An Active Pixel CMOS Separable Transform Image Sensor

Yu M. Chi[1,2], Adeel Abbas[3], Shantanu Chakrabartty[4] and Gert Cauwenberghs[2,1]

[1] Department of Electrical and Computer Engineering, University of California, San Diego, CA 92093

[2] Division of Biological Sciences, University of California, San Diego, CA 92093

[3] Thomson, Corporate Research, Princeton, NJ 08540

[4] Department of Electrical and Computer Engineering, Michigan State University, Lansing, MI 48824

Abstract— This paper presents a 128×128 charge-mode CMOS imaging sensor that computes separable transforms directly on the focal plane. The pixel is a unique extension of the widely reported Active Pixel Sensor (APS) cell. By capacitively coupling across an array of such cells onto switched capacitor circuits, computation of any unitary 2-D transform that is separable into inner and outer products is possible. This includes the Walsh, Hadamard and Haar basis functions. This scheme offers several advantages including multiresolution imaging, inherent de-noising, compressive sampling and lower integration voltage and faster readout. The chip was implemented on a $0.5 \mu m$ CMOS process and measures $9mm^2$ in MOSIS' submicron design rules.

Fig. 1.   Three-level dyadic wavelet decomposition.

I. INTRODUCTION

The digital image sensor finds its use in almost every field of engineering. Its applications vary from high bandwidth line of sight laser communication to microscopic imaging and its ubiquity ranges from being inside the human body to frontline in a battlefield. Due to these factors, the demand for high-resolution, low-noise imaging is expected to grow in the next decade. Likewise, the high demand for portable multimedia devices has imposed low power operation constraints in imaging systems.

Most commercial imagers are fabricated today on charge coupled device (CCD) process. The main advantages of CCD are low-noise and high fill factors. These devices are capable of accurate, high resolution and low noise imaging. However because they need a programmable processor to operate with, power consumption by a CCD is excessively large (typically in the order of Watts). Another problem is that CCD fabrication processes are much more expensive than digital CMOS processes.

CMOS based image sensors have been the subject of study for several years [1], because they can provide answers to some of the CCD limitations. Amongst several reported designs, the passive pixel sensor (PPS) [3] is the simplest. Because it is composed of only one transistor and a photodiode, the PPS allows highest fill-factor (and largest integration density). However a well-known limitation in PPS architecture is its vulnerability to noise, because the charge is extremely sensistive to disturbances on the column line. Active Pixel Sensor (APS) [4] and [5] is based on a slightly different formulation. The charge on the gate of a transistor is gradually removed by a reverse-biased photodiode. This amount of charge removed is directly related to the intensity of light.
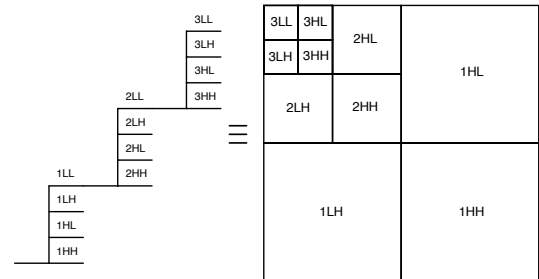
The APS has been successfully used in both voltage and current mode implementations. For voltage mode (e.g. [4]), the output is buffered using source-follower configuration. For current mode, the output is developed from gate-source voltage difference across a PMOS in saturation (for instance as in [6]).

In this paper, we present the design of a CMOS imager that capacitively adds outputs from APS cells to compute Wavelet transform of the image (on the focal plane). Similar work on successfully similar 2-D transforms on the focal plane have either used a PPS [13], were restricted to limited block sizes [7] or involved more complicated computational circuits [2]. In this work we first present, an full APS based wavelet transform imager that supports basis functions that span the full resolution of the pixel array. Experimental results from the fabricated chip are then shown, along with a discussion of applications towards compression and compressed sensing [11] [12] [13].

II. ALGORITHM FOR 2D HAAR TRANSFORM

The wavelet transform has emerged as a powerful tool in many signal and image processing applications. A three-level two-dimensional dyadic wavelet decomposition is shown in Fig. 1. Because of the high spatial correlation between pixels in practical images, most of the energy appears in the low-pass sub-bands. Each dyadic decomposition results in four sub-bands: HH captures diagonal edges, LH captures vertical edges, HL captures horizontal edges and LL captures all the low-pass energy. One popular implementation of this decomposition involves separable 2D scheme, where wavelet filters are applied along rows, followed by columns.
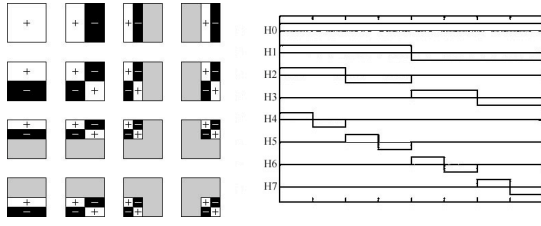
Fig. 2. Various 2-D basis functions for the transform (left) which can be constructed from the outer product of 1D (right) signals.

Fig. 3. Schematic of the active pixel cell. The output from pixels is capacitively summed on wire $\text{Vpix}_{(j)}$ through the coupling capacitor C_{pix}.

Haar is regarded as the simplest wavelet transform because its low-pass sub-band is the average and high-pass sub-band is the difference of data samples. In Fig. 2 (right), Haar coefficients for a 1D 8-point signal are shown. The scaling function (H0) is simply the sum of all data samples, while the wavelet functions (H1-H7) are differences at different scales. Similarly in Fig. 2 (left), Haar coefficients for a 2D, 4×4 image are displayed. Pixels are added wherever plus sign appears and subtracted where minus sign appears. Pixels in gray region do not contribute in computation. Thus all pixels is multiplied by either +1, −1 or 0 and then summed. Note that normalization is needed to make sure that dynamic range of the output remains in limits.

III. VLSI Implementation

Fig. 3 shows pixel configuration for our design. The pixel is based on the conventional APS design, however it adds one more transistor and a linear capacitor. At the start of computation cycle, RST (reset) goes high to charge up the gate of transistor M2. The photodiode discharges this gate, and the amount of discharge corresponds to the light intensity of pixel. $RS_{(i)}$ is a digital signal that controls row-select operation and is different for rows. Another important difference here is that biasing transistor M4 is present in each pixel. For power considerations, M4 has to be biased in sub-threshold. Our results indicate that the range of discharge on the gate of M2 does not have to be as large as in APS. Thus M4 can be biased at a smaller current than for a readout transistor in APS [8]. An M4 current of $38nA$ will set a lower bound on the power consumption to be $3.1mW$.

Shown in Fig. 4 is the peripheral circuit for a single column. Assume clocks S1/H1 and S2/H2 to be non-overlapping. The circuit consists of a popular switched capacitor amplifier. If $RS_{(i)}$ is low (i.e. M3 is off) during sample operation S1, the amplifier samples source follower output, $V_{(i,j)}^{int}$ where subscript (i,j) denotes location of the pixel. In the hold phase H1, if $RS_{(i)}$ goes high, the amount of charge transferred from each C_{pix} to C_1 is $C_{pix}(V_{(i,j)}^{int} - V^{aps})$. Using a very similar argument, if $RS_{(i)}$ goes high during S1 and low during H1, the amount of charge transfered is $C_{pix}(V^{aps} - V_{(i,j)}^{int})$. If $RS_{(i)}$ remains high during both S1 and H1, no charge is transfered. Therefore, if V^{aps} is tuned to be equal to pixel voltage when there is no light, the above equations correspond to a multiplication by +1, −1 or 0. The output of the amplifier Q1 (Fig. 4) can be written as

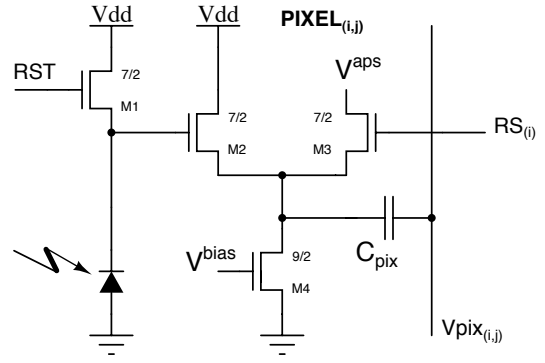$$V_{(j)}^{amp} = V^{ref} + \frac{C_{pix}}{C_1} \times \sum_i (V_{(i,j)}^{int} - V^{aps}) \times RS_{(i)} \quad (1)$$

In order to achieve a higher dynamic range, digital inputs are used to program capacitor C_1 to be either $4C_{pix}$, $16C_{pix}$ or $64C_{pix}$. Setting C_1 to $4C_{pix}$ will help resolve wavelets at a finer scale (in extreme case, only two pixels along a column contribute). When computing scaling function (i.e. summing charge of all the pixels), C_1 has to be $64C_{pix}$ to prevent the output $V_{(j)}^{amp}$ to saturate.

In what has been presented so far, we compute partial products along rows. We need to compute partial product along columns also. This is clearly evident from Fig. 2 (left) because haar wavelets are organized in a particular checkerboard like fashion. Moreover, for certain coefficients some of the columns have to be multiplied by zero. One way to achieve this is by multiplying $V_{(j)}^{amp}$ by +1, −1 or 0 and then summing (by capacitive sensing). We do this through complementary signals $CS_{(j)}$ and $\overline{CS}_{(j)}$ (CS is shorthand for column select) and second-stage switched capacitor amplifier (Q2). Assuming we are in H1 phase and S2 goes low, if $CS_{(j)}$ is low, we sample $V_{(j)}^{amp}$. After H2 goes high, if $CS_{(j)}$ goes high so that V^{ref} is connected to the bottom plate of the capacitor, the amount of charge transferred is $4C_{pix}(V_{(j)}^{amp} - V^{ref})$. Similarly, by setting $CS_{(j)}$ to be high during S2 and low during H2, an equal but opposite charge transfer can be achieved i.e. $4C_{pix}(V^{ref} - V_{(j)}^{amp})$. Connecting V^{ref} to the capacitor during S2 as well as H2 results in no charge transfer (multiplication by zero). Therefore during H2, the output of the amplifier Q2 is given by

$$W_{out} = \frac{4C_{pix}}{C_2} \times \sum_j (V_{(j)}^{amp} - V^{ref}) \times CS_{(j)}$$

Combining with (1), the above equation can be written as

$$W_{out} = V^{ref} + \frac{4C_{pix}^2}{C_2 C_1} \sum_j (\sum_i (V_{(i,j)}^{int} - V^{aps})RS_{(i)})CS_{(j)}$$

Where $RS_{(i)}$ and $CS_{(j)}$ are either +1, −1 or 0, depending upon coefficients being computed. The capacitor C_2 can be programmed to be either $16C_{pix}$, $64C_{pix}$ or $256C_{pix}$. Thus, if $C_1 = 4C_{pix}$ and $C_2 = 16C_{pix}$, a 1V variation in one of the pixels will cause about 0.0625V variation at the output.

IV. Experimental Results and Applications

The fabricated image sensor was integrated in a test environment consisting of a PCB, microcontroller for timing signal
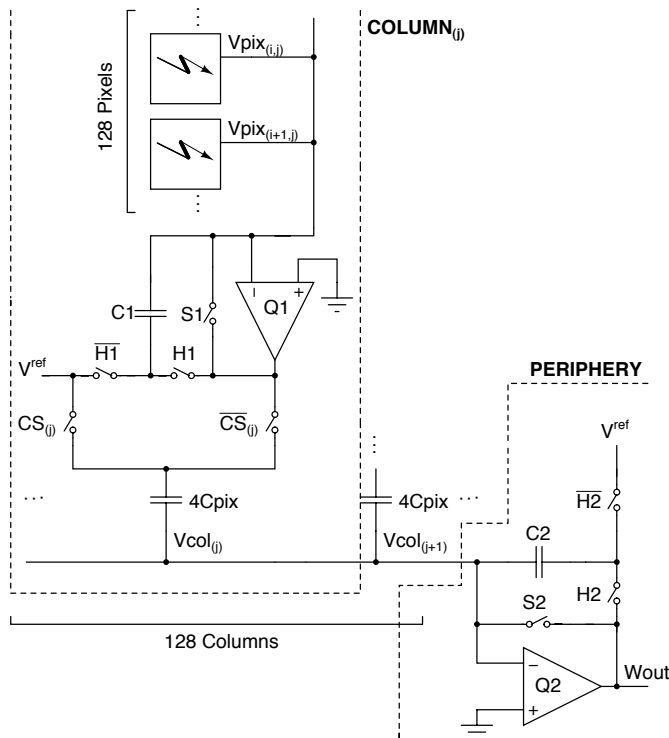
Fig. 4. Schematic for the peripheral cell for each column and the second-stage amplifier. The connection for the second stage amplifier is similar to the one used for the pixel to the first stage where all of the outputs of the first stage amplifier are capacitively summed.
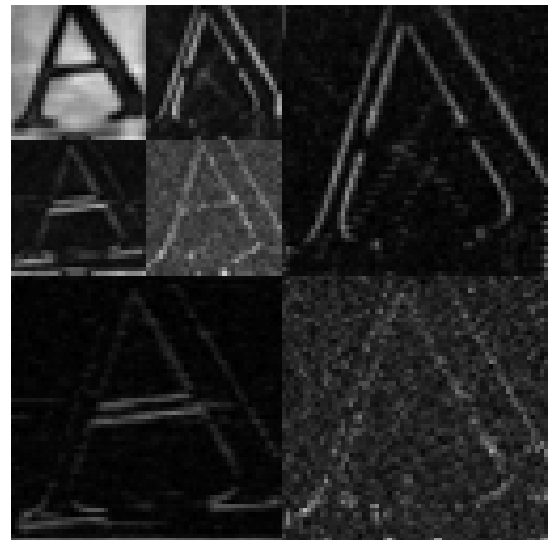

Fig. 5. Acquired 2-level Haar transform of the letter 'A' from imager. Intensity values in each subband were independently scaled to reveal the signal. Most of the sensor noise is captured by the highest frequency subbands.

generation and a 12-bit ADC to acquire the analog output of the computation circuits for transfer to a PC for analysis.

A. Wavelet Transform

Figure 5 shows a sample output of the wavelet transform where the coefficients for a 2-level dyadic Haar decomposition. The appropriate row/column outer-product vectors for each subband were loaded into the shift registers of the sensor from the microcontroller's memory and used to compute the Haar coefficients. A test image consisting of a printed letter 'A' was used under normal indoor lighting conditions.

During experimental testing, it was discovered that the readout circuit exhibited a large degree of temporal noise. The displayed image is a 16-frame time average to improve the clarity of the result.

In addition to the expected wavelet transformed output, it is interesting to examine the output of the two HH subbands, which exhibit the largest degree of fixed pattern noise (FPN). The presence of the FPN in the HH subbands and the lack of FPN in the HL and LH subbands implies that the greatest contribution to mismatch is the random fabrication error of the pixel coupling capacitor, C_{pix}, rather than the larger capacitors between the row/column outer-product computational blocks. This is useful in de-noising applications since it is expected that FPN will appear in the HH subbands and can possibly be thresholded out.

For compression applications, two approaches may be taken.

First coefficients can be acquired starting from the lowest frequency until a certain quality point or coefficient count can be reached. A more sophisticated approach is to intelligent select which coefficients to acquire using a technique like embedded zero tree coding [10]. Either case avoids having to fully readout every single coefficient location (the entire 128×128 array).

B. Compressive Sampling

Recent work has suggested a radically different approach to data compression. Rather than computing standard space-frequency transforms (DCT, Wavelet) followed by entropy coding, compressive sampling [11] [14] suggests that for images, a signal can be recovered from an incomplete set of randomly selected basis functions. The advantages here is that the data reduction occurs directly at the sensor, rather than later in the signal processing chain, hence reducing the computational load on the external processor. Unlike traditional compression architectures, the burden of computation is now shifted to the decoder, rather than encoder. For applications like low power sensors, this is an advantageous tradeoff since back end decoders are typically much more powerful and less constrained than the front end sensing node.

Successful compressive sampling imaging applications have used PPS sensors over smaller 16×16 blocks [13] which were not optimized for binary coefficients $(1, -1)$ or a mechanical array of micro-mirrors [12]. In this section we present an architecture for compressive sampling that utilizes a full APS sensor.

For compressive sampling, a Walsh rather than Haar basis was used since it consists of basis functions that span the entire imaging array, rather than the localized features of the Haar and is hence better suited for compressed sampling. Random 2-D basis functions were generated by randomly
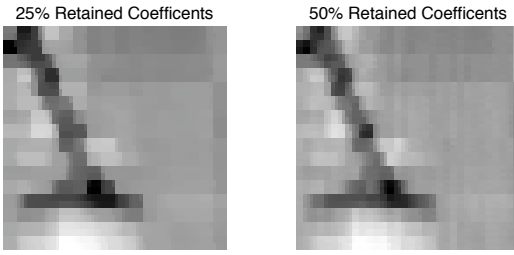
25% Retained Coefficents   50% Retained Coefficents

Fig. 6. Reconstructed image via Compressive Sampling [14] for retaining 1/4 and 1/2 of all the possible coefficients of a Walsh transform over a 64×64 pixel region.

loading different combinations of 1-D basis vectors into the row and column shift registers to construct the under-sampled measurement matrix.

Images are reconstructed by using solvers [14] that search for an the image that exhibits the least total variance which fits the observations from the measurement matrix, subject to an error factor to account for acquisition noise. Sample output from the compressive sampling scheme is shown in Figure 6. Due to memory and runtime limitations in the compressive sampling reconstruction algorithm [14] the image size was limited to a 64×64 patch at this time.

V. CONCLUSION

We present a 128×128 CMOS sensor which directly computes 2-D separable transform coefficients at the focal plane. Table I summarizes the main features of the fabricated chip. The image sensor computes transforms with basis functions up to the size of the full array resolution.

While the performance of binary valued transforms (Haar, Walsh) is typically inferior to more complex basis sets, the tools presented are still useful for a variety of image processing tasks like image compression via traditional transforms or the new compressive sampling. Because these transforms can be efficiently implemented in focal plane circuits, the sensor is ideally suited for applications calling for low-power electronics followed by low complexity signal processing.

TABLE I

SUMMARY OF THE CHIP CHARACTERISTICS

Technology	$0.5\mu m$ CMOS 3M2P
Area	$3mm \times 3mm$
Dimensions	128×128
Pixel Size	$17\mu m \times 17\mu m$
Fill factor	40%
Supply voltage	3.3V

Fig. 7. Micrograph of fabricated 3mm×3mm image sensor.

[3] C. Wang I. L. Fujimori and C. G. Sodini, "A 256×256 cmos differential passive pixel imager with fpn reduction techniques," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 2031–2037, Dec. 2000.
[4] R. H. Nixon et. al., "256×256 CMOS active pixel sensor camera on a chip," *IEEE Journal of Solid State Circuits*, vol. 31, pp. 2046–2052, Dec. 1996.
[5] R. Etienne-Cummings, "Single capacitor single contact active pixel sensor," *IEEE Intl. Symposium on Circuits and Systems*, vol. V, pp. 177–180, May 2000.
[6] V. Gruev and R. Etienne-Cummings, "Implementation of steerable spatiotemporal image filters on the focal plane," *IEEE Trans. on Circuits and Systems - II*, vol. 49, pp. 233–243, April 2002.
[7] L. Qiang and J. Harris, "A novel integration of on-sensor wavelet compression for a CMOS imager," *IEEE Intl. Symposium on Circuits and Systems*, May 2002.
[8] K. Salama and A. E. Gamal, "Analysis of active pixel sensor readout circuit," *IEEE Trans. on Circuits and Systems - I*, vol. 50, pp. 941–944, July 2003.
[9] B. Fowler H. Tian and A. E. Gamal, "Analysis of temporal noise in cmos photodiode active pixel sensor," *IEEE Journal of Solid State Circuits*, vol. 36, pp. 92–101, Jan. 2001.
[10] J.M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficents," *IEEE Transactions on Signal Processing*, vol. 41, No. 12, pp. 3445-3462 December 1993.
[11] E. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, March. 2008.
[12] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, T. Sun, K.F. Kelly and R.G. Baraniuk "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, March. 2008.
[13] R. Robucci, L. K. Chiu, J. Gray, J. Romberg, P. Hasler and D. Anderson "Compressive sensing on a CMOS separable transform image sensor," *IEEE Intl. Conference on Acoustics, Speech and Signal Processing*, March. 2008.
[14] E. Candes and J. Romberg "l_1-magic," *Software at www.l1-magic.org*, 2006.

REFERENCES

[1] E. R. Fossum, "CMOS image sensors: electronic camera on a chip," *IEEE Trans. On Electron Devices*, vol. 44, pp. 1689–1698, Oct. 1997.
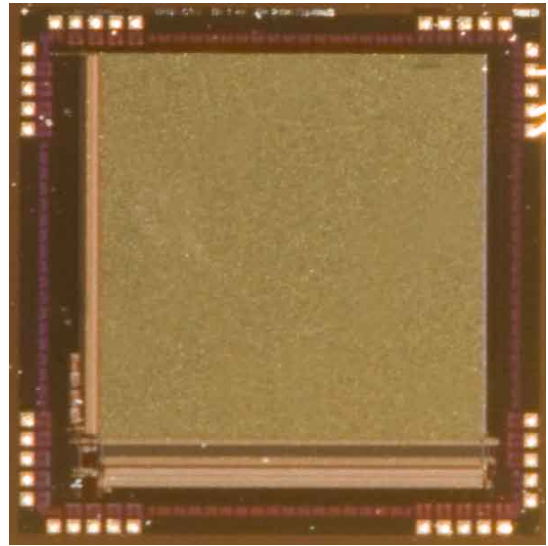[2] A. Olyaei and R. Genov, "Focal-Plane Spatially Oversampling CMOS Image Compression Sensor," *IEEE Trans. on Circuits and Systems I*, 45:1 26-34, 2007.