that order, as depicted in Fig. 2. Image acquisition and correlated double sampling (CDS) yield offset-compensated pixel output I_{xy} as de-

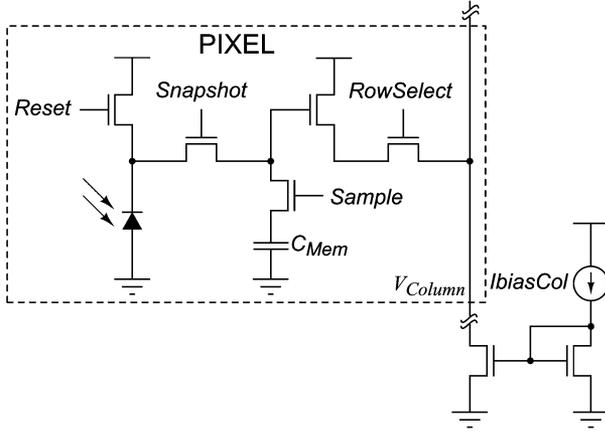


Fig. 3. Photodiode-based active pixel sensor circuit.

scribed in Section IV. A switched-capacitor sign unit circuit multiplies pixel output I_{xy} by the sign of a respective kernel coefficient via selecting the sampling sequence order yielding a sign-transformed pixel output S_{xy} in (6). Section V-A presents the sign unit circuit implementation. Weighted average of V adjacent pixel outputs in an image column is computed by combining oversampling quantization and selective distributed sampling of the sign-transformed pixel outputs to yield $\hat{T}_{ij,h}$, as discussed in Sections V-B and Sections V-C. $\hat{T}_{ij,h}$ is the digital representation of $T_{ij,h}$ in (5). The switch matrix routes the block matrix coefficients and their corresponding sign values bit-serially from a ring shift register, SR , with a sequence period of V values and spatial period of H columns, synchronously with image read-out clock $RowScan$ to the oversampling quantizers and sign units, respectively. The operation of the switch matrix is discussed in Section V-D. 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 \hat{T}_{ij} , which is the digital representation of T_{ij} in (4). Section V-E presents an implementation of this digital accumulation.

IV. IMAGE ACQUISITION

Image acquisition is performed by the active pixel array, the row control, and the CDS units. The active pixel comprises a resettable n^+ -diffusion- p -substrate photodiode, a selectable analog memory C_{Mem} , and a selectable source follower with shared column-parallel current source biased with $I_{biasCol}$ current as shown in Fig. 3. The analog memory is implemented as a MOS capacitor for higher density of integration inside the pixel and consequently a larger fill factor.

The row control unit generates the digital signals *Reset*, *Snapshot*, *Sample*, and *RowSelect* controlling the integration and readout phases. The active pixel array can be programmed for four different modes of operation. In the snapshot mode of operation, as shown in the timing diagram of Fig. 4, a snapshot of the focal-plane image is stored in the array's analog memories. The stored values are then read out on a row-by-row basis. In the rolling mode of operation, both integration and readout are performed row by row, as illustrated in Fig. 5. In the CDS mode of operation, as shown in the timing diagram of Fig. 6,

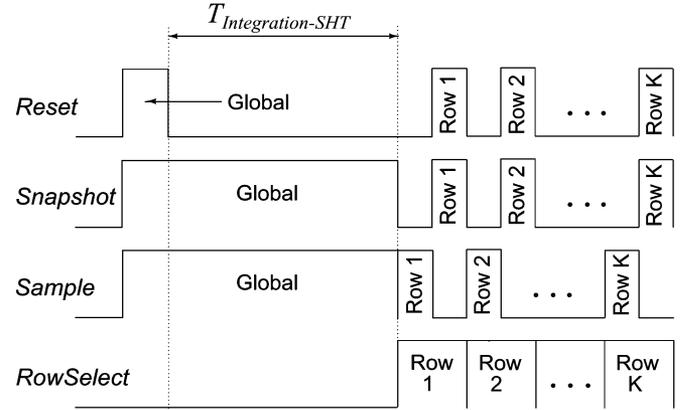


Fig. 4. Timing diagram for the snapshot mode of operation.

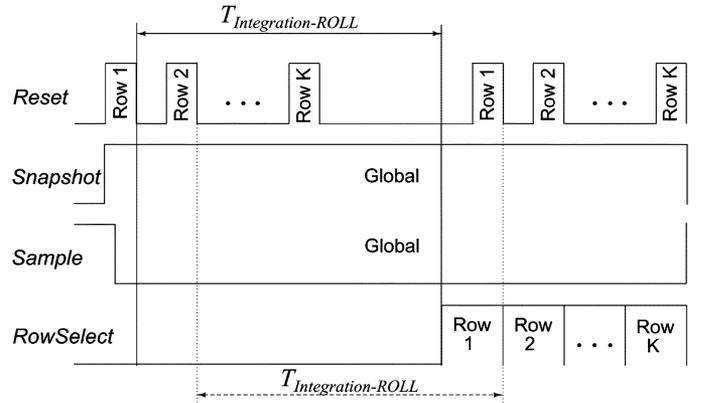


Fig. 5. Timing diagram for the rolling mode of operation.

the reset signals of all pixels are stored in the analog memories. The readout is then performed row by row. In the frame-differencing mode of operation, snapshot and rolling modes are combined to yield forward frame-differencing [18]–[20], which is widely employed in motion detection algorithms. The timing diagram for the frame differencing mode is shown in Fig. 7.

Fixed pattern noise (FPN) due to pixel-to-pixel transistor mismatch is one of the dominant sources of noise in CMOS imagers. A switched-capacitor CDS unit shown in Fig. 8 suppresses FPN by subtracting the reset pixel output from the sensed pixel output. The amplifier is a single-stage common-source cascoded nMOS amplifier. The clocks ϕ_{1CDS} and ϕ_{2CDS} are nonoverlapping. When ϕ_{1CDS} is high, the sensed pixel output is sampled on the input capacitor. When ϕ_{2CDS} is high, the reset pixel output is subtracted from the earlier sampled signal, with the difference I_{xy} produced at the CDS circuit output. In the frame-differencing mode of operation, the sampled pixel signal in a frame is subtracted from the pixel output generated in the consecutive frame to yield forward frame-differencing [14]. In the snapshot, rolling, and forward frame-differencing modes of operation, the CDS unit performs a double sampling operation which removes FPN, but does not affect the thermal (reset) noise. In the CDS mode of operation, the pixel reset and signal values are correlated; hence, both FPN and thermal noise are suppressed.

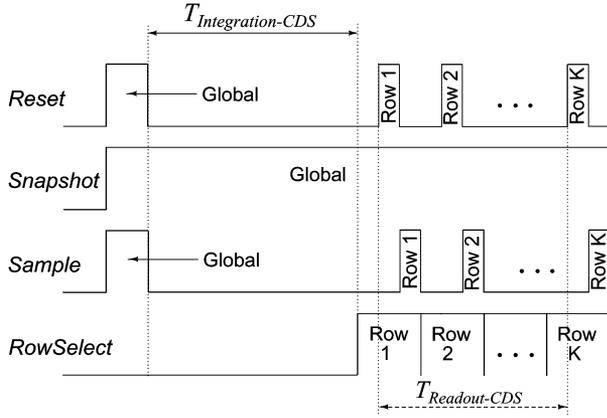


Fig. 6. Timing diagram for the CDS mode of operation. To achieve uniform integration time for all rows, $T_{\text{Integration-CDS}} \gg T_{\text{Readout-CDS}}$, which is often the case.

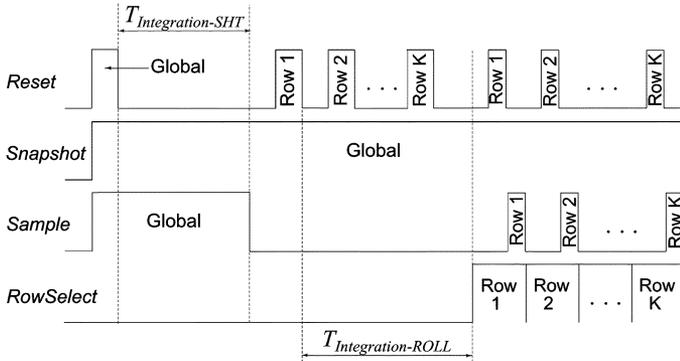


Fig. 7. Timing diagram for the frame-differencing mode of operation.

V. COMPUTATIONAL QUANTIZATION

This section presents a mixed-signal VLSI implementation of (6), (5), and (4). To simplify notation, in this section we consider (5) and (6) for a single image column segment (*i.e.*, for given i , h , and x) and the first row of the transformed image (*i.e.*, for $j = 1$ and $y = v$). This simplifies the partial sum in (5) to

$$T = \sum_{v=1}^V C_v I_v = \sum_{v=1}^V |C_v| S_v. \quad (7)$$

For a particular image row (*i.e.*, for given v), the sign-transformed pixel output in (6) further simplifies to

$$S = \text{sign}(C)I \quad (8)$$

where I is the raw output of a pixel.

A. Sign Unit

The sign unit shown in Fig. 9(a) is implemented as a switched-capacitor difference circuit. It applies the sign of the coefficient C to the input by selecting a switched-capacitor sampling sequence order as illustrated in the timing diagram in Fig. 9(b). This directly implements (8). The amplifier in the sign unit is the same as in the CDS circuit. For the sake of

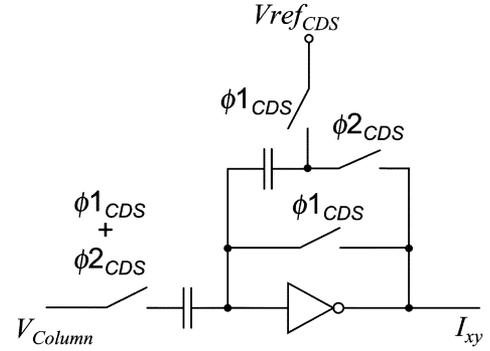


Fig. 8. CDS circuit.

simplicity the feedback capacitor reference voltage is shown as ground.

B. $\Delta\Sigma$ -Modulated Multiplying ADC

The spatially compressing image quantizer is implemented as a first-order incremental $\Delta\Sigma$ -modulated ADC extended to an oversampling multiplying ADC [21] as described in this section. The first-order incremental oversampling ADC is depicted in Fig. 10(a). It converts a sequence of analog samples into a digital word representing a quantized version of the average of all samples. It is comprised of a sample-and-hold (S/H) circuit, an integrator, a comparator and a decimating counter. The rectangular decimation window and initial reset of the accumulator avoid tones in the quantization noise spectrum at dc input that are characteristic of a conventional first-order DS modulator with a low-pass decimation filter. As shown in Fig. 10(b), this architecture can be combined with the sign unit and extended to perform both quantization and signed multiplication of the analog input with a digital word

$$C = \text{sign}(C)|C| \quad (9)$$

with

$$|C| = \sum_{i=0}^{N-1} c[i] \quad (10)$$

where $c[i]$ are unsigned unary coefficients of C .

Selective sampling of the sign-transformed pixel output S , controlled by the bit-serial unary sequence $c[i]$, yields an analog sequence $u[i] = Sc[i]$. The first-order modulator converts the sequence $Sc[i]$ into a bit stream $y[i]$ in N cycles, using a “resettable” (RST) analog integrator

$$w[0] = 0 \quad (11)$$

$$w[i+1] = w[i] + \alpha(Sc[i] - y[i]), \quad i = 0, \dots, N-1 \quad (12)$$

$$w[N+1] = w[N] - \alpha y[N] \quad (13)$$

and a single-bit quantizer

$$y[0] = -1 \quad (14)$$

$$y[i] = \text{sign}(w[i]), \quad i = 1, \dots, N \quad (15)$$

where α is the intrinsic gain of the integrator.

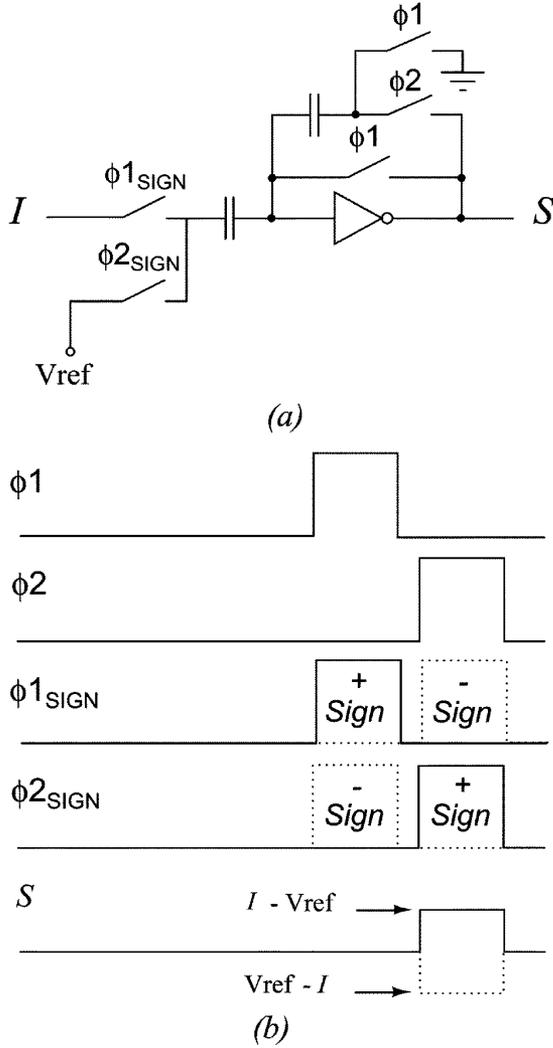


Fig. 9. (a) Sign unit circuit in the sensory processor. (b) Sign circuit timing diagram; for a continuous range of signed outputs $V_{ref} = \min\{I\}$.

A binary counter accumulates the bits $y[i]$ to produce a decimated output. The rectangular decimation window, and initial reset of the integrator, avoid tones in the quantization noise spectrum at dc input that are characteristic of a conventional first-order $\Delta\Sigma$ modulator with low-pass decimation filter [22]. The quantization error (conversion residue) is directly given by the final integrator value $(1/\alpha)w[N+1]$, as verified by summing (12) and (13) over i

$$\sum_{i=0}^N y[i] = \sum_{i=0}^{N-1} Sc[i] - \frac{1}{\alpha}w[N+1] \quad (16)$$

where

$$\sum_{i=0}^N y[i] = \hat{T}' \quad (17)$$

is the digital output.

This operation yields multiplication of the sign-transformed analog pixel output S with the unsigned digital coefficient $|C|$

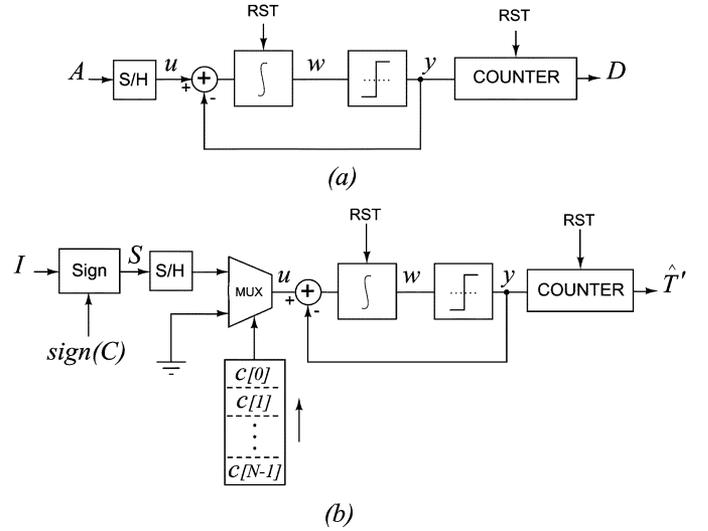


Fig. 10. (a) First-order $\Delta\Sigma$ incremental A/D converter. (b) $\Delta\Sigma$ -modulated multiplying ADC. The sign unit circuit shown in Fig. 9(a) also performs the S/H operation. Here, the S/H cell is explicitly shown for clarity.

defined in (10), while a digital output resolution of $\log_2(N)$ bits is warranted

$$\hat{T}' = |C|S + q' \quad (18)$$

which in combination with (8) and (9) yields

$$\hat{T}' = CI + q' \quad (19)$$

where

$$|q'| = \left| \frac{1}{\alpha}w[N+1] \right| \quad (20)$$

is the multiplication quantization noise. Higher resolution at lower oversampling ratio N can be obtained using higher-order incremental conversion [23].

As the input is amplitude modulated with unary signed coefficients, an error in the amplitude of these coefficients can contribute to the noise. A noise analysis due to this nonideality is given in [24] where amplitude modulation of an analog sequence with a Hadamard sequence is utilized in the design of a Nyquist-rate $\Delta\Sigma$ -modulated ADC. This noise is deemed negligible in this design as a simple analog multiplexer is employed to modulate the input.

C. $\Delta\Sigma$ -Modulated Weighted Averaging ADC

The multiplying ADC architecture in Section V-B can be extended to perform weighted averaging. Accumulation is achieved by sampling V adjacent pixels in one column, $I_1, I_2 \dots I_V$, without resetting the integrator or the binary counter. A discrete-time index v is thus introduced.

The architecture of the oversampling weighted averaging ADC is depicted in Fig. 11. The bit stream $y[v_i]$ is now generated for V inputs each sampled N times

$$w[0] = 0 \quad (21)$$

$$w[v(i+1)] = w[v_i] + \alpha(S[v]c[i, v] - y[v_i]), \quad i = 0, \dots, N-1; \quad v = 1, \dots, V \quad (22)$$

$$w[V(N+1)] = w[VN] - \alpha y[VN] \quad (23)$$

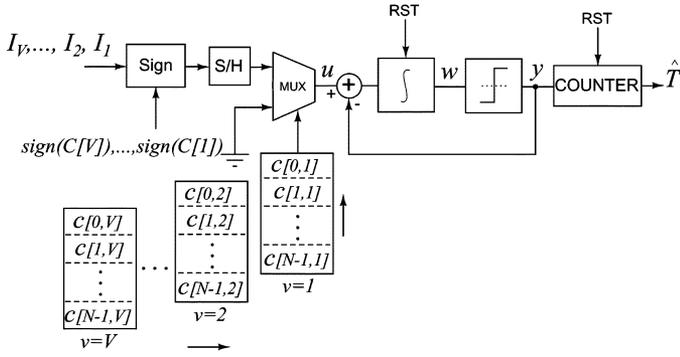


Fig. 11. $\Delta\Sigma$ -modulated weighted averaging ADC for $j = 1$ and given i and h . The ADC samples V adjacent pixels in one column, weights each by coefficient C_v , and concurrently quantizes their sum.

and a single-bit quantizer

$$y[0] = -1 \quad (24)$$

$$y[v_i] = \text{sign}(w[v_i]), \quad i = 1, \dots, N; \quad v = 1, \dots, V. \quad (25)$$

The quantization error $(1/\alpha)w[V(N+1)]$ is obtained similarly

$$\sum_{v=1}^V \sum_{i=0}^N y[v_i] = \sum_{v=1}^V \sum_{i=0}^{N-1} S[v]c[i, v] - \frac{1}{\alpha}w[V(N+1)] \quad (26)$$

where

$$\sum_{v=1}^V \sum_{i=0}^N y[v_i] = \hat{T} \quad (27)$$

and the notation is consistent with that of (11)–(16).

This realizes the computation of a weighted sum of sign-transformed pixel outputs $S[v]$ with the unsigned digital coefficients $|C[v]|$, defined in (10), with an output resolution of $\log_2(VN)$ bits

$$\hat{T} = \sum_{v=1}^V |C[v]|S[v] + q \quad (28)$$

which in combination with (8) and (9) yields

$$\hat{T} = \sum_{v=1}^V C[v]I[v] + q \quad (29)$$

where

$$|q| = \left| \frac{1}{\alpha}w[V(N+1)] \right| \quad (30)$$

is the weighted averaging quantization error. We arrive at expression (28) for the digital weighted sum, \hat{T} , which is a discrete-time equivalent of (7).

Optimization of the number of oversampling cycles based on particular block matrix coefficients can enhance the computational throughput of the architecture. When the maximum absolute value of coefficients in a single row of a block matrix (scaled to all integers) N_V is less than N , the oversampling computational cycle is stopped on the N_V -th sample and continued on the next row.

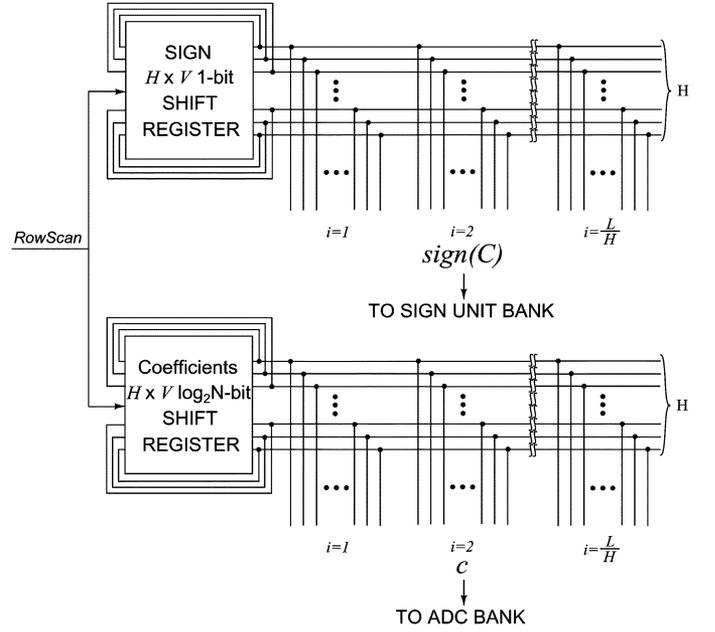


Fig. 12. Block diagram of the switch matrix.

Further improvements of the computational throughput of this architecture are achieved by employing an algorithmic $\Delta\Sigma$ -modulated ADC [25]. When the maximum of sums of absolute values of coefficients in all columns of a block matrix (scaled to all integers) is less than VN , the oversampling computational cycle is stopped once all of the coefficients have been fed in. Higher resolution bits are then obtained by subsequent algorithmic residue resampling and extended counting on the residue [25].

D. Switch Matrix

The switch matrix routes the H different time-dependent block matrix coefficients and sign signals to L/H groups of adjacent column-parallel ADCs and sign unit circuits, respectively. The block diagram of the switch matrix is shown in Fig. 12. A total of V pixel rows are sampled while V coefficients are being shifted out from the shift register. The coefficients are looped back to the shift register input to maintain V -row time period. Each kernel coefficient is stored in a binary format of length $\log_2(N)$ -bits and is digitally oversampled to yield its unary representation of length N bits, to match the sampling mechanism of an oversampling ADC and correspondingly weight each pixel output.

E. Digital Accumulation

Re-introducing the spatial indexes i , j , and h back into (29) yields the general expression for the column-wise weighted average

$$\hat{T}_{i,j,h} = \sum_{v=1}^V C_{hv}I_{xy} + q_{i,j,h}$$

where $q_{ij,h}$ is the column-wise weighted averaging quantization noise with standard deviation $\sigma_{ij,h}$.

A simple digital delay and adder loop performs spatial accumulation over H adjacent ADC outputs in the digital domain as they are read out

$$\hat{T}_{ij} = \sum_{h=1}^H \hat{T}_{ij,h} = \sum_{h=1}^H \sum_{v=1}^V C_{hv} I_{xy} + q_{ij} \quad (31)$$

where q_{ij} is the transformed image quantization noise with standard deviation

$$\sigma_{ij} \approx \sqrt{\sum_{h=1}^H \sigma_{ij,h}^2}$$

for *i.i.d* noise and large H , yielding an additional improvement in the signal-to-noise ratio (SNR). This mixed-signal VLSI computation realizes a block matrix transform in (1) with $H \leq 8$. The switch matrix size scales linearly with H . Maximum of H equal to 8 is chosen here to strike a balance between the switch matrix implementation complexity and the overall functionality. The area overhead of sign unit circuits, switch matrix and digital accumulator scales linearly with the imager size and becomes small for large K and L . As computing is interleaved with quantization, the extra computational time and thus power dissipation are small compared to those of raw image quantization in a conventional CMOS digital imager.

VI. COMPARATIVE EXAMPLE

This section compares the presented architecture with a conventional approach where column-parallel algorithmic ADCs performing no computation are employed and an additional peripheral serial digital multiplier and accumulator performs video compression. It is assumed that the kernel is a square matrix of size V with M -bit coefficients and the frame rate is the same in both cases. The comparison is performed for one column only as the two-dimensional computation can be partitioned such that multiplication is performed in the vertical dimension, and only V additions per kernel are performed in the horizontal dimension.

The first-order incremental $\Delta\Sigma$ -modulated ADC requires a number of clock cycles exponential with the number of bits of resolution, M . This is a disadvantage compared to the algorithmic ADC which requires a number of clock cycles proportional to M . On the other hand, the SNR of the spatially oversampling ADC is much higher than that of the algorithmic ADC for the same resolution and the same energy-per-cycle due to in-pixel and inter-pixel oversampling and subsequent noise shaping. In thermal noise limited circuits, power dissipation is linear with SNR. Thus, for the same SNR power dissipation of the oversampling ADC can be reduced below that of the algorithmic ADC.

The numeric comparison depends on the degree of vertical overlap of kernels in subsequent computations. In the worst case, corresponding to the highest number of computations, the subsequent kernels overlap by $V - 1$ pixels in the vertical di-

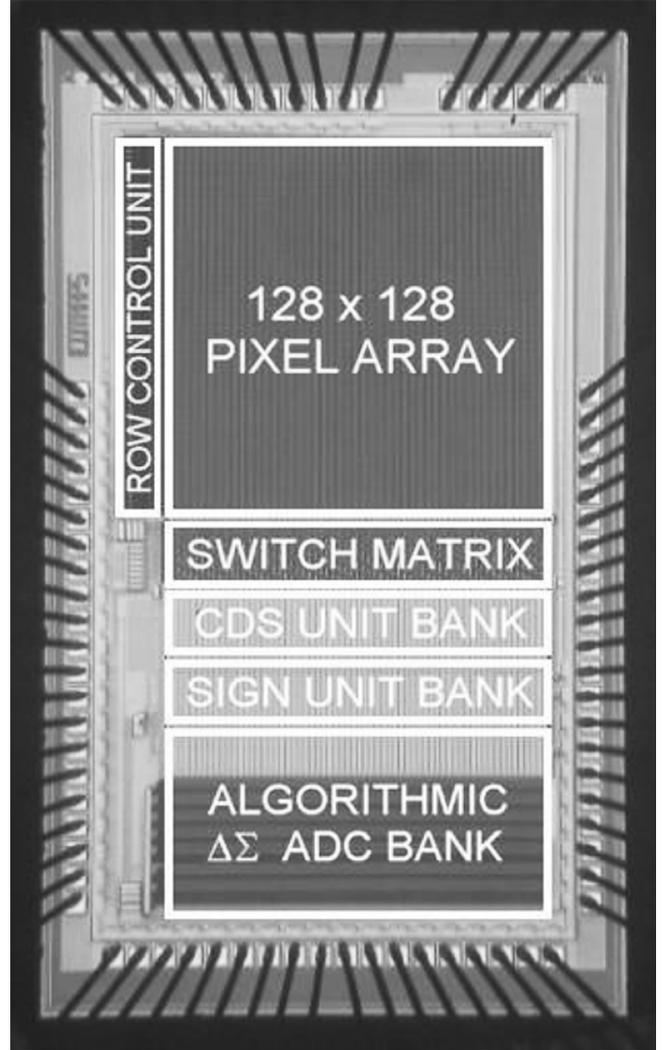


Fig. 13. Die micrograph of the focal-plane spatially oversampling CMOS image compression sensor. The integrated 3.1 mm \times 1.9 mm prototype was fabricated in a 0.35- μ m CMOS technology.

mension. Assuming $V = 8$ and $M = 8$, in the worst case, the power dissipation of the $\Delta\Sigma$ modulated spatially oversampling ADC is 63 percent of the power dissipation of the algorithmic ADC for the same SNR. In the nominal case, when the kernels do not overlap, the power dissipation of the spatially oversampling ADC is only eight percent of the power dissipation of the algorithmic ADC. This assumes multiplication and addition accuracy of 8 bits as necessary for many image compression tasks. In addition, the conventional approach requires a serial digital multiplier and an adder. At HDTV 1020i imager resolution, a computational throughput of several billions of operations per second is required and would need to be delivered by the digital multiplier and adder at the cost of significant additional power dissipation and integration area. In the proposed approach, besides savings in the ADC power dissipation, the need for such a high-throughput DSP is eliminated.

VII. RESULTS

Experimental results are obtained from a 0.35- μ m CMOS prototype containing a 128 \times 128-pixel array and a bank of

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

128 column-parallel algorithmic $\Delta\Sigma$ -modulated ADCs. Fig. 13 shows the die micrograph of the image compression sensor. Table I summarizes its electrical and optical characteristics. The value of parameters V and H are programmable in the range of 2 to 8. Any transform in this size range with signed digital coefficients can be computed. Two-dimensional Haar wavelet transform, a block matrix transform commonly used in image compression ([26],[27]), is chosen here as a simple example to illustrate the functionality of the presented imager implementation.

Fig. 14(a) shows an image acquired by the pixel array with 25-ms integration time. The algorithmic $\Delta\Sigma$ -modulated ADC performs distributed image sampling and concurrent signed weighted average quantization, realizing a one-dimensional spatial Haar wavelet transform. Two oversampling phases each of length $N = 32$ clock cycles are interleaved with a single algorithmic residue resampling cycle. Image read-out and computational quantization are characterized off-line in two sequential steps. A digital delay and adder loop implemented off-chip in digital domain performs spatial accumulation over multiple ADC outputs. This amounts to computing a two-dimensional Haar wavelet transform. Fig. 14(b) depicts experimentally measured two-dimensional one-, two-, and three-level Haar wavelet transforms of the original image. Fig. 14(c) shows the reconstructed images of the corresponding Haar wavelet transforms. The reconstructed images of one-level Haar transform are compared in Fig. 15 for various peak signal-to-noise and compression ratios.

The horizontal resolution of the imager is limited only by maximum scan-out clock frequency for a given frame rate as is the case in conventional imagers. Area and power dissipation scale linearly with the horizontal imager size. In the vertical dimension, all pixels have to be sampled within the a given frame period as set by the programmable spatial kernel with parameters H , V , and coefficients C as well as the imager resolution with parameters L , K , in (1)–(3). When computing the discrete cosine transform using 64 8×8 blocks at 30 frames/s, the sensory processor is projected to yield a computational throughput of 4 GMACS when scaled to HDTV 1080i resolution. The throughput is based on a conservative quantizer sampling rate of 40 ksp/s and a pixel integration time of 5 ms. If a higher resolution in the vertical dimension is required, either the integration time has to be reduced, or the ADC sampling rate has to increase.

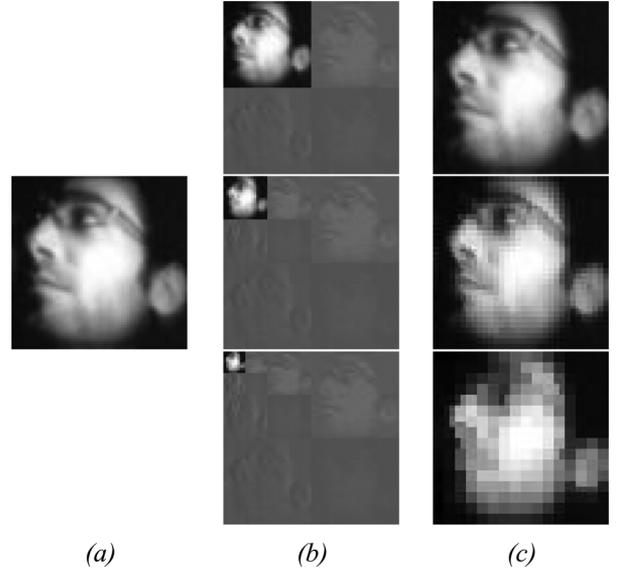


Fig. 14. (a) Image captured by the CMOS image compression sensor at 30 fps. (b) Experimentally recorded one-level (top), two-level (center), and three-level (bottom) Haar wavelet transforms of the image in (a) computed on the CMOS image compression sensor. (c) Reconstructed images for one-level (top), 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.

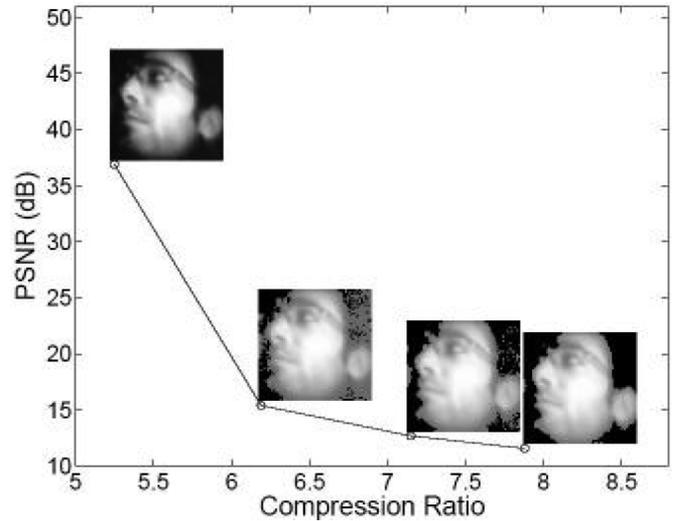


Fig. 15. Reconstructed images obtained by decompression of the experimentally computed one-level transform of the original image [top of Fig. 14(b)] for varying compression thresholds.

VIII. CONCLUSION

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 power dissipation lower than that of a conventional digital imager while the need

for a peripheral DSP is eliminated. The experimental results obtained from a 0.35- μm 128 \times 128-pixel CMOS prototype validate the utility of the design for large-scale focal-plane signal processing.

ACKNOWLEDGMENT

The authors thank the CMC foundry service for chip fabrication.

REFERENCES

- [1] H. Yamauchi, S. Okada, K. Taketa, T. Ohyama, Y. Matsuda, T. Mori, T. Watanabe, Y. Matsuo, Y. Yamada, T. Ichikawa, and Y. Matsushita, "Image processor capable of block-noise-free JPEG2000 compression with 30 frames/s for digital camera applications," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC'03)*, 2003, vol. 1, pp. 467–477.
- [2] C. Chen, Z. Yang, T. Wang, and L. Chen, "A programmable VLSI architecture for 2-D discrete wavelet transform," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'00)*, 2000, vol. 1, pp. 619–622.
- [3] U. Sjoström, I. Defilippis, M. Ansoorge, and F. Pellandini, "Discrete cosine transform chip for real-time video applications," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'90)*, 1990, vol. 2, pp. 1620–1623.
- [4] S. E. Kemeny, R. Panicacci, B. Pain, L. Matthies, and E. R. Fossum, "Multiresolution image sensor," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 4, pp. 575–583, Aug. 1997.
- [5] Q. Luo and J. G. Harris, "A novel integration of on-sensor wavelet compression for a CMOS imager," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'02)*, Scottsdale, AZ, May 26–29, 2002.
- [6] 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, no. 12, pp. 2030–2041, Dec. 1997.
- [7] V. Gruiev and R. Etienne-Cummings, "Implementation of steerable spatiotemporal image filters on the focal plane," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 49, no. 4, 2002.
- [8] Z. Zhou, B. Pain, and E. R. Fossum, "Frame-transfer CMOS active pixel sensor with pixel binning," *IEEE Trans. Electron Devices*, vol. 44, pp. 1764–1768, 1997.
- [9] A. Bandyopadhyay, J. Lee, R. Robucci, and P. Hasler, "A 80 μ W/frame 104 \times 128 CMOS imager front end for JPEG compression," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'05)*, 2005, pp. 5318–5321.
- [10] E. Artyomov, Y. Rivenson, G. Levi, and O. Yadid-Pecht, "Morton (Z) scan based real-time variable resolution CMOS image sensor," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 947–952, Jul. 2005.
- [11] E. Artyomov and O. Yadid-Pecht, "Adaptive multiple resolution CMOS active pixel sensor," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'04)*, 2004, vol. 4, pp. 836–839.
- [12] T. Hamamoto, K. Aizawa, Y. Egi, T. Hamamoto, M. Hatori, H. Maruyama, and J. Yamazaki, "On sensor image compression," in *IEEE Signal Processing Society Workshop on VLSI Signal Processing*, 1995, pp. 61–69.
- [13] K. Aizawa, H. Ohno, Y. Egi, and M. H. Yamazaki, "Image sensor for compression and enhancement," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 3, pp. 543–548, 1997.
- [14] U. Mallik, M. Clapp, E. Choi, G. Cauwenberghs, and R. Etienne-Cummings, "Temporal change threshold detection imager," in *IEEE Int. Solid-State Circuits Conf. (ISSCC'05)*, 2005, vol. 1.
- [15] D. Leon, S. Balkir, K. Sayood, and M. W. Hoffman, "A CMOS imager with pixel prediction for image compression," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'03)*, 2003, vol. 4, pp. 776–779.
- [16] A. Graupner, J. Schreiter, S. Getzlaff, and R. Schuffny, "CMOS image sensor with mixed-signal processor array," *IEEE J. Solid-State Circuits*, vol. 38, pp. 948–957, 2003.
- [17] E. R. Fossum, "CMOS image sensor: Electronic camera-on-a-chip," *IEEE Trans. Electron Devices*, vol. 44, no. 10, pp. 1689–1698, Oct. 1997.
- [18] V. Milirud, L. Fleshel, W. Zhang, G. Jullien, and O. Yadid-Pecht, "A wide dynamic range CMOS active pixel sensor with frame difference," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'05)*, 2005, vol. 1, pp. 588–591.
- [19] A. Dickinson, B. Ackland, E.-S. Eid, D. Inglis, and E. R. Fossum, "A 256 \times 256 CMOS active pixel image sensor with motion detection," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC'95)*, 1995, pp. 226–227.
- [20] A. Simoni, G. Torelli, F. Maloberti, A. Sartori, S. E. Plevridis, and A. N. Birbas, "A single-chip optical sensor with analog memory for motion detection," *IEEE J. Solid-State Circuits*, vol. 30, no. 7, pp. 800–806, Jul. 1995.
- [21] A. Olyaei and R. Genov, "Algorithmic delta-sigma modulated FIR filter," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'06)*, Kos, Greece, May 21–24, 2006.
- [22] J. Robert, G. C. Temes, V. Valencic, R. Dessoulavy, and P. Deval, "A 16-bit low-voltage CMOS A/D converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 2, pp. 157–163, Apr. 1987.
- [23] O. J. A. P. Nys and E. Dijkstra, "On configurable oversampled A/D converters," *IEEE J. Solid-State Circuits*, vol. 28, no. 7, pp. 736–742, Jul. 1993.
- [24] I. Galton and H. T. Jensen, "Delta-sigma modulator based A/D conversion without oversampling," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 42, no. 12, pp. 773–784, Dec. 1995.
- [25] G. Mulliken, F. Adil, G. Cauwenberghs, and R. Genov, "Delta-sigma algorithmic analog-to-Digital conversion," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'02)*, Scottsdale, AZ, May 26–29, 2002.
- [26] A. Olyaei and R. Genov, "CMOS wavelet compression imager architecture," in *Proc. IEEE CAS Emerging Technologies Workshop*, St. Petersburg, Russia, Jun. 23–24, 2005.
- [27] —, "Mixed-signal CMOS haar wavelet compression imager architecture," in *Proc. Midwest Symp. Circuits Syst. (MWSCAS'05)*, Cincinnati, OH, Aug. 7–10, 2005.



Ashkan Olyaei (SM'05) received the B.A.Sc. degree in electrical engineering with honors from the University of British Columbia, Vancouver, BC, Canada, and the M.A.Sc. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2003 and 2006, respectively.

He has held engineering positions at Broadcom Canada Corporation, Richmond, BC, Canada, in 2001, and GF Strong Rehabilitation Centre, Vancouver, BC, Canada, in 2002. He is currently a Design Engineer in the Wireless Department of

Marvell Semiconductors Inc. Santa Clara, CA.



Roman Genov (SM'96–M'02) received the Ph.D. degree from Johns Hopkins University, Baltimore, MD, in 2003.

He is presently an Assistant Professor in the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. His research interests include analog and digital VLSI circuits, systems and algorithms for energy-efficient signal processing with applications to electrical, chemical and photonic sensory information acquisition, biosensor arrays, neural interfaces, parallel signal processing, adaptive computing for pattern recognition, and implantable and wearable biomedical electronics.

Dr. Genov received a Best Presentation Award at IEEE IJCNN 2000, a Student Paper Award at IEEE MWSCAS 2000, Canadian Institutes of Health Research (CIHR) Next Generation Award in 2005, and the Dalsa Corporation Componentware Award in 2006. He is an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.