A Low-Power CMOS Analog Vector Quantizer

Gert Cauwenberghs and Volnei Pedroni

***Abstract*—We present a parallel analog vector quantizer (VQ) in 2.0-μm double-poly CMOS technology and analyze its energetic efficiency. The prototype chip contains an array of 16 $\times$ 16 charge-based distance estimation cells, implementing a 16 analog input, 4-b coded output VQ with a mean absolute difference (MAD) distance metric. The distance cell including dynamic template storage measures 60 $\times$ 78 μm^2. The output code is produced by a 16-cell winner-take-all (WTA) output circuit of linear complexity which selects the winning template with constant power-delay product, independent of input levels and scale. Experimental results demonstrate 34 dB analog input dynamic range and 0.7 mW power dissipation at 3 μs cycle.**

***Index Terms*— CMOS integrated circuits, low-power design, vector quanitzation.**

I. Introduction

VECTOR quantization (VQ) [1] is a common technique for efficient digital coding of analog data, with applications to pattern recognition and data compression in vision, speech, and beyond. The implementation of VQ involves a search among a set of vector templates for the one which best matches the input vector, according to a given distance metric.

Efficient hardware implementation requires a parallel search over the template set and a fast selection and encoding of the "winning" template. Several analog very large scale integration (VLSI) vector quantizers have been developed in recent years, e.g., [2]–[7]. Distinct features of the 16 $\times$ 16 VQ chip presented here include a mean absolute difference (MAD) distance metric and a winner-take-all (WTA) of linear complexity with global positive feedback [6].

II. Architecture

The VQ consists of three main parts in the usual format: a core which computes the distances between the input vector and each of the template vectors; a winner-take-all stage which correspondingly selects the template closest to the input; and a demultiplexing output stage encoding the winning template index.

Additionally, the VQ contains address selection and write enable circuitry to store an input vector into a particular template location. The storage of templates is dynamic in analog format and needs to be refreshed periodically in the present implementation to counteract leakage. Suitable on-chip techniques for local long-term analog storage can be applied and integrated in the present design, e.g., implementing periodic binary quantization [8] and partial incremental refresh [9].

The core of the VQ consists of a 16 $\times$ 16 2-D array of distance estimation cells, configured to interconnect columns and rows according to the vector input components and template outputs. Each cell computes in parallel the absolute difference distance between one component x_j of the input vector $\mathbf{x}$ and the corresponding component y_j^i of one of the template vectors $\mathbf{y}^i$

$$d(x_j, y_j^i) = |x_j - y_j^i|, \qquad i, j = 1 \cdots 16. \tag{1}$$

The MAD distance between input and template vectors is accumulated along rows

$$\hat{d}(\mathbf{x}, \mathbf{y}^i) = \frac{1}{16} \sum_{j=1}^{16} |x_j - y_j^i|, \qquad i = 1 \cdots 16 \tag{2}$$

and presented to the WTA, which selects the single winner

$$k^{\mathrm{WTA}} = \arg\min_i \hat{d}(\mathbf{x}, \mathbf{y}^i). \tag{3}$$

All computations in the VQ processor are performed in parallel, including the distance estimations and the winner-take-all search. It is by now well known that parallel architectures allow energetically more efficient implementation in CMOS for a given computational bandwidth requirement. Indeed, CMOS circuits can be operated in a more energetically efficient mode at lower speeds by operating in the subthreshold MOS regime or by adjusting the supply voltage bias conditions [10].

Fig. 1 shows the micrograph of the VQ chip. The modular structure of each of the four parts (distance estimation, winner-take-all, output encoding, and address decoding) allows the expansion of the design toward larger dimensions without the need to redesign or resize portions of the cells. It is important to note that unlike for conventional (digital) parallel processors, the computational throughput here is additive and scales roughly with the number of cells in each part. Since power dissipation is also additive across cells, the energy per unit computation remains approximately constant, invariant with scale.

III. VLSI Implementation

Both MAD distance and WTA cells operate in clocked synchronous mode using a precharge/evaluate scheme in the voltage domain. The approach followed here offers a wide analog voltage range of inputs and templates at low-power weak-inversion MOS operation and a fast and decisive settling of the winning output using a single communication line for

Manuscript received September 16, 1996; revised January 28, 1997. This work was supported in part by ONR and DARPA through MURI Grant N00014-95-1-0409.

G. Cauwenberghs is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA.

V. Pedroni is with LAC–UFPR/COPEL, Polytechnic Center of UFPR, 80001-970 Curitiba-PR, Brazil.

Publisher Item Identifier S 0018-9200(97)05295-5.

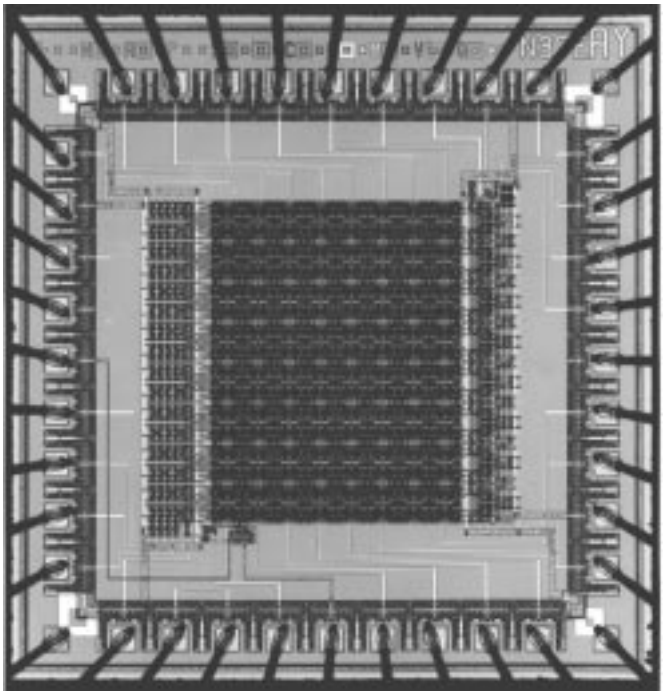

Fig. 1. Chip micrograph. The core contains the array of 16 × 16 distance cells. The parallel WTA section is on the right side of the array.

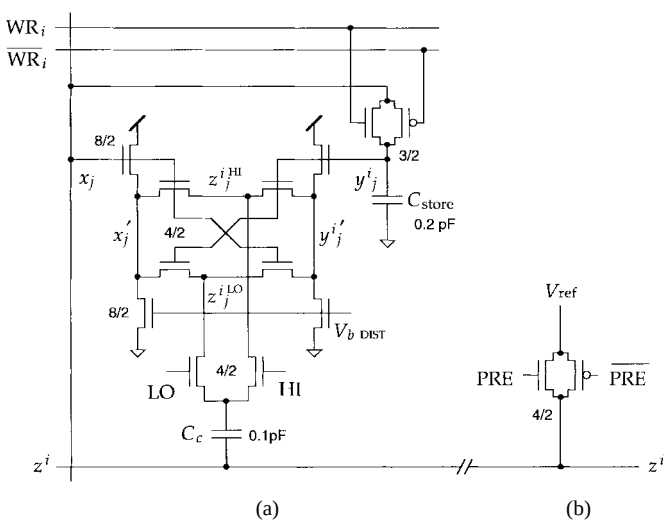

Fig. 2. Schematic of mean absolute difference cell.

global positive feedback. The output encoding and address decoding circuitry are implemented using standard CMOS logic.

A. Distance Estimation Cell

The schematic of the distance estimation cell, replicated along rows and columns of the VQ array, is shown in Fig. 2(a). The cell contains two source followers, which buffer the input voltage x_j and the template voltage y^i_j. The template voltage is stored dynamically onto C_{store}, written or refreshed by activating WR_i while the y^i_j value is presented on the x_j input line. The WR_i and $\overline{\text{WR}_i}$ signal levels along rows of the VQ array are driven by the address decoder, which selects a single template vector $\mathbf{y}^i$ to be written to with data presented at the input $\mathbf{x}$ when WR is active.

Additional lateral transistors connect symmetrically to the source follower outputs x'_j and $y^{i\prime}_j$. By means of resistive division, the lateral transistors construct the maximum and minimum of x'_j and $y^{i\prime}_j$ on $z^{i\,\text{HI}}_j$ and $z^{i\,\text{LO}}_j$, respectively. In particular, when x_j is much larger than y^i_j, the voltage $z^{i\,\text{HI}}_j$ approaches x'_j and the voltage $z^{i\,\text{LO}}_j$ approaches $y^{i\prime}_j$. By symmetry, the complementary argument holds in case x_j is much smaller than y^i_j. Therefore, the differential component of $z^{i\,\text{HI}}_j$ and $z^{i\,\text{LO}}_j$ approximately represents the absolute difference value of x_j and y^i_j

$$\begin{aligned} z^{i\,\text{HI}}_j - z^{i\,\text{LO}}_j &\approx \max\left(x'_j, y^{i\prime}_j\right) - \min\left(x'_j, y^{i\prime}_j\right) \\ &= \left|x'_j - y^{i\prime}_j\right| \approx \kappa\left|x_j - y^i_j\right| \end{aligned} \tag{4}$$

with κ the MOS back gate effect coefficient.

The MAD distances (2) are obtained by accumulating contributions (4) along rows of cells through capacitive coupling, using the well-known technique of correlated double sampling. To this purpose, a coupling capacitor C_c is provided in every cell, coupling its differential output to the corresponding output row line. In the precharge phase, the maximum values $z^{i\,\text{HI}}_j$ are coupled to the output by activating HI, and the output lines are preset to reference voltage V_{ref} by activating PRE, Fig. 2(b). In the evaluate phase, PRE is deactivated, and the minimum values $z^{i\,\text{LO}}_j$ are coupled to the output by activating LO. From (4), the resulting voltage outputs on the floating row lines are given by

$$\begin{aligned} z^i &= V_{\text{ref}} - \frac{1}{16}\sum_{j=1}^{16}\left(z^{i\,\text{HI}}_j - z^{i\,\text{LO}}_j\right) \\ &\approx V_{\text{ref}} - \kappa\frac{1}{16}\sum_{j=1}^{16}\left|x_j - y^i_j\right|. \end{aligned} \tag{5}$$

The last term in (5) corresponds directly to the distance measure $\hat{d}(\mathbf{x}, \mathbf{y}^i)$ in (2). Notice that the negative sign in (5) could be reversed by interchanging clocks HI and LO, if needed. Since the subsequent WTA stage searches for maximum z^i, the inverted distance metric is what is needed for VQ.

Characteristics of the MAD distance estimation (5), measured directly on the VQ array with uniform inputs x_j and templates y^i_j, are shown in Fig. 3. The expanded view in Fig. 3(b) clearly illustrates the effective smoothing of the absolute difference function (4) near the origin, $x_j \approx y^i_j$. The smoothing is caused by the shift in x'_j and $y^{i\prime}_j$ due to the conductance of the lateral coupling transistors connected to the source follower outputs in Fig. 2(a) and extends over a voltage range comparable to the thermal voltage kT/q depending on the relative geometry of the transistors and current bias level of the source followers. The observed width of the flat region in Fig. 3 spans roughly 60 mV and shows little variation for bias current settings below 0.5 μA. Tuning of the bias current allows to balance speed and power dissipation requirements,

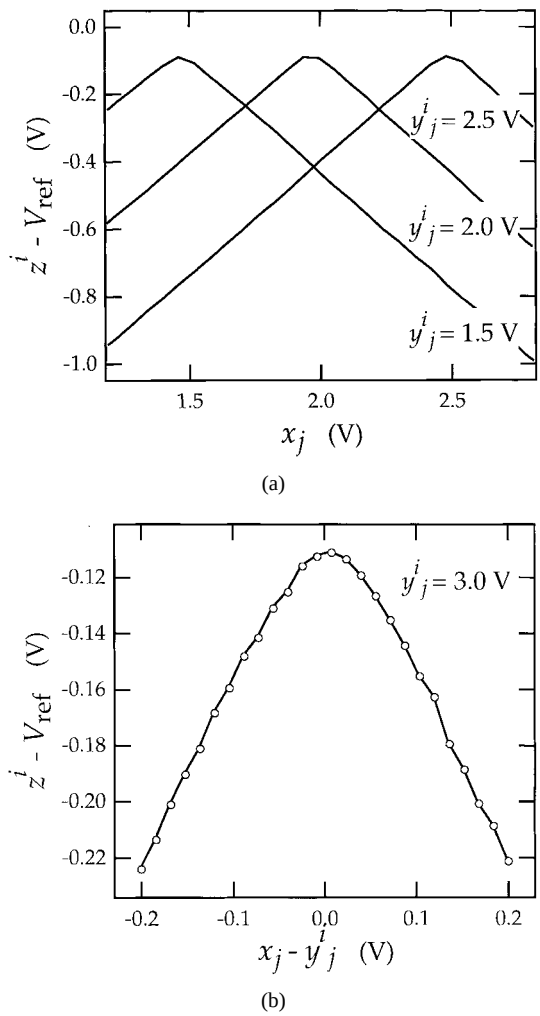

Fig. 3. Measured absolute difference distance characteristics (a) for various values of y^i_j and (b) expanded view.

since the output response is slew-rate limited by the source followers.

B. Winner-Take-All Circuitry

The circuit implementation of the WTA function combines the compact sizing and modularity of a linear architecture as in [3], [11], and [12] with positive feedback for fast and decisive output settling independent of signal levels, as in [4] and [5]. Typical positive feedback structures for WTA operation use a logarithmic tree [5] or a fully interconnected network [4], with implementation complexities of order $O(n \log n)$ and $O(n^2)$, respectively, n being the number of WTA inputs. The present implementation features an $O(n)$ complexity in a linear structure by means of globalized positive feedback, communicated over a single line.

The schematic of the WTA cell, receiving the input z^i and constructing the digital output d_i through global competition communicated over the COMM line, is shown in Fig. 4. The global COMM line is source connected to input transistor Mi and positive feedback transistor Mf and receives a constant bias current $I_{b\mathrm{WTA}}$ from Mb1. Locally, the WTA operation is governed by the dynamics of d'_i on (parasitic) capacitor

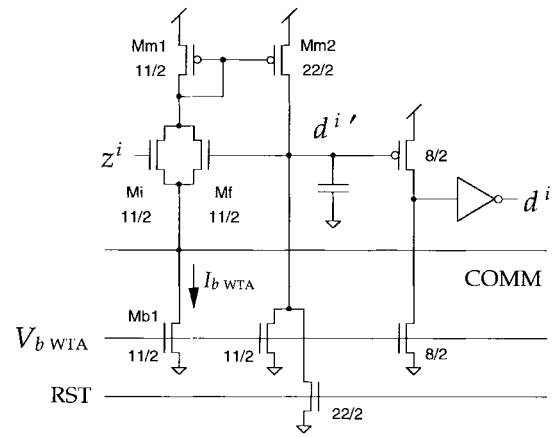

Fig. 4. Circuit schematic of winner-take-all cell.

C_p. A high pulse on RST, resetting d'_i to zero, marks the beginning of the WTA cycle. With Mf initially inactive, the total bias current $nI_{b\mathrm{WTA}}$ through COMM is divided over all competing WTA cells, according to the relative z^i voltage levels, and each cell fraction is locally mirrored by the Mm1–Mm2 pair onto d'_i, charging C_p. The cell with the highest z^i input voltage receives the largest fraction of bias current and charges C_p at the highest rate. The winning output is determined by the first d'_i reaching the threshold to turn on the corresponding Mf feedback transistor, say $i = k$. This threshold voltage is given by the source voltage on COMM, common for all cells. The positive feedback of the state d'_k through Mf, which eventually claims the entire fraction of the bias current, enhances and latches the winning output level d'_k to the positive supply and shuts off the remaining losing outputs d'_i to zero, $i \neq k$. The additional circuitry at the output stage of the cell serves to buffer the binary d'_i value at the d_i output terminal.

In principle, more than one winner could exist at equilibrium. In practice, this is almost never the case, by nature of the combined positive feedback and global renormalization in the WTA competition. Other variants on this WTA circuit can be found in [13].

Fig. 5 illustrates the operation of a 16-element WTA, obtained by Spice simulation using the extracted parameters from the layout. The winning cell receives an input 1 mV higher than the other inputs, which are all identical (2 V). The current $I_{b\mathrm{WTA}}$ is 120 nA.

Test measurements conducted on a separately fabricated 16-element WTA array, identical to the one used on the VQ chip, have confirmed single-winner WTA operation with a response time below 0.5 μs at less than 2 μW power dissipation per cell. The input offset is in the 10–20 mV range and is mostly due to mismatch in transistor Mi. This is not a problem of great concern for the VQ here, owing to the wide voltage range and linear MAD metric of the distance estimates z^i at the input (see discussion below).

IV. Analysis of Energy Efficiency

We define energetic efficiency in the usual context of computation, in terms of the energy required per unit of

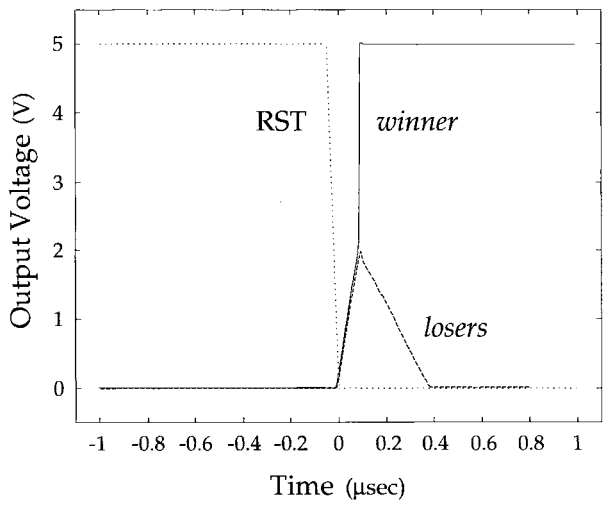

Fig. 5. Simulated response of an array of 16 winner-take-all cells, for a winning cell with input 1 mV higher than all other cells. $I_{b\mathrm{WTA}} = 120$ nA.

computation or, equivalently, the product of power dissipation and cycle time for a given processing element in the circuit which contributes that unit of computation. This is clearly a more meaningful measure of energy efficiency than power dissipation alone, which does not take into account the amount of time required to perform the unit of computation during which the power is dissipated. We will use this measure as a base of comparison between alternatives.

A. Distance Estimation Cell

The most significant contribution of dissipated power in the distance estimation cell is the static power in the source followers, each supplying a constant current $I_{b\mathrm{DIST}}$ set by the $V_{b\mathrm{DIST}}$ bias voltage. This current charges C_c when needed, and thus contributes indirectly to dynamic dissipation as follows. Let t_{cycle} be the cycle time per computation, and let α represent the corresponding fraction devoted to the distance estimation prior to winner-take-all selection. Also, let $V_{\max} < V_{\mathrm{dd}}$ be the maximum allowable voltage swing of the templates y_j^i and inputs x_j. Then the worst case scenario, accounting for current-limited slew, dictates the condition

$$I_{b\mathrm{DIST}} > C_c \frac{V_{\max}}{\alpha t_{\mathrm{cycle}}} \tag{6}$$

which yields an estimate of the lower bound on the energy dissipated per distance estimation cell per computation cycle

$$E_{\mathrm{DIST}} \approx 2 I_{b\mathrm{DIST}} V_{\mathrm{dd}} t_{\mathrm{cycle}} > \frac{2}{\alpha} C_c V_{\max} V_{\mathrm{dd}}. \tag{7}$$

This condition is consistent with the general form of dynamic dissipation in MOS-switched capacitive networks.

B. Winner-Take-All Cell

Similarly, the power dissipation in the WTA cell is dominated by the static current sources $I_{b\mathrm{WTA}}$ charging the output capacitor in Fig. 4. Similar conditions on slew-rate limited fall times as in (6) apply, again placing a lower bound on the dissipated energy E_{WTA} per cell and per cycle. Taking into account the mirror ratios depicted in Fig. 4, the estimate of dissipated energy per cell per cycle is obtained as

$$E_{\mathrm{WTA}} \approx 2.72 I_{b\mathrm{WTA}} V_{\mathrm{dd}} t_{\mathrm{cycle}} > \frac{2.72}{1-\alpha} C_{\mathrm{WTA}} V_{\mathrm{dd}}^2 \tag{8}$$

where C_{WTA} denotes the (parasitic) capacitance of the WTA output node d_i'.

It is important to note that, unlike some other WTA implementations, the time $(1-\alpha)t_{\mathrm{cycle}}$ required for settling does not depend on the number of inputs, nor on the relative input levels, and approximately equals $C_{\mathrm{WTA}} V_{\mathrm{dd}} / I_{b\mathrm{WTA}}$. Implementations without positive feedback suffer from metastability problems which cannot be resolved by increasing the gain of the input stage or by increasing the current $I_{b\mathrm{WTA}}$.

C. Estimates and Comparisons

The relative contribution of energy consumption in the winner-take-all stage becomes smaller as the number of input vector components increases. We consider the limit of large input vector dimensions and neglect the WTA contribution altogether.

In order to minimize E_{DIST} in (7), $V_{\max}$ and C_c cannot be chosen arbitrarily small as to compromise dynamic range and signal-to-noise requirements. $V_{\max}$ directly determines the largest resolvable signal, and the smallest resolvable signal $V_{\min}$ depends inversely on C_c through effects of thermal noise and switch injection noise. With $V_{\mathrm{dd}} = 5$ V, $V_{\max} = 4$ V, $C_c = 0.1$ pF, and $\alpha \approx 0.5$, the estimated energy dissipation per cell per cycle is 8 pJ. This is at least a factor 1000 lower than the equivalent dissipated energy per unit computation in a typical advanced low-power microprocessor or DSP.

From the other samples of VQ implementations discussed before and listed in the references, only the VQ processor in [4] has a superior energy efficiency, owing to the unusually compact charge-domain circuit design implementing the distance operations. While more efficient, this design does not allow straightforward programming and refresh of the templates. The design presented here combines the advantageous energetic efficiency of charge-domain computation with the convenience of standard write and refresh interfacing of the stored templates.

V. System-Level Performance

To characterize the performance of the entire VQ system under typical real-time conditions, the chip was presented a periodic sequence of 16 distinct input vectors $\mathbf{x}(i)$, stored and refreshed dynamically in the 16 template locations $\mathbf{y}^i$ by circularly incrementing the template address and activating WR at the beginning of every cycle. The test vectors represent a single triangular pattern rotated over the 16 component indexes with single index increments in sequence. The fundamental component $x_0(i)$ is illustrated on the top trace of the scope plot in Fig. 6. The other components are uniformly displaced in time over one period by a number of cycles equal to the index $x_j(i) = x_0(i - j \bmod 16)$. Fig. 6 also displays the VQ output waveforms in response to the triangular input sequence, with the desired parabolic profile for the analog distance output

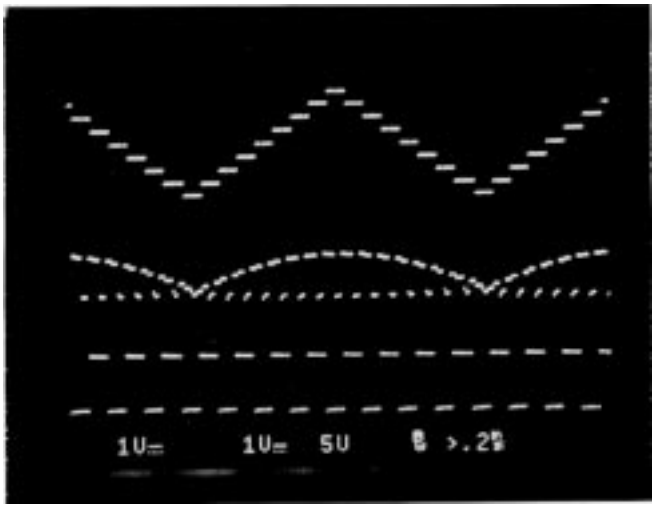

Fig. 6. Scope plot of VQ waveforms. Top: analog input x_0. Center: analog distance output z^0. Bottom: least significant bit of encoded output.

z^0 and the expected alternating bit pattern of the WTA least significant output bit.

The triangle test performed correctly at speeds limited by the instrumentation equipment, and the dissipated power on the chip measures 0.7 mW at 3-μs cycle time (including template write operations) and 5-V supply voltage. With 16 $\times$ 16 VQ distance cells, this translates to a dissipated energy of 10 pJ per elementary computation (component-wise distance estimate and accumulate). The measured energy consumption per cell confirms the prediction of energy efficiency in the above analysis.

An experimental measure for the dynamic range of analog input and template voltages was obtained directly by observing the smallest and largest absolute voltage difference still resolved correctly by the VQ output uniformly over all components. By tuning the voltage range of the triangular test vectors, the recorded minimum and maximum voltage amplitudes for 5 V supply voltage are $V_{\min} = 87.5$ mV and $V_{\max} = 4$ V, respectively. The estimated analog dynamic range $V_{\max}/V_{\min}$ is thus 45.7, or roughly 34 dB. The dynamic range is limited by transistor mismatches in the implementation, but also by the smoothing of the MAD measure characteristic (1) near the origin in Fig. 3(b), implying a "dead zone" of roughly $\pm$20 mV comparable to typical MOS V_T mismatch.

We note that a similar limitation of dynamic range applies to other distance metrics with vanishing sensitivity near the zero crossover point as well, the popular mean square error (MSE) formulation in particular. The MSE metric is frequently adopted in VQ implementations using strong inversion, squaring MOS circuitry [3], [5]. Due to the relatively wide flat region of the MSE distance function near the origin, it fails to discriminate subtle differences between templates which are more pronounced with a MAD metric. This may explain why VQ implementations with MSE formulation are relatively sensitive to template entries in the codebook which are concentrated around zero, as observed for instance in [5] even though the circuits used there are insensitive to MOS mismatches. The implementation presented here has no provisions to compensate for mismatches, since their effect is comparable to that of smoothing in the MAD distance metric. For VQ applications requiring a large number of input components or code vectors, this is still adequate provided the maximum SNR at reconstruction as limited by codebook quantization is less than the 34 dB caused by imperfections in the implementation, which is typically the case for VQ at high ratio of data compression.

TABLE I
FEATURES OF THE VQ CHIP

Technology	2 μm p-well double-poly CMOS
Supply voltage	+ 5 V
Power dissipation	
VQ chip	0.7 mW (3 μsec cycle time)
Dynamic range	
inputs, templates	34 dB
Area	
VQ chip	2.2 mm X 2.25 mm
distance cell	60 μm X 78 μm
WTA cell	76 μm X 80 μm

VI. CONCLUSION

We implemented a low-power VLSI prototype of an analog vector quantizer in 2-μm CMOS technology. A MAD metric is implemented for the distance computations using simple charge-based circuitry. By virtue of the MAD metric, a fairly large (34 dB) analog dynamic range of inputs and templates has been obtained. Likewise, fast and unambiguous settling of the WTA outputs, using global competition communicated over a single wire, has been obtained by adopting a compact linear circuit structure to implement the positive feedback WTA function.

The implemented architecture is fully modular and can be directly extended toward larger dimensions of the vector field and the template set by extending the array of cells. The response time and consumed energy per computation are approximately invariant to scale under resizing of the array dimensions.

The energy dissipation of the chip was analyzed and verified experimentally. The VQ is found to dissipate 10 pJ per cell per cycle, corresponding to one unit of distance computation (per input component per template) in the VQ computation. A summary of the chip features of the 16 $\times$ 16 vector quantizer is presented in Table I.

REFERENCES

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.
[2] B. Hochet, V. Peiris, S. Abdot, and M. J. Declercq, "Implementation of a learning Kohonen neuron based on a new multilevel storage technique," *IEEE J. Solid-State Circuits*, vol. 26, pp. 262–267, 1991.
[3] W. C. Fang, B. J. Sheu, O. T. C. Chen, and J. Choi, "A VLSI neural processor for image data-compression using self-organization networks," *IEEE Trans. Neural Networks*, vol. 3, no. 3, pp. 506–518, 1992.

[4] Y. He and U. Cilingiroglu, "A charge-based on-chip adaptation Kohonen neural network," *IEEE Trans. Neural Networks*, vol. 4, no. 3, pp. 462–469, 1993.

[5] G. T. Tuttle, S. Fallahi, and A. A. Abidi, "An 8b CMOS vector A/D converter," in *ISSCC Tech. Dig.*, 1993, vol. 36, pp. 38–39.

[6] G. Cauwenberghs and V. Pedroni, "A charge-based CMOS parallel analog vector quantizer," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, vol. 7, pp. 779–786, 1995.

[7] M. Konda, T. Shibata, and T. Ohmi, "Neuron-MOS correlator based on Manhattan distance computation for event recognition hardware," in *Dig. Int. Symp. Circuits and Syst.*, Atlanta, GA, 1996, vol. IV, pp. 217–220.

[8] G. Cauwenberghs, "A micropower CMOS algorithmic A/D/A converter," *IEEE Trans. Circuits Syst. I*, vol. 42, no. 11, pp. 913–919, 1995.

[9] G. Cauwenberghs and A. Yariv, "Fault-tolerant dynamic multi-level storage in analog VLSI," *IEEE Trans. Circuits Syst. II*, vol. 41, no. 12, pp. 827–829, 1994.

[10] E. Vittoz, "Low-power design: Ways to approach the limits," in *1994 IEEE ISSCC Proc.*, 1994, vol. 37, pp. 14–18.

[11] J. Lazzaro, S. Ryckebusch, M. A. Mahowald, and C. A. Mead, "Winner-take-all networks of $O(n)$ complexity," in *Advances in Neural Information Processing Systems*. San Mateo, CA: Morgan Kaufman, vol. 1, pp. 703–711, 1989.

[12] A. G. Andreou, K. A. Boahen, P. O. Pouliquen, A. Pavasovic, R. E. Jenkins, and K. Strohbehn, "Current-mode subthreshold MOS circuits for analog VLSI neural systems," *IEEE Trans. Neural Networks*, vol. 2, no. 2, pp. 205–213, 1991.

[13] V. Pedroni, "Inhibitory mechanism analysis of complexity $O(n)$ MOS winner-take-all networks," *IEEE Trans. Circuit Syst. I*, vol. 42, no. 3, pp. 172–175, 1995.