A 16-channel Noise-Shaping Machine Learning Analog-Digital Interface

Fred N. Buhler[1], Adam E. Mendrela[1], Yong Lim[1,2], Jeffrey A. Fredenburg[3] and Michael P. Flynn[1]
[1]University of Michigan, Ann Arbor, MI, [2]Samsung Electronics, Yongin, Korea, [3]Movellus Circuits, Ann Arbor, MI

Abstract

A 16-channel machine learning digitizing interface embeds Inner-Product calculation within a Delta-Sigma Modulator (IPDSM) array canceling quantization noise and noise shaping the multiplicand. The prototype, with 16 independent IPDSM channels occupies a core area of 0.95mm^2 in 65 nm CMOS. Each channel performs up to 100M multiplications/s. The system is demonstrated with a standard machine learning scheme for image recognition. It achieves the same classification accuracy for the MNIST set of hand-written digits as with the same algorithm on floating point DSP.

Introduction

IoT devices are collecting and transmitting an ever-increasing amount of data to monitor health, the environment and manufacturing. For example, IoT devices can send vital data and real time monitoring from a patient to hospital staff but this requires large bandwidth and low power operation. Machine learning can overcome bandwidth and power limitations by decreasing the amount of transmitted data through feature extraction, or classification [1] at the sensor. A challenge is that these need energy intensive and accurate inner-product multiplication of the input signal with a pre-calculated basis vector. Traditionally, this involves the digitization of the sensor signal, communication of large amounts of raw ADC data, and then extensive DSP. A compelling approach is to embed machine learning functions within the digitization process. [2] uses a multiplying SAR ADC, however the multiplying SAR approach is limited to low accuracy of around 7 effective bits due to noise, nonlinearity and mismatch. Our new approach achieves classification accuracy equivalent to floating point DSP by embedding the Inner-Product calculation within a Delta-Sigma Modulator (IPDSM) array. The prototype achieves 1.6G 1 bit multiplications/s over 16 channels. Accurate digitization and an effective oversampled multiplicand resolution of 14 bits enables recognition accuracy similar to floating point DSP.

Machine Learning Noise-Shaping Architecture

Our new architecture embeds inner-product calculation, the key function in machine learning, within a ΔΣ modulator to achieve the effective accuracy of conventional DSP, at a fraction of the power, area and data bandwidth. This is done with little extra front-end power or area overhead. The inner-product, IP, is the vector multiplication of the signal sequence, X, with a basis vector φ:

$$P = \varphi X = \sum_{i=0}^{n-1} \varphi_i X$$

Conventionally, computing an inner-product on sensor data requires an ADC to digitize the sensor signal, followed by a digital multiply-and-accumulate (MAC) function. We combine accurate multiplication and digitization by introducing mixing within the modulator loop as shown the simplified first-order depiction in Fig. 1(a). This mixing is implemented either as a pass through or an inversion of the differential signal path – effectively multiplying by ±1. The digitized inner-product is simply the accumulation of the sequence of digital outputs from the one-bit quantizer.

The new architecture has several unique and fundamental advantages over conventional digitization followed by DSP:

First, a unique advantage is that the IPDSM performs digitization and inner-product calculation with almost no quantization error. This is unlike a conventional ΔΣ modulator which simply shapes quantization noise for later DSP filtering. Simple time domain analysis (Fig. 1(a)) of a first order loop shows that almost all of the quantization error terms ($e_{(1)}$, $e_{(2)}$ etc.) cancel.

Second, the multiplicand in the IPDSM can achieve arbitrary resolution through a sequence of ±1 values and is noise-shaped for very high precision. The quantization noise of the multiplicand falls outside the signal transfer function (STF) of the main modulator and is rejected. The multiplicand is noise shaped to achieve an effective resolution of 14 bits. For comparison, [2] is limited to a 4 bit multiplicand.

Third, instead of the complicated filters and decimators that are required after a conventional ΔΣ modulator, the digital output of the modulator is processed by a simple accumulator. The IPDSM directly generates the 16 bit inner-product reducing the data output rate by four orders of magnitude over that of a traditional ΔΣ modulator.

Finally, this approach of combining digitization and inner-product calculation has several practical advantages. Implementation of ±1 mixing is inherently linear. Also, compared to a conventional ΔΣ modulator, the only extra cost is the control of the mixing sequence.

Implementation

In the prototype, each modulator is implemented as a second order structure with a 1 bit quantizer (Fig. 1(b). A second-order modulator with ring amplifiers [3] achieves the highest resolution and high bandwidth. A single-bit quantizer resolution ensures high accuracy of the feedback DAC without dynamic element matching (DEM). The amplifiers (Fig. 2) are three-stage ring-amplifiers for efficiency and wide output swing. Furthermore, the autozero of the ring amplifier suppresses 1/f noise which would otherwise contaminate the inner-product calculation. The mixing values for each modulator are stored in a dedicated 8K SRAM. The switch network, shown in Fig 3, implements mixing. Accumulation of the ΔΣ outputs, buffering, and time-interleaving of the output as well as synchronization are handled by a digital controller.

Measurements

The prototype (Fig. 4), with 16 independent IPDSMs is implemented in 65nm CMOS and occupies 0.95 mm^2. Each IPDSM performs up to 100M multiplications per second. The system was evaluated with a standard machine learning algorithm that recognizes handwritten digits from the MNIST database [4]. This algorithm requires extensive and accurate inner-product calculation.

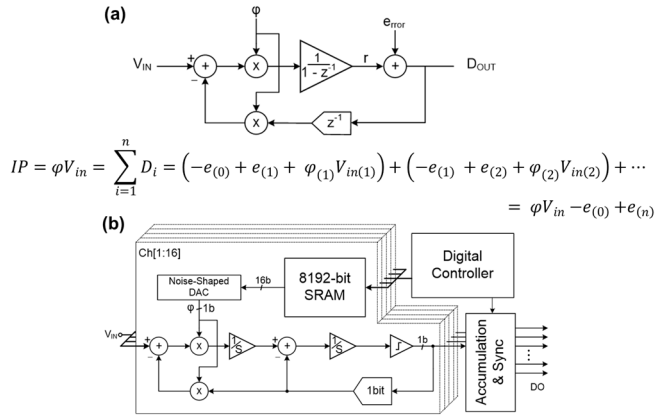
The machine learning process is summarized in Fig 5. In this experiment, 3000 thresholded training images are used to generate a Principal Component Analysis (PCA) basis. The

978-1-5090-0635-9/16/$31.00 ©2016 IEEE

images expressed in the PCA basis are then used to train an Error Correction Output Code (ECOC) Machine Learner with one vs all coding. The IPDSM φ matrix is the product of the machine learner support vectors and the PCA basis.

As shown in Fig 6, IPDSM achieves the same 88% accuracy as floating point DSP on MNIST digit recognition. The plot compares the accuracy of IPDSM, with both fixed point DSP and floating point DSP. With a signal bandwidth of 800k multiplications/sec, the accuracy of IPDSM is equal to that of floating point DSP or 15 bit fixed point.

Fig 7 summarizes the measured performance and compares with the state of the art.



$$IP = \varphi V_{in} = \sum_{i=1}^{n} D_i = \left(-e_{(0)} + e_{(1)} + \varphi_{(1)} V_{in(1)}\right) + \left(-e_{(1)} + e_{(2)} + \varphi_{(2)} V_{in(2)}\right) + \cdots$$
$$= \varphi V_{in} - e_{(0)} + e_{(n)}$$

Fig. 1(a): A simplified first order modulator. 1(b) A block diagram of the IPDSM system.

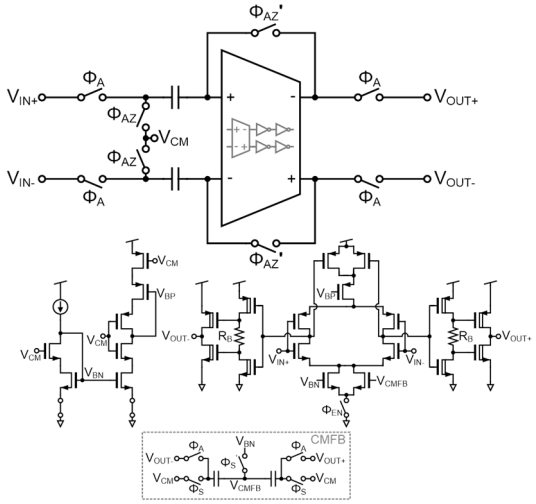


Fig. 2: Ring amplifier schematic with auto-zeroing switch configuration.

Metric	[2]	[5]	This System
Technology	0.13 µm	0.13 µm	65 nm
Num. of Channels	1	64	16
Area (per channel)	0.106 mm^2	0.0938 mm^2	0.0594 mm^2
Signal BW	20 kHz	1 kHz	100 kHz
SNDR	45.8 dB	40.6 dB	69.1 dB
ADC ENOB	7.31	6.5b	11.2b
Sampling Freq.	20 kHz	2 kHz	100 MHz
OSR	-	-	128
FOM _S	151.4	146.6	161.6
Power/ch. (analog)	286.8 nW	28 nW	241 uW
Power/ch. (digital)	376.8 nW		2.7 uW
Product Dynamic Range	40b	8b	16b
Multiplicand Resolution	4b	6b	14b
Application	Feature Extraction	Compressive Sensing	Feature Extraction

Fig 7: Performance summary and comparison with other published works.

Acknowledgments

This work was supported in part by the National Science Foundation and Intel ISRA grant.

References

[1] T. Hastie, et al., The elements of statistical learning
[2] J. Zhang, et al., *ISSCC*, 2015
[3] Y. Lim, et al., *ISSCC*, 2015
[4] LeCun, et al. "The MNIST database of handwritten digits." 1998
[5] D. Gangopadhyay, et al., *JSSC*, F

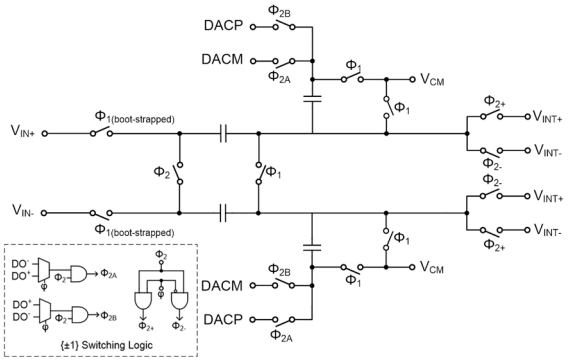


Fig 3: The switch network implementing ±1 multiplication.

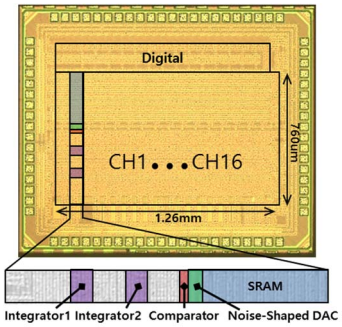


Fig. 4: Die microphotograph.

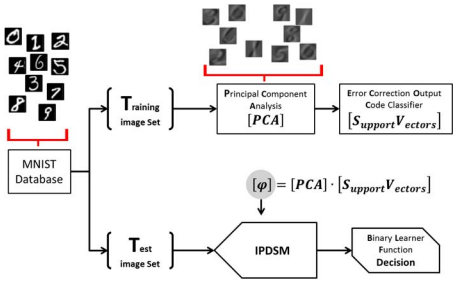


Fig. 5: Machine learning setup. IPDSM calculates the inner product and sends it to the BLF.

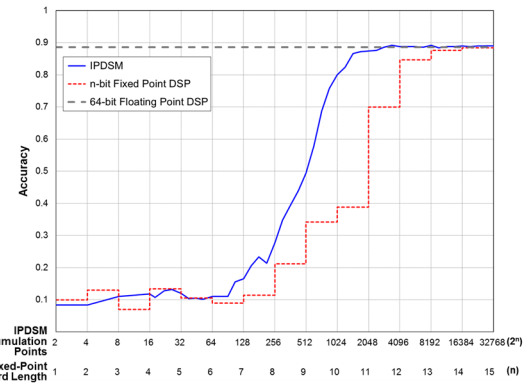


Fig. 6: Machine learning results of IPMDSM versus fixed and floating point DSP of MNIST digit classification.