Mixed-Signal CMOS Wavelet Compression Imager Architecture

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 imager architecture implements $\Delta\Sigma$-modulated block matrix transforms, such as Haar wavelet transform, on the focal plane, for real-time video compression. The active pixel array is integrated with a bank of column-parallel first-order incremental oversampling analog-to-digital converters (ADCs). Each ADC performs column-wise distributed focal-plane sampling and concurrent signed weighted average quantization, realizing a one-dimensional spatial filter. A digital delay and adder loop performs spatial accumulation over multiple adjacent ADC outputs. This amounts to computing a two-dimensional block matrix transform, with no overhead in time and negligent overhead in area compared to a baseline digital imager system. The architecture is experimentally validated on a 0.35 micron CMOS prototype with a bank of first-order incremental oversampling ADCs computing Haar wavelet transform of an emulated pixel array output. The architecture yields simulated computational throughput of 1.4 GMACS with SVGA imager resolution at 30 frames per second.**

I. Introduction

Block matrix transforms such as discrete cosine transform (DCT) or discrete wavelet transform (DWT) require computationally intensive spatial filtering of the image demanding vast computing resources for real-time video compression [1]. Mixed-signal CMOS imaging technologies allow to perform low-cost, low-power signal processing directly on the focal plane, eliminating the need for an external processor [2]. The analog circuitry and intrinsic parallelism of such computing architecture yields energy efficiency and throughput well above those of modern digital processors, allowing to perform complex computations with low power consumption in real time. Digital components sustain the accuracy and configurability of such systems.

Various active pixel spatial filtering techniques suitable for on-focal-plane analog VLSI implementation of block matrix transforms have been developed. Switched capacitor implementations use charge sharing to compute average sum and difference but require matched capacitors and have limited scalability [3], [4], [5]. Current-mode implementations [6], including those with weighted averaging [7], use zero-latency current-mode addition but employ multiple matched current mirrors at the expense of increased pixel area. Current integration and gain-stage voltage summation [8], [9] utilized in variable resolution imaging do not allow for weighted averaging and require additional column-parallel amplifiers. All of the aforementioned architectures perform computing in analog VLSI domain and require an extra analog-to-digital converter (ADC) to provide the digital output.

We present a mixed-signal VLSI architecture of a digital CMOS imager computing block matrix transforms on the focal plane for real-time video compression. Our approach combines weighted spatial averaging and oversampling quantization in a single $\Delta\Sigma$-modulated analog-to-digital conversion cycle, making focal-plane computing an intrinsic part of the quantization process. The approach yields no overhead in time and negligent overhead in area compared to a baseline digital imager employing oversampling quantization. The architecture yields 1.4 GMACS throughput with SVGA imager resolution at 8-bit output resolution. The rest of this paper is organized as follows. Section II gives an overview of Haar wavelet transform, a simple example of a block matrix transform for video compression. Section III presents the architecture and circuits of the sensory Haar wavelet compression processor. Section IV presents experimental and simulated results validating functionality and computational throughput of the sensory processor architecture.

II. Haar Wavelet Transform Compression

Two-dimensional Haar wavelet transform, a block matrix transform commonly used in image compression, is chosen here as a simple example to illustrate the functionality of the presented imager architecture. Results obtained for other block matrix transforms, such as DCT, will be reported elsewhere.

By extracting horizontal, vertical and diagonal edges, Haar wavelets [10] register the relationship between intensities among neighboring pixels in different orientations and hence form a "ratio template" [11]. The ratio template yields information about spatial gradient of image intensity, not just its absolute value. As a result, a more detailed object description is generated at a cost of lower spatial resolution. 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.

The one-dimensional Haar wavelet is composed of the scaling function ϕ(t), or the father wavelet, and the wavelet prototype function ψ(t) also known as the the mother wavelet:

0-7803-9197-7/05/$20.00 © 2005 IEEE.

$$\phi(t) = \begin{cases} 1 & \text{if -1}\leq \text{t} \leq 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\psi(t) = \begin{cases} 1 & \text{if } 0 < \text{t} \leq 1 \\ -1 & \text{if -1} \leq \text{t} \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The combination of scaling and wavelet prototype functions in a two-dimensional space yields the following scalar, horizontal, vertical, and diagonal two-dimensional Haar wavelet functions:

$$\begin{aligned} \phi(x,y) &= \frac{1}{4}\phi(x)\phi(y), \\ \psi^H(x,y) &= \frac{1}{4}\psi(x)\phi(y), \\ \psi^V(x,y) &= \frac{1}{4}\phi(x)\psi(y), \\ \psi^D(x,y) &= \frac{1}{4}\psi(x)\psi(y), \end{aligned}$$

where the $\frac{1}{4}$ coefficient is applied so that the maximum value of the wavelet transform stays within the image intensity range. The equivalent spatial kernels of the scalar and wavelet functions for the first-level Haar wavelet are:

$$\mathbf{\Phi_1} = \frac{1}{4}\begin{pmatrix} +1 & +1 \\ +1 & +1 \end{pmatrix} \quad (1)$$

$$\mathbf{\Psi_1^H} = \frac{1}{4}\begin{pmatrix} +1 & -1 \\ +1 & -1 \end{pmatrix} \quad (2)$$

$$\mathbf{\Psi_1^V} = \frac{1}{4}\begin{pmatrix} +1 & +1 \\ -1 & -1 \end{pmatrix} \quad (3)$$

$$\mathbf{\Psi_1^D} = \frac{1}{4}\begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix} \quad (4)$$

The transformation of the image fragment $\mathbf{I}$ into the feature $\mathbf{X}$ is of the linear form:

$$X_{ij} = \sum_{h=1}^{K}\sum_{v=1}^{K} C_{hv}\ I_{xy}, \quad (5)$$

$$x = h + (i-1)K, \quad i = 1, 2, \ldots, \frac{H}{K}, \quad (6)$$

$$y = v + (j-1)K, \quad j = 1, 2, \ldots, \frac{V}{K}, \quad (7)$$

$$K = 2^l, \quad l = 1, \ldots, L, \quad (8)$$

where $C_{hv} \in \{-1, +1\}$ are the Haar block matrix coefficients comprising a square spatial kernel; *H* and *V* are the image fragment horizontal and vertical sizes, multiples of the kernel size K; h and v are the horizontal and vertical Haar kernel indices; i and j are the indices of the Haar transformed image; L is the number of levels of Haar transform, and

$$\mathbf{C} \in \{\mathbf{\Phi_1}, \mathbf{\Psi_1^H}, \mathbf{\Psi_1^V}, \mathbf{\Psi_1^D}\}. \quad (9)$$

L-level Haar wavelet features for $L > 1$ can be obtained by repetitive use of level-one transform, or directly using $3L + 1$ spatially averaging Haar wavelet coefficient kernels of size $K = 2^l$, with $l = 1, \ldots, L$.

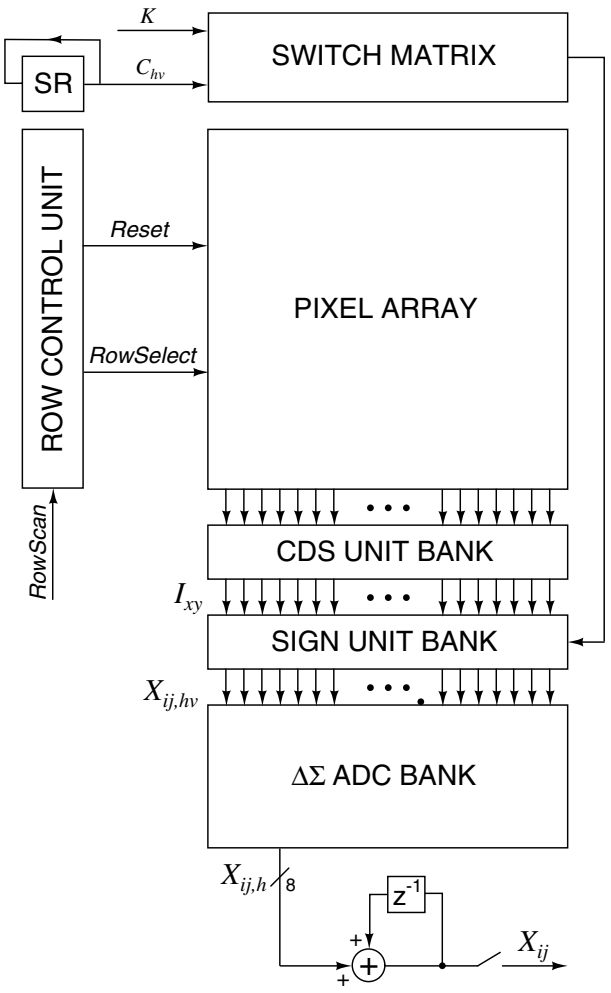

Fig. 1. Sensory wavelet processor top-level architecture.

III. Sensory Haar Wavelet Compression Processor

Image acquisition and wavelet transform computation are both performed on the focal plane of the sensory processor. Its mixed-signal VLSI architecture is depicted in Fig. 1.

A. Analog Image Acquisition

Analog image acquisition is performed by the active pixel array, the row control and the correlated double sampling (CDS) units. The active pixel comprises a resetable n^+-diffusion–p-substrate photodiode and a selectable source follower with shared column-parallel current source biased with *IbiasCol* current as shown in Fig. 2.

The row control unit generates the digital signals *Reset* and *RowSelect* controlling the integration and readout phases. Fixed pattern noise (FPN) due to pixel-to-pixel transistor mismatch is one of the dominant sources of noise in CMOS imagers. Switched-capacitor CDS unit shown in Fig. 3 suppresses FPN by subtracting the reset pixel output from the sensed pixel output. The clocks $\phi 1_{CDS}$ and $\phi 2_{CDS}$ are non-overlapping. When $\phi 1_{CDS}$ is high, the sensed pixel output

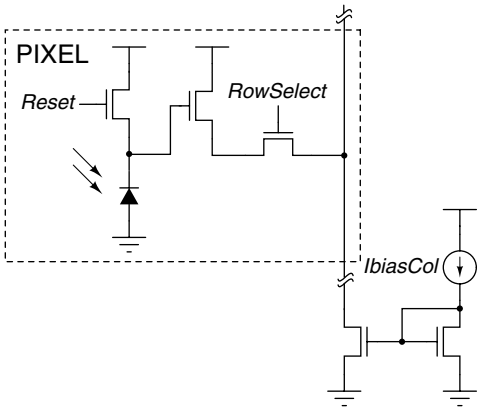

Fig. 2. Photodiode active pixel circuit.

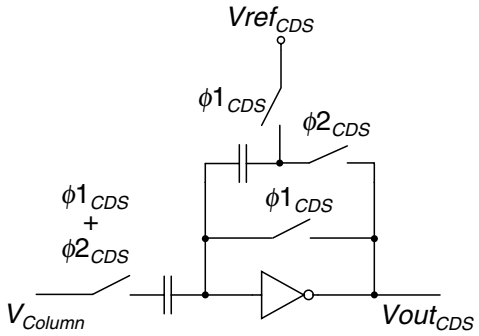

Fig. 3. Correlated double sampling (CDS) circuit.

is sampled on the input capacitor. When $\phi 2_{CDS}$ is high, the reset pixel output is subtracted from the earlier sampled signal, with the difference produced at the CDS circuit output. The amplifier is a single-stage nMOS cascoded amplifier.

B. On-Focal-Plane Haar Wavelet Transform

As detailed in Section II, L-level Haar wavelet transform requires correlating an image with $3L+1$ square spatial kernels of sizes $K = 2^l$, $l = 1, \ldots, L$. For $L = 1$, the Haar kernels are of the form (1)–(4). The presented mixed-signal VLSI architecture computes multiple kernel-image correlations in parallel within a single oversampling quantization cycle, with negligent space overhead and no time overhead as compared to a standard digital imager.

The active pixel array is integrated with a bank of column-parallel oversampling analog-to-digital converters (ADCs) as shown in Fig. 1. Each ADC is configured to sample K adjacent pixels in a column within a single quantization cycle. Uniform pixel sampling yields a digital representation of a column-wise spatial average of the acquired image. Correlation of a K-pixel-long fragment of an image column with a signed unity-amplitude one-dimensional kernel of size K is computed by combining the uniform pixel sampling and pixel output sign inversion. A double sampling sign unit circuit determines the sign of each pixel output by selecting a switched-capacitor sampling sequence order. Sign values are presented bit-serially

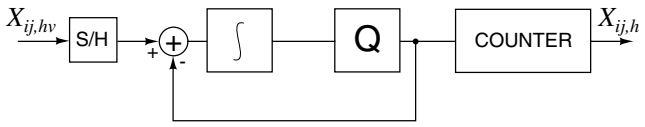

Fig. 4. Block diagram of a first-order $\Delta\Sigma$-modulated incremental A/D converter with comparator, sample-and-hold circuit, integrator and counter.

from a ring shift register, *SR*, with a sequence period of K values, synchronously with image read-out clock *RowScan*. This amounts to computing one-dimensional column-parallel spatial Haar wavelet transform:

$$X_{ij,h} = \sum_{v=1}^{K} X_{ij,hv} = \sum_{v=1}^{K} C_{hv}\ I_{xy}, \tag{10}$$

where the notation is consistent with that of (6)–(8).

The oversampling quantizer is implemented as a first-order incremental $\Delta\Sigma$-modulated ADC. The implementation is compact as the decimating filter is a simple digital counter. The block diagram of the ADC is depicted in Fig. 4.

The switch matrix routes the K different time-dependent sign signals to H/K groups of adjacent column-parallel sign unit circuits. A simple digital delay and adder loop performs spatial accumulation over K adjacent ADC outputs as they are read out:

$$X_{ij} = \sum_{h=1}^{K} X_{ij,h} = \sum_{h=1}^{K}\sum_{v=1}^{K} C_{hv}\ I_{xy}. \tag{11}$$

This mixed-signal VLSI computation realizes a two-dimensional Haar wavelet transform in (5) of up to L levels. The area overhead of sign unit circuits, switch matrix and digital accumulator scales linearly with the imager size and is negligent. As computing is interleaved with quantization, no extra computational time is required above that of raw image oversampling quantization in a baseline digital imager.

IV. Results

The architecture is experimentally validated on a 0.35 micron CMOS prototype containing a bank of 8-bit oversampling analog-to-digital converters (ADCs) computing Haar wavelet transform. Image sensory stimulus emulating an analog pixel array output is presented to the input of the ADC. Proper pixel value sign inversion is performed on the image a priori. The first-order incremental $\Delta\Sigma$-modulated ADC performs distributed image sampling and concurrent signed weighted average quantization, realizing a one-dimensional spatial Haar wavelet transform. A digital delay and adder loop implemented in software performs spatial accumulation over multiple ADC outputs. This amounts to computing a two-dimensional Haar wavelet transform. Fig. 5*(b)* depicts experimentally measured two-dimensional one-, two-, and three-level Haar wavelet transforms of *Audrey*. Fig. 5*(c)* shows the reconstructed images of the corresponding Haar wavelet transforms. The reconstructed images after being compressed using the one-level

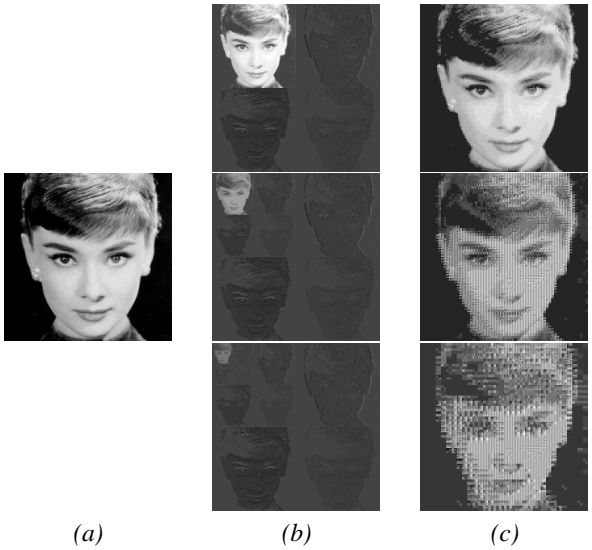

Fig. 5. *(a)*, *Audrey*; *(b)*, experimentally measured *Audrey*'s Haar transforms; and *(c)*, reconstructed images for one-level (*top*), two-level (*center*), and three-level (*bottom*) Haar transforms. Compression ratios from top to bottom are: 8.55, 16.61, and 21.58.

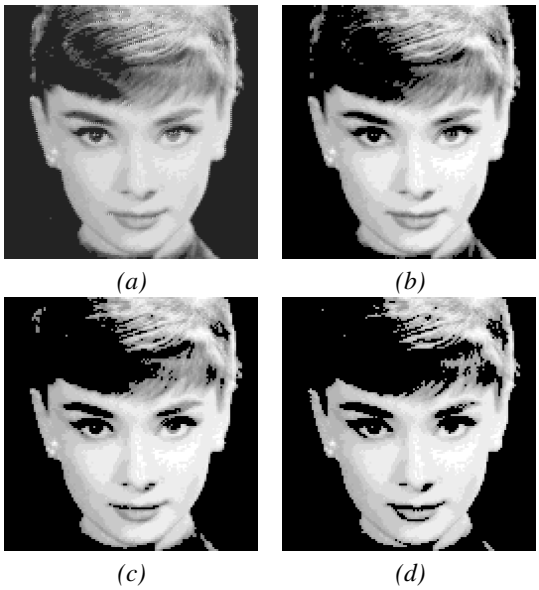

Fig. 6. Reconstructed images from experimentally measured one-level Haar wavelet transform of *Audrey* with peak signal-to-noise and compression ratios: *(a)*, PSNR=24.92 dB, CR=8.55; *(b)*, PSNR=23.07 dB, CR=9.64; *(c)* PSNR=19.14 dB, CR=10.81; *(d)* PSNR=16.29 dB, CR=12.04.

Haar transform are compared in Fig. 6 for different compression ratios. Peak signal-to-noise ratios and compression ratios are reported in the figure.

The photodiode active pixel is simulated to have the worst case integration time of 10 msec. Assuming no pipelining, at 30 frames per second frame rate, additional 23 msec are available for Haar wavelet image processing. Based on a conservative quantizer sampling rate of 40 ksps, the computational throughput is estimated to be 1.4 GMACS for SVGA image resolution. The estimate is given for the three-level Haar wavelet transform computed on the presented architecture employing $\Delta\Sigma$-modulated algorithmic ADC [12].

V. Conclusions

We present a mixed-signal VLSI architecture 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 $\Delta\Sigma$-modulated analog-to-digital conversion cycle, making focal-plane computing an intrinsic part of the quantization process. The approach yields no overhead in time and negligent overhead in area compared to a baseline digital imager employing oversampling quantization. The architecture functionality has been experimentally validated on a bank of first-order incremental $\Delta\Sigma$-modulated analog-to-digital converters fabricated in a 0.35 μm CMOS technology. The architecture yields 1.4 GMACS simulated computational throughput at SVGA imager resolution and 8-bit output resolution.

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," IEEE International Solid-State Circuits Conference (ISSCC'03), vol. **1**, 2003.

[2] E.R. Fossum, "CMOS image sensor: Electronic camera-on-a-chip," *IEEE Trans. Electron Devices,* vol. **44**, pp 1689-1698, 1997.

[3] 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.

[4] 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.

[5] F. Saffih, and R. Hornsey, "Multiresolution CMOS image sensor," Technical Digest of SPIE Opto-Canada 2002, p. 245.

[6] A. Fish, and O. Yadid-Pecht, "CMOS current/voltage mode winner-take-all circuit with spatial filtering," IEEE Int. Symp. on Circuits and Systems (ISCAS'01), vol. **3**, pp 636-639, 2001.

[7] 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.

[8] 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.

[9] M. Tabet, and R. Hornsey, "CMOS image sensor camera with focal plane edge detection," *Canadian Conference on Electrical and Computer Engineering,* vol. **2**, pp 1129-1133, 2001.

[10] G. Kaiser, "The fast Haar transform," *IEEE Potentials,* vol. **17** (2), pp 34-37, 1998.

[11] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian Detection Using Wavelet Templates," in *Computer Vision and Pattern Recognition*, pp 193-199, 1997.

[12] 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.