An Experimental Evaluation of Error Spectrum Shaping Applied to Mixed-Signal Image Convolutions

Brent Buchanan, *Member, IEEE,* and Martin Brooke, *Member, IEEE*

***Abstract*—In this experimental research paper, analog circuits are manufactured and their performance evaluated for suitability in performing FIR image filtering with the convolution kernel and data altered via error spectrum shaping and oversampling so as to preclude corruption by circuit imperfections. Similar to the one–dimensional (1-D) binary signals output by Sigma Delta analog-to-digital converters, the representational noise caused by the circuit's resolution inaccuracy is pushed into an unused portion of the spectrum in these analog signals, permitting the inband portion of the oversampled signal to be more effectively represented and processed by imperfect circuits. An analysis of image convolutions performed using the circuits' data establishes that ESS is successful at reducing the computational error in certain analog image convolutions.**

***Index Terms*—Analog CMOS, analog signal processing, error spectrum shaping, FIR image convolutions, noise shaping.**

I. Introduction

MOST electronic signal sensing and processing systems attempt to maintain a high signal-to-noise ratio (SNR) throughout their internal data paths, and this dictates that the circuits operating on these signals then similarly maintain high levels of precision, functional performance, and integrity. For many systems, this typically involves raising circuit performance to the level of sustaining the signal content, in contrast to subordinating the signals' traits to fit those of the components. Efforts toward this end can be expensive and generally produce designs that are catastrophically intolerant of even a single element's immoderate variance.

A different approach is through the adroit manipulation of signal data and noise, where better matching of signals with device characteristics greatly simplifies hardware complexity. Switching power supplies and Sigma Delta ($\Sigma\Delta$) analog-to-digital converters (ADCs) are examples of where such signal conditioning reduces the overall obligations on a substantial portion of their analog circuitry.

Unlike their charge-coupled device (CCD) predecessors, CMOS imagers can incorporate a wide range of signal processing circuitry on the same die with the photodetectors, making cameras or imaging systems on a chip possible. As a fundamental signal processing mechanism, convolutions are going to be one of the operations that vision systems must perform, making the development of high performance convolution architectures an elemental precursor to the construction of complex CMOS imaging and vision systems.

Use of a mixed-signal design approach allows incorporating the best features of both analog and digital circuits in such systems. Specifically, analog circuits are very much faster and significantly smaller than their digital counterparts, but analog computations can be expected to include uncontrollably variable offset, gain, and nonlinearity. In an image convolution, such errors would be manifest as spatial noise and so would be subject to spatial manipulations. Error spectrum shaping (ESS) is a signal processing tool that dissipates this kind of representational error, making it possible to tradeoff enough circuit complexity for signal complexity that small, nonideal analog CMOS circuits can perform image convolutions with sufficient accuracy to be useful.

A. Analog CMOS Image Convolutions

Numerous analog convolution architectures have been constructed in both CCD and CMOS VLSI processes. What typically prevents these devices from performing general image filtering tasks is that either there is inadequate control over their convolution kernel coefficients, the spatial support of their convolution kernels is far too small to affect any but the very highest image frequencies, or computational inaccuracy attributable to device mismatch severely limits output SNR.

Various imagers/processors that smooth or bandpass the incident image via an assortment of passive, active, and nonlinear resistive networks have been constructed: Mahowald [1] and Mead [2], Bair *et al.* [3], Boahen *et al.* [4], Kobayashi *et al.* [5], and Harris *et al.* [6]. While all of these networks are scalable, and smoothing is an incidence of a convolution, the utility of this particular filtering operation is extremely limited, its range and specific characteristics are somewhat difficult to control, and it is not readily extensible to general convolutions.

Chong *et al.* [7], use current mirrors to perform a 3×3 DOG/LOG convolution, and Ward *et al.* [8] have used current mirrors to convolve an incident image with a 3×3 Sobel kernel [9]. The kernel size is too small to support filters beyond the very highest image frequencies, and is not well scalable to either general convolution coefficients or easily scalable to larger kernels. Nilson [10] comparably uses Gilbert multipliers to provide and transmit convolution kernel weights. While all of the weights in this filter are negative, the scheme could be extended to include positive coefficients. Again, kernel expansion is limited; note that the smallest on/off-center surround

Manuscript received August 22, 2001; revised May 8, 2003. This paper was recommended by Associate Editor A. Petraglia.

B. Buchanan was with the Georgia Institute of Technology, Atlanta, GA 30322 USA. He is now with Philips Semiconductors, Longmont, CO 80501 USA (e-mail: brent.buchanan@philips.com).

M. Brooke was with the Georgia Institute of Technology, Atlanta, GA 30322 USA. He is now with Duke University, Durham, NC 27708 USA.

Digital Object Identifier 10.1109/TCSII.2003.820236

fields (i.e., convolution kernels) in the human retina have been shown to have well over an order of magnitude more input samples [11].

Allen *et al.* [12] row-wise shift the incident image into a pipeline that buffers the previous two rows. Per-column cells convolve the three-row subimage with the three hexagonal equivalents of Prewitt's directional gradients [10] using current mirrors, with the results assembled off-chip. Though less so than in fully parallel implementations, spatial support and kernel coefficients are very limited.

Nitta *et al.* use variable photodetector response to scale the pixel value [13]. This method permits only a single multiplication per pixel and so can only compute one convolution output sample at a time, though it has image-wide spatial support.

B. **Remedial Concepts**

The conceptual complexities of discrete filtering, analog CMOS design, and noise shaping are far from trivial, and are sufficient to prevent including a genuinely adequate introduction or review in a paper of this scope and length. Among many others available in the literature: Oppenheim *et al.* [14] provide a thorough investigation of one-dimensional (1-D) discrete signal filtering with Dudgeon *et al.* [15] extending the concepts to multiple dimensions; Candy *et al.* [16] provide a thorough description of noise shaping in their review of oversampling methods for analog-to-digital (A/D) and digital-to-analog (D/A) converters (with the extension to multiple dimension signals largely being an inclusion of dimension indices to the otherwise 1-D signal variables); and, Allen *et al.* [17] provide a current overview of analog CMOS design concepts.

ESS is essentially a generalized term for the noise shaping technique that is used in such sampling operations as $\Sigma\Delta$ quantization [16] and digital half-toning [18]. It is a mechanism for transforming point-wise resolution information into a signal's temporal or spatial domain [19], [20], whereby representational error is accumulated and then used to influence the choice of subsequent output sample values. One everyday use of ESS that even most laypeople are familiar with is the sequence of charges that traditionally occur when purchasing like items that are priced to include fractions of a currency's smallest denomination. For example, purchasing the first in a collection of like items priced at 3/$1 (i.e., 33 1/3¢ each; $1=100¢) typically results in a charge of 34¢ (where the vendor's quantization algorithm is to always round up; rounding to the nearest integer works equally as well for this demonstration), with the overcharge of 2/3¢ then being subtracted from the cost of the next like item. The second-like item is then charged at 33¢, with the accumulated overcharge of 1/3¢ being subtracted from the cost of the third-like item. Once a third like item is charged at 33¢ the accumulated error is zero, and the sequence of 34¢, 33¢, 33¢ then repeats as additional like items are purchased. Another widely familiar application of ESS is the use of leap days to fit a calendar made of whole days onto a year that technically contains 365.2425 days: a leap day is injected into the calendar every four years, omitted once every hundred years (i.e., every 36 524.25 days, of which 24 were leap days), but not so omitted every 400 years (reducing the accumulated error to zero then every 400 years).

Of interest from an engineering perspective are both the spectral performance and noise behavior of such operations. While the specifics are dependent on the ESS algorithm used, the goal is to shape the spectrum of the representational error such that it is minimized in the frequency band of interest at the expense of increasing it elsewhere in the spectrum. Oversampling can be used to provide additional bandwidth specifically intended for the disposal of this noise (the "disposal band"). As an example, the differentiation of quantization error in first-order $\Sigma\Delta$ oversampled A/D conversion serves to attenuate the portion of the quantization noise that is in the signal band ('inband') by displacing a great deal of it into the spectrum that lies above the input signal's Nyquist frequency. Though the total noise energy remains fixed (i.e., the full-spectrum SNR remains unchanged), the portion of it inband is substantially reduced, in turn increasing the inband SNR. At whatever point in the signal processing sequence it is convenient, if even necessary at all, the inband portion of the signal can be extracted with a bandpass filter. In the case of first-order $\Sigma\Delta$ oversampled A/D conversion, the pass band of interest is from DC to the input's Nyquist frequency, so a low pass filter is used in any such restorative step. Returning to the ESS example of like items priced at 3/$1, note that low-pass filtering (i.e., averaging) the output sequence of 34¢, 33¢, 33¢ reproduces the input sequence of 33 1/3¢, 33 1/3¢, 33 1/3¢.

With convolution in the time/image domain being equivalent to multiplication in the frequency domain, the primary interest is in the behavior of the operations that occur inband, with whatever events or chaos that occurs in the disposal band, where noise operates upon noise, a largely independent and irrelevant matter. Thus, the convolution of an ESS'd image with an ESS'd kernel will produce a result with ESS-like characteristics: the signal of interest confined to a particular spectral band and substantial noise energy outside of that.

C. **Research Overview**

The fundamental hypothesis in this effort is that ESS can be used to enhance the inband SNR of image convolutions performed with inaccurate analog circuits. This is motivated by the observation that the most sophisticated image acquisition and processing systems in existence today are constructed of biological components that are, relative to man-made devices, poorly matched in size and response, vary in quality with time and other conditions, and operate at bandwidths that barely exceed O(1kHz) [11]. Rather than attempt to modify the circuit architecture with improvements to some enhanced level of performance or to use the set of convolution coefficients and image sample values that lie nearest those that are ideal, ESS has been used to select a subtly different set of comparably inexact coefficients and image samples, creating coefficient/data sets that in particular have spectral characteristics such that the noise they produce lies outside of the signal band.

A convolution consists of many multiplications and a summation. Salient features of the particular current-mode convolution architecture that is used in this research are diagramed in Fig. 1. Since the current-mode summation can be performed virtually without error using only wires and a node, the quality of the

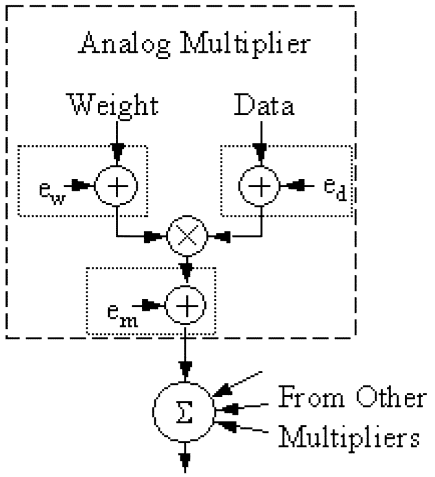


Fig. 1. Spatial noise model of current-mode analog convolution.

multiplier array largely sets the performance of the operation. Various error sources corrupt the array multipliers' result: e_w and e_d, both of which contain quantization noise and random spatial error, and e_m, which in an aggregate sense (i.e., measured from multiplier to multiplier) consists of only random error. The fundamental assumption made here is that the quantization noises and e_w's random error component can be attenuated with ESS and that the effects of the remaining error sources are subject to noise reduction with oversampling.

To evaluate this hypothesis, a simple first- and second-order ESS algorithm have been selected (see the following section), and then step-wise following the signal flow in Fig. 1, its performance sequentially quantified against the error sources and their increasingly combined effects. First, ESSs inband attenuation of both quantization noise and random gaussian error was evaluated (i.e., its effect on models of e_w and e_d). Second, convolutions of images corrupted by quantization and random gaussian error were performed and measured with and without ESS (i.e., its effect on the overall system model short e_m). Lastly, convolutions of ESS images and kernels using actual circuit measurements were performed; note that this final test suplants the convolution model in Fig. 1 in that it includes the effects of any potentially unmodeled error sources, operations, or relationships. Space constraints limit this paper to only discussing this last step, with the previous two largely superceded by its results.

To obtain the circuit data, CMOS circuits that perform the array multiplier's function have been designed and constructed, and measurements made of their performance to quantify the effects of the noise sources. Lastly, an analysis of image convolutions performed using the circuits' data establishes that ESS is successful at reducing the computational error in analog image computations.

II. Error Spectrum Shaping

ESS is a mechanism for transforming point-wise resolution information into the spatial domain [19], [20]. It is a generalized term for the noise shaping done in $\Sigma\Delta$ ADC [16] and is commonly employed in half-tone printing [18] and other aspects of signal processing. For example, both Figs. 2 and 3 are

Fig. 2. Lenna $512^2\times$ 1-b simple binary quantization.

Fig. 3. Lenna $512^2\times$ 1-b first-order ESS quantization.

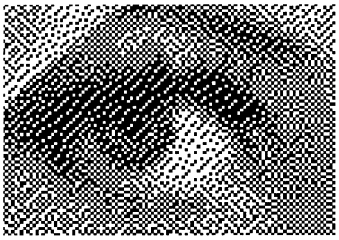


Fig. 4. Detail section of lenna $1024^2\times$ 1-b first-order ESS quantization. Squint to approximate the restorative effect of low pass filtering.

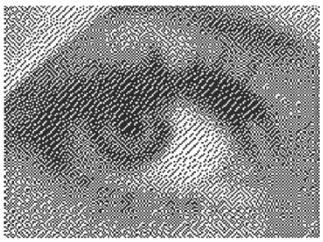


Fig. 5. Detail section of lenna $2048^2\times$ 1-b. Squint to approximate the restorative effect of low pass filtering.

1-b images derived from an 8-b image (note: the various steps involved in transferring these images into print may have damaged them). In the ESS image, the otherwise lost point-wise resolution data is encoded in the high frequency variation of the samples; low-pass filtering (i.e., local averaging) then is the appropriate step necessary to restore the original image. Oversampling provides additional spectrum where such resolution information can be stored; compare Fig. 4 with its higher frequency counterpart, Fig. 5. The greater the ESS algorithm's spatial support (i.e., polynomial order), the more efficiency that it has in capturing the resolution information; see Fig. 8.

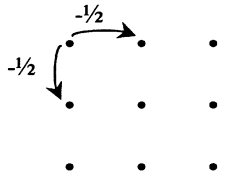


Fig. 6. Propagation of error in the first-order ESS algorithm.

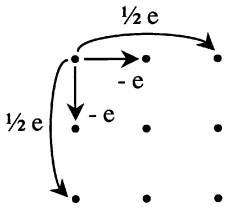


Fig. 7. Propagation of error in the second-order ESS algorithm.

Recall that in first-order 1-D Sigma Delta A/D conversion that the output signal is the sum of the input signal and the first-order difference of the error signal:

$$y_i = x_i + (e_i - e_{i-1})$$

Expanding this to two-dimensional (2-D) with the error signal's energy evenly divided between the dimensions yields a simple first-order 2-D ESS algorithm (strictly speaking, this is first order per dimension)

$$\begin{aligned} y_{i,j} =& x_{i,j} + \frac{1}{2}(e_{i,j} - e_{i-1,j}) + \frac{1}{2}(e_{i,j} - e_{i,j-1}) \\ =& x_{i,j} + e_{i,j} - \frac{1}{2}e_{i-1,j} - \frac{1}{2}e_{i,j-1} \end{aligned}$$

This first-order ESS error signal's propagation is graphically illustrated in Fig. 6. This is the first-order ESS algorithm used in this research; Figs. 3–5 illustrate its application.

A simple second-order ESS algorithm can be similarly constructed (strictly speaking, this is second-order per dimension):

$$\begin{aligned} y_{i,j} =& x_{i,j} + \frac{1}{2}(e_{i,j} - 2e_{i-1,j} + e_{i-2,j}) \\ &+ \frac{1}{2}(e_{i,j} - 2e_{i,j-1} + e_{i,j-2}) \\ =& x_{i,j} + e_{i,j} - e_{i-1,j} - e_{i,j-1} + \frac{1}{2}e_{i-2,j} + \frac{1}{2}e_{i,j-2} \end{aligned}$$

This second-order ESS error signal's propagation is graphically illustrated in Fig. 7. This is the second-order ESS algorithm used in this research.

Much more elaborate, efficient, and isotropic ESS algorithms obviously exist, error-diffusion half-toning algorithms have been reported with several dozen coefficients [21], but these two are adequate for the task at hand.

The standard SNR equation used for quantifying the results of image processing operations is

$$\mathrm{SNR} = 10\mathrm{Log}\left[\frac{\sum (s_{ij})^2}{\sum (y_{ij} - s_{ij})^2}\right]$$

where s_{ij} represents the original image samples, y_{ij} represents altered image samples, and the summations are taken over the image's entire range [22]. Being a point-wise process (i.e., it says nothing about the spatial interaction between pixels), this encompasses the complete spectral content of each image, and counts all of the energy in the original image as "signal" regardless of its actual bandwidth. ESS'd signals contain a substantial noise element outside of the signal band that is inappropriately counted as "signal" when using this metric. What is desired instead is to only count the inband portion of the spectrum as signal.

Parseval's theorem equates the total energy in a sequence to its total spectral energy

$$\sum |x[n]|^2 = \frac{1}{\mathrm{N}} \sum |\mathrm{X}(n)|^2, \quad \mathrm{N} = \text{sequence length}$$

This provides a mechanism for converting the standard point-wise SNR metric from sample space to frequency space:

$$\mathrm{SNR} = 10\mathrm{Log}\left[\frac{\sum |\mathrm{S}_{ij}|^2}{\sum |\mathrm{Y_{ij}} - \mathrm{S}_{ij}|^2}\right]$$

Rather than compute this over the entire spectrum of the image, it can be limited to any spectral band of interest, such as the inband component of ESS'd images.

III. Circuits

The circuit's purpose in this research is two-fold: to provide real-world error data for evaluation in ESS convolutions (e_w, e_d, and e_m in Fig. 1), and to insure the inclusion of any noise sources or operations omitted from the spatial noise model of Fig. 1.

A. Archtetural Considerations

Each output point in a convolution consists of many multiplications and a summation. Image data is typically nonnegative (ranging from 0 to full-scale), while convolution weights may be negative or even complex, necessitating multiquadrant multiplications and the potential for a complex summation.

The multisummand addition encourages current-mode summations, where any number of values can be added using only wires and a node. Particularly advantageous to such current mode additions is that they are not necessarily affected by MOS device mismatch so long as the driving circuits have a high enough output resistance; in MOS design, output resistance can be made as arbitrarily high as desired using cascoding. Other advantages to current-mode computation are that currents can be easily replicated and conveyed with negligible loss. Also, current-mode multiplication can be compactly performed with current mirrors.

Noise shaping of the kernel coefficients requires the ability to store and alter their values. Currents can not be directly stored, and though they could be generated from a stored voltage, temporary storage of analog values has its general problems: circuit complexity, decay, destructive read-out, and corruption; all of these are issues that threaten to distort and degrade the results of this research and are beyond its scope. Digital storage satisfactorily addresses these needs, and though digital quantization produces yet another noise source, its effects can be easily calculated and then accounted for in the final results.

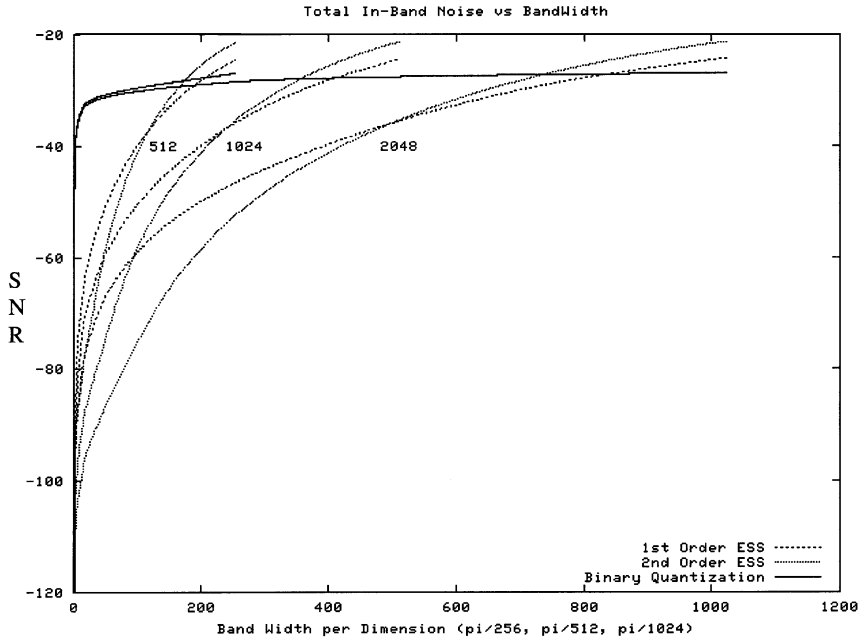


Fig. 8. Total inband quantization noise (dB) versus signal bandwidth, 3-bit lenna 512^2, 1024^2, and 2048^2 images.

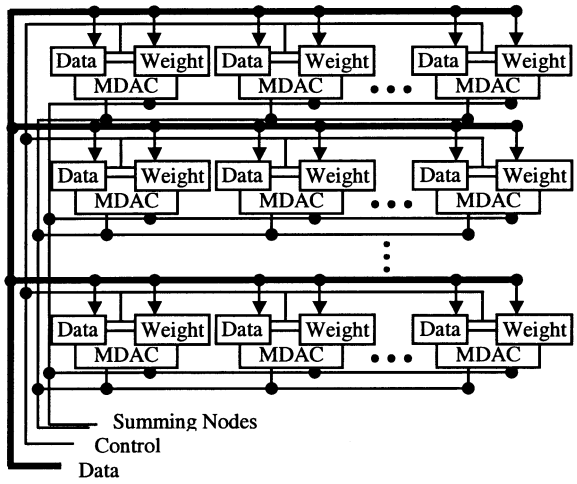


Fig. 9. Convolution circuit architecture.

With the summation portion of a current-mode convolution unlikely to cause significant error, the inital issue defaults to resolving the effect of noise shaping on the performance of the multipliers. To that end, measuring the noise characteristics of a multiplier array is the focus of the fabricated circuits.

B. Circuit Architecture

Fig. 9 illustrates the block diagram of the selected architecture for evaluating ESS performance on analog-computed CMOS VLSI image convolutions. This circuit computes a single convolution point at a time using an array of parallel multiplying analog to digital converters (MDAC). The weight and data values are loaded via a common data bus into registers, and the positive and negative summation components are collected on independent nodes.

While this structure is primarily intended as a vehicle to measure mismatch in the array of multipliers, it could easily be extended to performing real-time convolutions in any spatial dimension by replacing the data bus and data registers with a chain of shift registers such that signal data would be swept past the convolution kernel, with one convolution output sample for each shift of the registers. In an image convolution, a succession of additional shift registers could maintain the pixels' spatial relationship and only retain the portion of the image that would be subsequently used (Fig. 10), effectively spiraling the kernel around the 2-D signal as it is row-wise shifted into the shift register array. For imagers that serialize their data, such a pipelined structure would not interfere with the preexisting data-rate bottleneck and only imposes a time shift on the results.

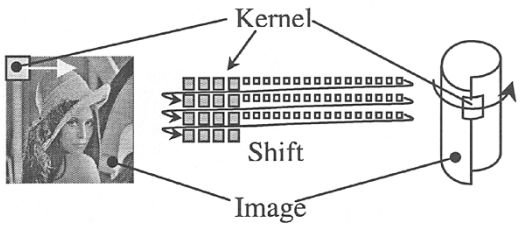


Fig. 10. Pipelined convolution processor.

In practice, the convolution kernel coefficients would be adjusted using ESS to both correct for the individual hardware's variations and the resolution loss from quantization (i.e., the two components of e_w in Fig. 1); the resultant values would then be stored in the weight registers for the duration of a convolution. Since each sample of the image is swept through the kernel and eventually visits each MDAC, no hardware-specific correction can be made to the image values, though such error is still prone to reduction through oversampling. Nonetheless, the image's reduced quantization necessitates the use of ESS to preserve its point-wise resolution.

The weight's sign bit directs the output current of each MDAC onto the appropriate summing node. Splitting the positive and negative components into two nodes simplifies

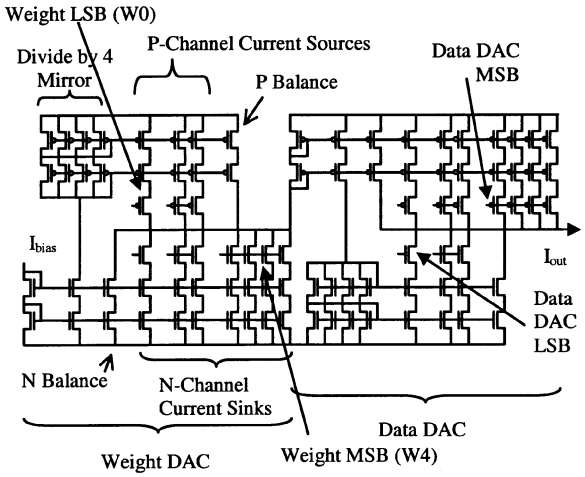


Fig. 11. MDAC schematic.

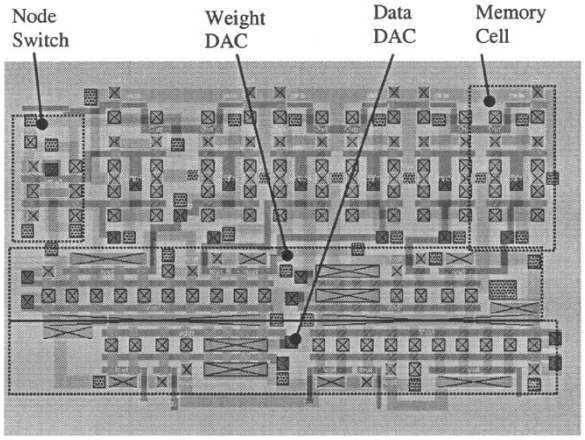


Fig. 12. MDAC layout.

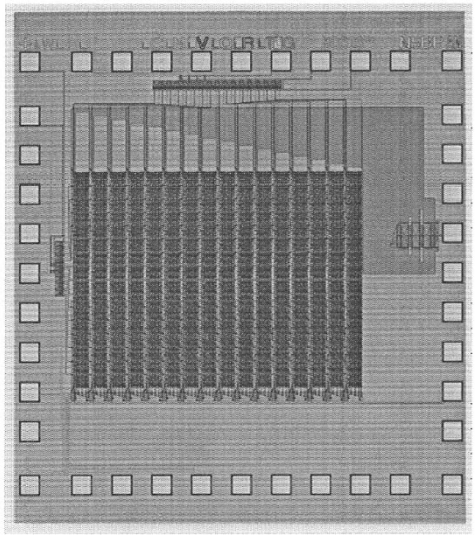


Fig. 13. 16 × 16 MDAC array IC.

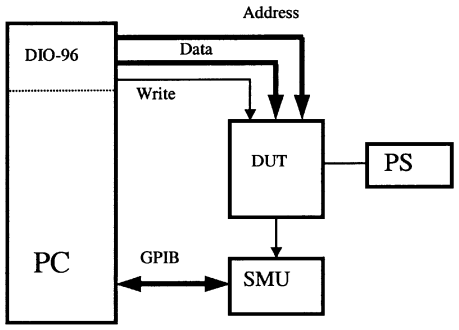


Fig. 14. Automated data collection system.

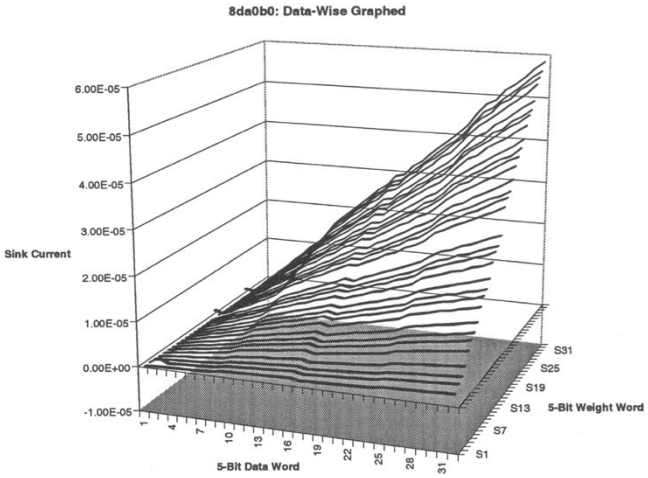


Fig. 15. Example MDAC output sweep.

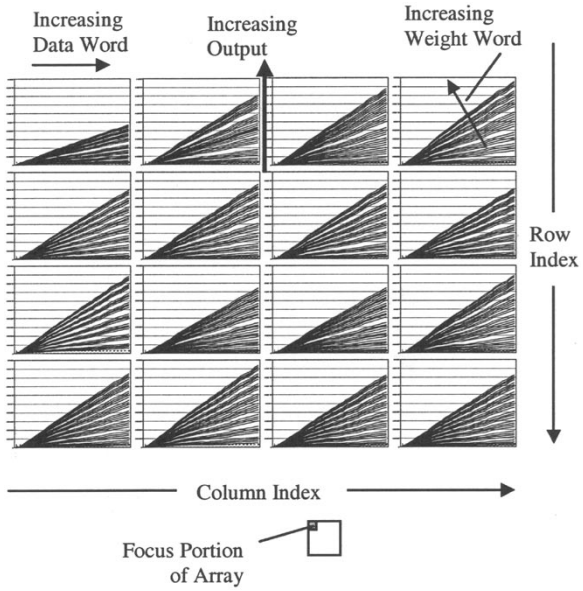


Fig. 16. Example output sweeps from one corner of an MDAC Array IC; **0–200 μA** vertical span subplots. Each line represents a constant weight value.

the multiplier's design, helps to avoid saturation, and provides additional dynamic range for the output. Complex convolutions would require four summing nodes. In addition to representing the two polarity sums, the two output nodes provide a way to isolate the output of a particular MDAC from all of the others during test. Note that there is a precedent for this type of splitting: it is done in virtually all biological vision systems since neurons can only represent magnitude, necessitating parallel channels to independently carry the positive (e.g., "on-center" surrounds [11]) and negative ("off-center") information.

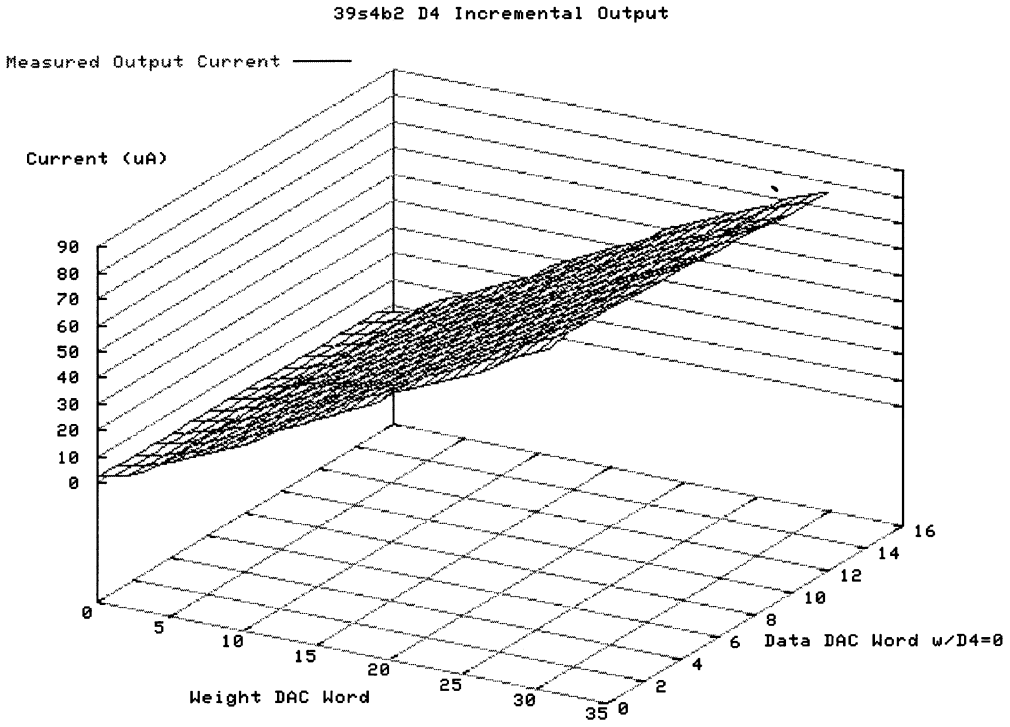


Fig. 17. Incremental I_{out} per bit for data MSB.

The process of consolidating the summing nodes' data is dependent on the needs of the downstream processing steps and represents a degree of design freedom that is not central to this research. A single opamp or current mirror could be used to combine them, if desired. Whatever methodology might be used, it is necessary that the voltage on these nodes remain within the limits allowed by the MDACs output resistance.

C. MDAC Design

The circuits shown in the left and right halves of Fig. 11 are CMOS current-mode 5-b Digital to Analog Converters (DAC); they are CMOS mirror images, but otherwise identical to each other. This DACs primary advantage is that it is very small in size, consisting of only 46 minimum geometry transistors.

Focusing on the weight DAC, the input current, I_{bias}, is cascode mirrored nine times. Triple cascoding was believed to be unnecessary overkill considering the accuracy expected of this circuit in most processes, though it might have utility in conditions where the output node voltage is under-controlled and it would prevent undesirable feedback through the output resistance. Three sets of these current sinks are gated on/off in groupings of four, two, and one, respectively sinking four, two, and one times a current equal to I_{bias} from the output node. Another of the *N*-channel current sinks drives a divide-by-4 *P*-channel mirror.

The remaining *N*-channel current sink works with the *P*-channel current sources to mimic the behavior of finer resolution *N*-channel current sinks. By using the *P*-channel sources to displace current that is already being drawn, additional current mirroring back to *N*-channel devices is avoided along with its attendant mismatching. Each of the *P*-channel sources can provide one quarter of the current drawn by an *N*-channel sink, leaving a balance of zero net current available on the output node when the two least significant bits (W0 and W1) are low. As the *P*-channel current sources turn off, the *N*-channel balancing sink must turn to the output node for an equivalent current: with the LSB (W0) high, $\frac{1}{4}$ I_{bias} net is drawn from the output node, and with W1 high $\frac{1}{2}$ of I_{bias} net is drawn. Like the *N*-channel balancing sink, the *P*-channel balancing current source is always on and is there so that the DAC sinks currents ranging from 0–7$\frac{3}{4}$ times I_{bias} rather than $\frac{1}{4}$-8 times I_{bias}.

By using the output of the Weight DAC as the bias input to the Data DAC, the overall computation performed by the MDAC circuit is

$$I_{out} = I_{bias}\frac{(\mathrm{WEIGHT})(\mathrm{DATA})}{16}$$

where *WEIGHT* is the numerical value of the 5-b binary input to the Weight DAC and *DATA* is the numerical value of the 5-b binary input to the Data DAC.

A representative layout of the MDAC occupies the lower half of Fig. 12. To avoid undue error contributions or inadvertently whitening its inherent error, various efforts were taken during creation of the layout to minimize expected geometrical and structural mismatch: compact design, shared well, central location of the mirror input transistors, etc.

D. MDAC Array IC

This IC contains a 16×16 array of MDACs, each of which has its own set of storage registers for retaining both the Sign + Weight (6 b) and unsigned 5-b Data values; its photomicrograph is shown in Fig. 13. This chip was fabricated in HPs CMOS26 G 0.8 μm N-well process using the MOSIS fabrication service. Both MDAC registers read from a single chip-wide data bus and

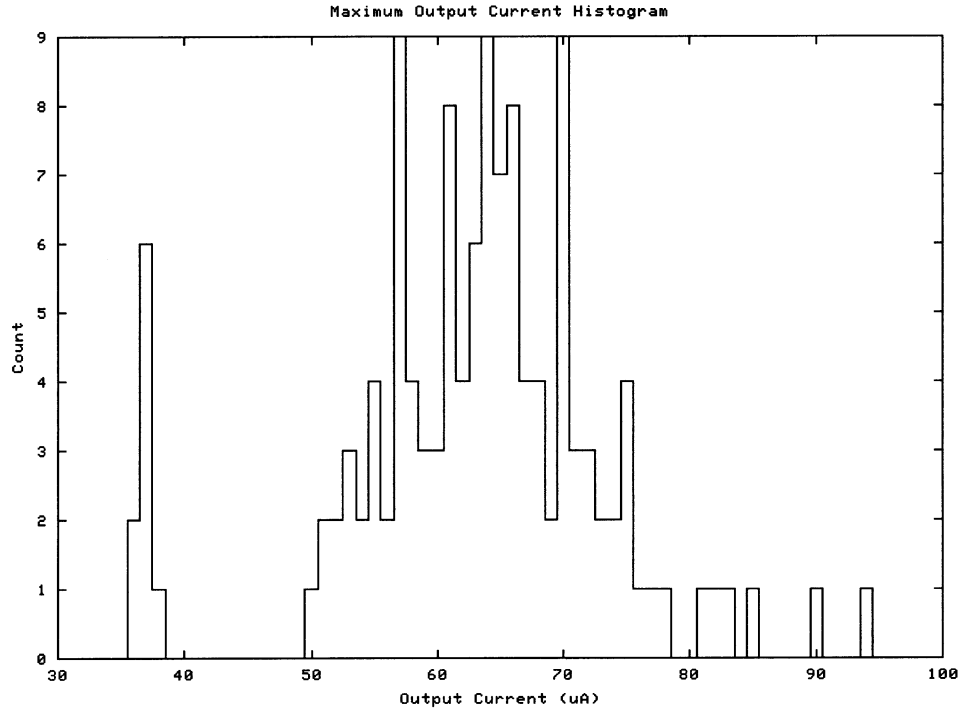


Fig. 18. Histogram of MDAC array maximum outputs.

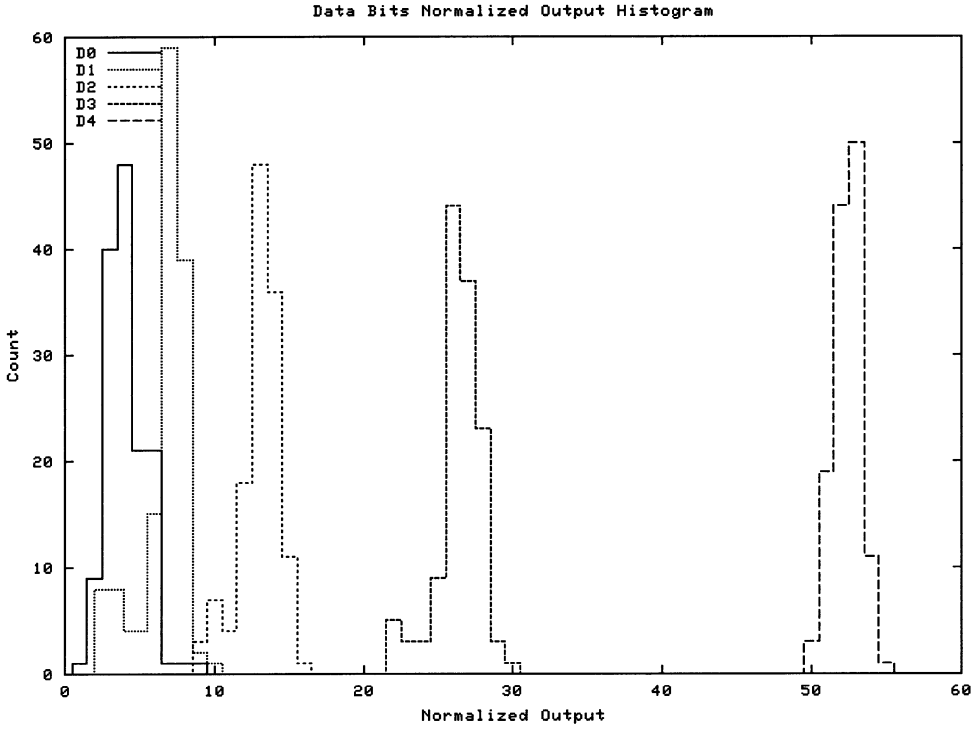


Fig. 19. Histogram of the weight bits (with the data word at maximum and each MDACs maximum output normalized to a value of 100). MSB on right, LSB on left.

the specific register is selected by row, column, and weight/data address lines. In addition to representing the two polarity sums, the two output nodes provide a way to isolate the output of a particular MDAC from all of the others.

Two different mechanisms were used to provide the reference (bias) current for the MDAC array. The 128 neighboring MDACs that occupy row addresses 8 through 15 are biased with a voltage that is distributed via a network of metal interconnect (voltage biased). The advantage of this approach is that it uses a minimum of silicon area to distribute the biasing signal. Its disadvantage is that the currents drawn from this line will produce a voltage drop that will in turn be reflected in the signal available at the input to the biasing circuits.

The 128 MDACs that occupy row addresses 0 through 7 are biased via individual current lines (current biased). While these MDACs are expected to be biased closer together than the

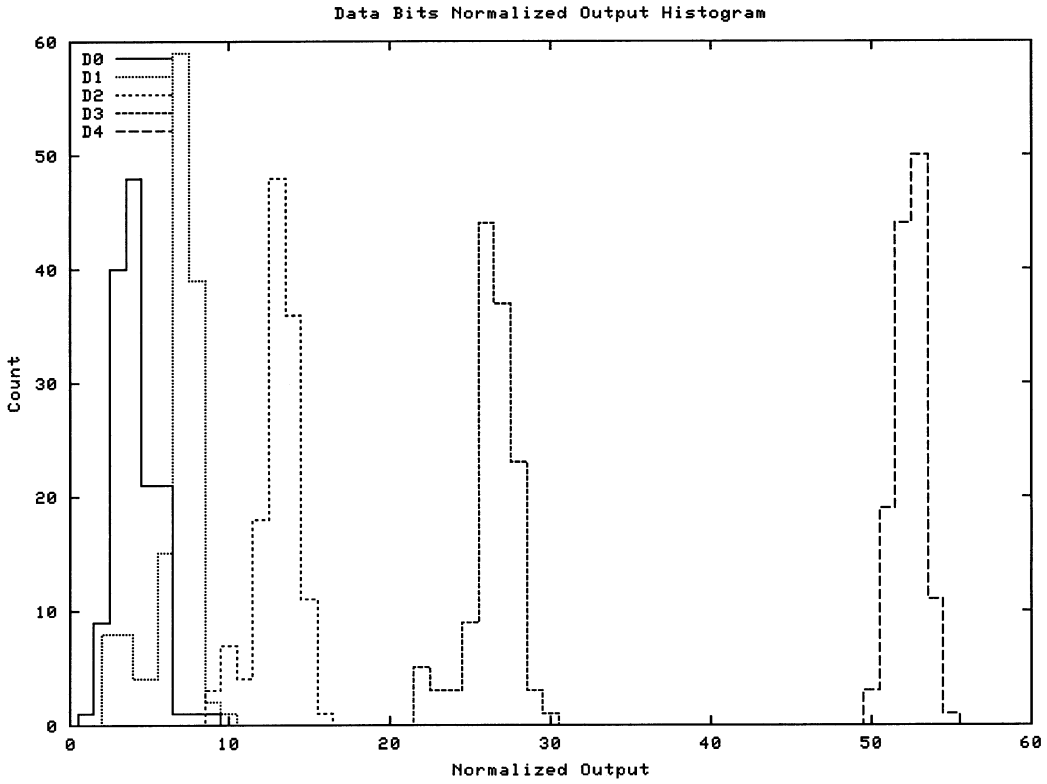


Fig. 20. Histogram of the data bits (with the weight word at maximum and each MDACs maximum output normalized to a value of 100) MSB on right, LSB on left.

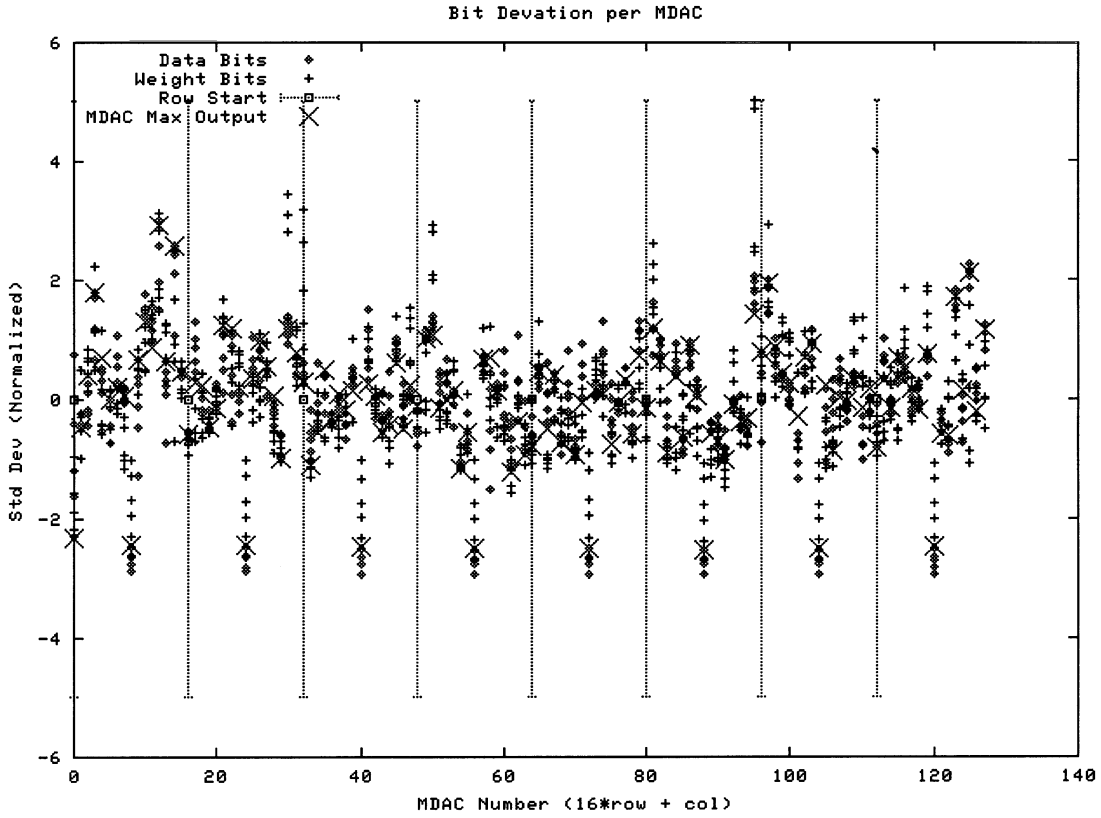


Fig. 21. Normalized per bit and maximum-output distributions plotted against MDAC physical location.

voltage biased group, this method requires substantially more metal interconnect to route the 128 individual current lines and a massive current mirror (visible on the right side of Fig. 14) to create the currents for distribution.

IV. Testing and Measurement

The working data for this research comes from the MDAC Array IC and its 16×16 array of MDACs. In addition to the array of MDACs, it also contains the row and column decoders, data and weight + sign registers for each MDAC, and the two different biasing mechanisms outlined in Section III.D.

Ignoring output directioning, each MDAC is capable of 1024 distinct input configurations. Twenty-five of these MDAC Array IC chips were fabricated, each with 256 MDACs, making a total of roughly 6 $\frac{1}{2}$ million inspection points to measure under several sets of differing biasing conditions each.

A. Automated Data Collection

An automated data collection system was built to measure the MDAC Array ICs performance; its block diagram is in Fig. 14. A National Instruments DIO-96 digital IO card placed in a generic brand 33 MHz 486 PC was used to drive digital signals (Row and Column Addresses, the Data Bus, and Write Pulse) into the IC under test. A Keithly 236 Source Measurement Unit (SMU) was used to collect the analog output data. Fifteen-turn potentiometers set up as voltage dividers were used to supply each of the necessary bias voltages and a HP 6205C Power Supply powered the chip and biasing dividers.

A custom LABView program was developed to control the automated system. After an initial sweep to clear each of the Weight and Data registers of any random power-on contents and to tie each of the MDACs to the same output node ("node B"'), each of the 256 MDACs was singly connected to the other output node ("node A") and swept through all 1024 possible input combinations. The Keithly SMU was set to measure the current sourced from output node A, and the collective results were written to a data file per MDAC.

B. MDAC Response

Fig. 15 is a 3-D graph of a data sweep collected from the MDAC at row 0 and column 0 of MDAC Array IC #8 ("8da0b0"). Note that the output is predominately linear, contains a minor element of irregularity (e.g., mismatch noise), and successfully implements the multiplication function.

Full data sweeps of the entire 16×16 MDAC array of each of the working chips have been recorded. Fig. 16 displays a representative set of sweeps for the MDACs in rows 0 through 3 and columns 0 through 3 of chip #8 with the SMU sinking current while biased to 2.50 V and $\mathrm{I_{bias}}$ tied to 3.90 V. The response is fairly linear in general, and with the exception of MDAC (0,0), which is only operating across roughly half of the typical output span, all of the MDACs are comparably responding. Various pathologies are represented here: the central gap in the spread of the lines in the graphs for MDACs (0,0), (0,2), (1,1), (2,1), and (2,2) indicate that their W4 bit (Weight MSB) is somewhat stronger than normal; similarly, MDAC (0,1) has a stronger than normal W3 bit, as evidenced by the upward displacement of both of its W3-on regions; several of the graphs have weight curves grouped into eight bunches of four lines each, MDAC (2,0) for example, which indicates the presence of some inordinate mismatch in their P-balance mirrors, etc.

Fig. 22. MDAC-based convolution result: $1024^2$1-b first-order ESS image convolved with 64^2 MDAC array-specific first-order ESS DOG.

Fig. 23. Desired goal: result of convolving a floating point 64^2 DOG with 8-b 1024^2 image.

C. Differential Nonlinearity

Referring to Fig. 11, the incremental output in current as a given Data bit turns on should be the same for a fixed Weight word and regardless of how the other data bits are set; any deviation would appear as differential nonlinearity (DNL). Extending the DNL concept to MDACs, as the Weight word increases, this incremental output per Data word bit should also increase proportionally. As plotted in Fig. 17, for D4 (the Data word's MSB) of the MDAC at row 4 column 2 of chip #8 (a random selection), the incremental output current, $\mathrm{I_{out}}$, largely exhibits this behavior.

D. Error Distribution

A histogram of maximum outputs $(\mathrm{Data} = 31, \mathrm{Weight} = 31)$ from the MDACs in the current biased half of an individual IC is plotted in Fig. 18. The central distribution largely follows a gaussian pattern.

The histogram of output values for each weight bit (with the Data word set to the maximum of 31, and the MDACs maximum output normalized to a value of 100,) is plotted in Fig. 19. Again, each of the significant bits cluster in essentially Gaussian-like distributions. A similar histogram for the Data bits is in Fig. 20.

To visually inspect for correlation between these distributions, each of the Weight and Data bit clusters and the maximum output histogram have been normalized to zero mean and unit variance and then superimposed on a per MDAC basis; the result appears in Fig. 21. This reveals that eight of the nine MDACs in the outlying cluster that form the leftmost group in Fig. 18 are all in the same column of the array (with the remaining such defect at MDAC (0,0), as visible in Fig. 16,).

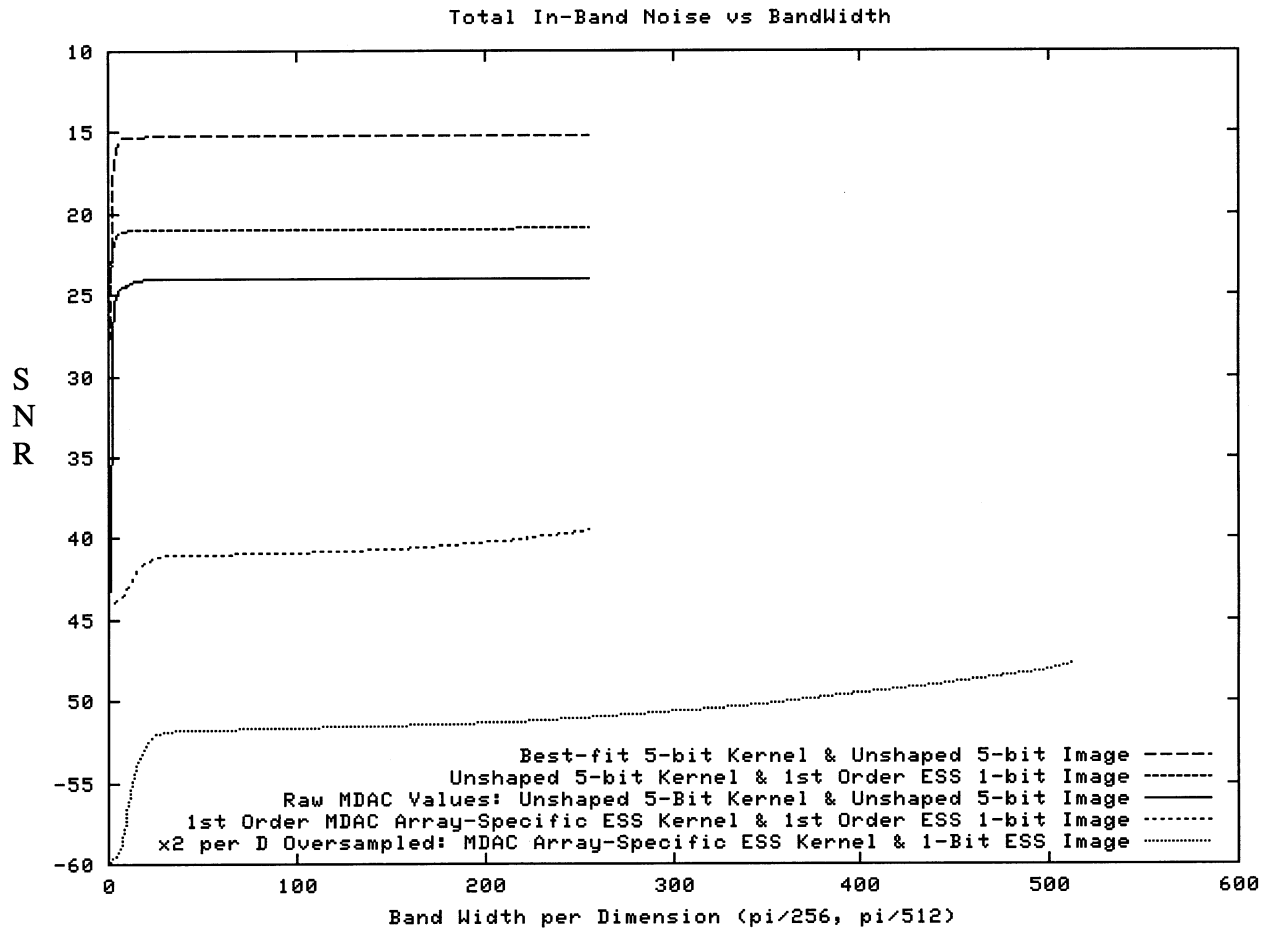


Fig. 24. Total inband noise for selected MDAC-based convolutions, chip #8. Curves from top to bottom (a) Best-fit 5-b kernel and unmodified 5-b image. (b) Unmodified 5-b kernel and 1-b first-order ESS image. (c) Unmodified 5-b kernel and unmodified 5-b image. (d) MDAC array-specific first-order ESS kernel and first-order ESS 1-b image. (e) Oversampled MDAC array-specific first-order ESS kernel and first-order ESS 1-b image. 32^2 DOG and 512^2 image for (a)-(d), 64^2 DOG and 1024^2 image for (e).

V. Evaluation

The measured MDAC data has been used to form the core of a convolution simulator. After loading a single ICs MDAC data into a four-dimensional (4-D) look-up table (addressed by row, column, 5-b weight word, and 5-b data word), each multiplication in a convolution is simulated by retrieving the associated output measurement. With the top and bottom halves of the fabricated arrays biased differently, the largest set of measured MDAC responses with uniform statistics is only 8×16, so to simulate larger kernels, the array data has been repeatedly tiled.

Fig. 22 is the result of convolution simulation using a 64^2 Difference of Gaussians Filter (DOG; as a bandpass filter, it is representative of many linear filters) that has been first-order ESS fit to the recorded MDAC outputs of one IC array and a 1-b first-order ESS image. For comparison, Fig. 23 is the targeted ideal result (floating point kernel and 8-b image).

Fig. 24 compares the inband noise for several types of MDAC-based convolutions: no noise shaping (the baseline, line c: 5-b DOG and 5-b image); only the image noise shaped (line b: 5-b DOG and 1-b first-order ESS image); only the kernel coefficients modified so as to select the MDAC output nearest the ideal value (line a: "best-fit" 5-b kernel and unmodified 5-b image); both the kernel and image noise shaped (line d: MDAC Array-Specific first-order ESS DOG and 1-b first-order ESS image), and both the $2\times$ oversampled per dimension kernel and $2\times$ oversampled per dimension image noise shaped (line e: oversampled MDAC array-specific first-order ESS kernel and first-order ESS 1-b image). A 15-dB improvement over the baseline occurs with noise shaping both signals, and a further 10 dB above that with additional oversampling.

Table I lists the inband SNR for various convolutions performed using the recorded data from a single chip. Identical calculations were performed for four other chips, with similar results [23]; space limitations preclude their inclusion here. In all cases, the baseline (i.e., "signal") image was computed using floating point values for the kernel coefficients and 8-b for the image samples.

With one clear and considerable exception, the benefit from noise shaping the kernel and/or image is generally only minor using this metric under these conditions. The exception is for the MDAC array-specific first-order ESS kernels and 1-b first-order ESS images, which across all chips and filter frequencies shows a substantial improvement that further increases with oversampling; these particular results are highlighted with bold text in the table.

Though it is reasonable to expect that decreases in the image's sample resolution would produce a corresponding decrease in

TABLE I
TOTAL IN-BAND SNR OF CONVOLUTION RESULTS FOR MDAC DATA FROM CHIP #8

Nyquist Rate		Kernel Size		
512^2 Image	Image Bits	64^2	32^2	16^2
Unmodified Kernel & Unmodified Image	5	23.96	24.02	21.95
	4	22.86	24.95	22.16
	3	22.21	23.93	22.58
	2	23.71	16.76	19.89
	1	19.54	20.00	19.87
Best-Fit Kernel & Unmodified Image	5	15.20	13.95	12.48
	4	14.66	13.81	12.51
	3	15.11	14.03	12.92
	2	17.62	16.61	15.98
	1	10.82	10.78	9.67
Unmodified Kernel & 1st Order ESS Image	5	23.95	23.81	21.81
	4	23.63	25.26	22.17
	3	26.03	24.83	23.57
	2	27.57	25.91	27.08
	1	20.94	26.45	21.00
1st Order ESS Kernel & 1st Order ESS Image	5	25.05	24.62	21.30
	4	26.28	25.42	21.51
	3	25.14	23.88	20.85
	2	27.00	25.43	20.00
	1	**40.78**	**34.57**	**25.37**
2nd Order ESS Kernel & 1st Order ESS Image	5	26.17	24.86	20.05
	1	31.91	30.74	26.70
Oversampled		**Kernel Size**		
1024^2 Image	Image Bits	128^2	64^2	32^2
Unmodified Kernel & Unmodified Image	5	24.01	21.91	23.32
	1	19.56	18.97	21.31
1st Order ESS Kernel & 1st Order ESS Image	5	25.31	24.13	23.26
	1	**51.56**	**46.24**	**37.63**
2048^2 Image	Image Bits	256^2	128^2	64^2
1st Order ESS Kernel & 1st Order ESS Image	1	-	**52.39**	**44.6**

the convolution's inband SNR, that is not revealed in these figures, where no general trend in either direction is apparent. Similarly, noise shaping the images showed no appreciable advantage when using an unmodified kernel.

The best-fit kernels, where the coefficients are selected to produce the MDAC weight closest to the ideal output regardless of which particular 5-b address produces it, could also be expected to out-perform the unmodified kernels, but there is no case when they did not instead do worse, often considerably.

In general, first-order ESS of both the kernel and the image produced only minor improvements over the unmodified kernel and unmodified image resultsions). The performance of the 1-b ESS images with the first-order ESS kernel, however, is substantially better than at the other resolutions.

A potential reason for this is that the 1-b images contain a substantial number of "0"valued coefficients, probably somewhat more so than the other resolution images, which do not contribute to the random component of e_d as much as do nonzero values. Detracting from this hypothesis is that there is no improvement of inband SNR as the images' bit resolution decreases (and the number of "0" value coefficients rises), but instead only a substantial jump at the 1-b level.

A more likely explanation for the superior results at the level of 1-b image resolution is that the levels used in the ESS feedback loop do not match those ultimately used in the MDAC array, creating another noise source (just as with multilevel $\Sigma\Delta$ ADCs, which similarly perform poorly compared to their 1-b counterparts).

The second-order ESS kernels do not perform as well as the first-order, though the improvement at 1-b image resolution appears to hold for them as well.

A further distinguishing characteristic of the 1-b first-order ESS images is that additional inband SNR improvement accompanies oversampling, which is not the case for at least the 5-b images. The high-pass nature of the first-order differential ESS is also apparent in these 1-b image results, where the lower the center of the DOGs passband (where the ESS image has better captured its information), the higher the resulting inband SNR.

VI. CONCLUSION

ESS has been demonstrated as a mechanism for reducing inband noise expected from quantization and circuit mismatch in mixed-signal FIR image filters, with the best results occurring when 1-b ESS images are convolved with noisy 5-b weights that are selected via ESS. The ESS algorithms used here were very crude, with the first-order filter using only two coefficients and the second-order filter using only four. Though they have been useful in demonstrating a successful utilization of ESS, their efficiency is very low and the second-order algorithm used here proved to be unstable on occasion. Refinement of the ESS algorithms and complete identification of the individual noise floors should be accomplished before another generation of circuits is attempted. Also, an assumption was made that noise in the current-mode addition would be negligible and this remains to be verified.

REFERENCES

[1] M. Mahowald, "Silicon retina with adaptive photoreceptors," in *Proc. Soc. Optical Int. Engineering (SPIE)* , vol. 1473, 1991, pp. 52–8.

[2] C. Mead, *Analog VLSI and Neural Systems*. New York: Addison-Wesley, 1989.

[3] W. Bair and C. Koch, "Real-Time motion detection using an analog VLSI zero-crossing chip," in *Proc. Soc. Optical Int. Engineering (SPIE)* , vol. 1473, 1991, pp. 59–65.

[4] K. A. Boahen and A. G. Andreou, "A contrast sensitive silicon retina with reciprocal sysapses," *Advances Neural Inform. Processing* , vol. 4, pp. 762–72, 1992.

[5] H. Kobayashi, T. Matsumoto, T. Yagi, and K. Tanaka, "Light-adaptive architectures for regularization vision chips," *Neural Networks*, vol. 8, no. 1, pp. 87–101, 1995.

[6] J. Harris, C. Koch, J. Luo, and J. Wyatt, "Resistive fuses: Analog hardware for detecting discontinuities in early vision," in *Analog VLSI Implementations of Neural Systems*. Norwell, MA: Kluwer, 1989, pp. 27–56.

[7] C. Chong, C. Salama, and K. Smith, "Real-Time edge detection and image segmentation," *Analog Integrated Circuits Signal Processing*, vol. 2, no. 2, pp. 117–30, Apr. 1992.

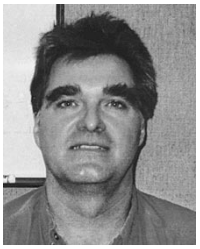
[8] V. Ward and M. Syrzycki, "VLSI implementation of receptive fields with current-mode signal processing for smart vision sensors," *Analog Integrated Circuits Signal Processing*, vol. 7, no. 2, pp. 167–179, Mar. 1995.

[9] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentic-Hall, 1989.

[10] C. D. Nilson, R. B. Darling, and R. B. Pinter, "Shunting neural network photodetector arrays in analog CMOS," *IEEE J Solid-State Circuits*, vol. 29, pp. 1291–1296, Oct. 1994.

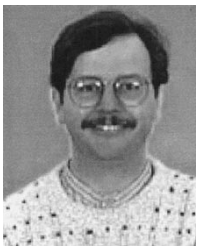
[11] J. E. Dowling, *The Retina: An Approachable Part of the Brain*. Cambridge, MA: Belknap Press, 1987.

[12] T. Allen, C. Mead, F. Faggin, and G. Gribble, "Orientation-Selective VLSI retina," in *Proc. Soc. Optical Int. Engineering (SPIE)*, vol. 1001, 1998, pp. 1040–6.
[13] Y. Nitta, J. Ohata, S. Tai, and K. Kyuma, "Optical neurochip for image processing," *Electron. Commun. Japan, Part 2* , vol. 78, no. 1, pp. 10–20, Jan. 95.
[14] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[15] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[16] J. C. Candy and G. C. Temes, *Oversampling Delta-Sigma Data Converters*. Piscataway, NJ: IEEE Press, 1992.
[17] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed. New York: Oxford Univ. Press, 2002.
[18] P. P. Jones, "Evolution of halftoning technology in the United States patent literature," *J Electronic Imaging*, vol. 3, no. 3, pp. 277–92, 1994.
[19] N. D. Stacy, "Error Spectrum Shaping Quantization," Ph.D. dissertation, Brigham Young Univ., Provo, Utah, 1993.
[20] T. Hinamoto, S. Karino, and N. Kuroda, "Error spectrum shaping in two-dimensional digital filters," *Electron. Commun. Japan*, pt. 3, vol. 78, no. 7, 1995.
[21] P. G. Roetling and R. P. Loce, "Digital halftoning," in *Digital Image Processing Methods*, New York: Marcel Dekker , 1994, pp. 363–413.
[22] H. R. Kang, *Digital Color Halftoning*. Piscataway, NJ: IEEE Press, 1999, pp. 357–95.
[23] B. Buchanan, "A Mixed-Signal CMOS VLSI Image Convolution Circuit using Error Spectrum Shaping," Ph.D. dissertation, Georgia Inst. Technology, Atlanta, GA, 2001.



Brent Buchanan (M'02) received the B.S. degree in applied physics (Hons.), the B.E.E. (Hon.), and the M.S. degrees in electrical engineering, and Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, in 1988, 1991, and 2001, respectively.

He is a Veteran of the United States Marine Corps and is currently an Analog Designer with the Tactical Library Development team of Philips Semiconductors Technology Center Group, Longmont, CO.



Martin Brooke (M'87) received the B.E. degree from The University of Auckland in New Zealand in 1981 (Hons.), and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1984 and 1988 respectively.

He is currently with Duke University, Durham, NC, where he is a Professor. His research focus is in high-speed high-performance signal processing and current projects include learning neural network hardware development; neural network prediction of turbulent flow; focal plane image processing hardware development; 155 Mb/s, 622 Mb/s, and Gb/s digital CMOS transceiver circuits for low-cost fiber optic communication; nonlinear filtering algorithms for telecommunications; nonlinear analog to digital converter design; accurate modeling of high speed circuit parasitics; and statistically relevant device models for accurate prediction of high performance integrated circuit yield. He has been awarded four U.S. patents, and has published articles in a variety of technical journals.