

MIXED ANALOG-DIGITAL IMAGE PROCESSING CIRCUIT BASED ON HAMMING ARTIFICIAL NEURAL NETWORK ARCHITECTURE

Stéphane Badel, Alexandre Schmid and Yusuf Leblebici

Swiss Federal Institute of Technology, Microelectronic Systems Laboratory
CH – 1015 Lausanne, Switzerland

ABSTRACT

A versatile integrated circuit based on the Hamming artificial neural network (ANN) architecture is presented. The circuit operation relies on capacitive processing of sum-of-products terms, complemented with digital post-processing allowing various complex functions to be processed on chip, with a minimal hardware overhead. VLSI realization and measurements are discussed.

1. INTRODUCTION

Over the past decades, artificial neural network (ANN) algorithms have proven to be efficient tools to solve some computationally intensive signal processing problems. The complexity of artificial neural network algorithms has opened the door to their hardware implementation with the main target of accelerating all computations involved.

The hardware Hamming ANN implementation proposed in this paper addresses the issues related to severe limitations in terms of circuit resolution, temporary memory size and their access schemes, realization of high interconnect density by the implementation of analog atomic elements, each processing some specific neural functions, and to be repeated into a regular structure forming a high-performance processing unit. Digital post processing is used in a hardware-friendly realization in order to enhance the functionality of the proposed device.

The Hamming ANN discussed in this paper is based on a compact implementation this approach, with the goal of constructing a building block for high-speed and low-power image processing applications.

2. A HAMMING ARTIFICIAL NEURAL NETWORK

A Hamming ANN is a two layer feed-forward neural network which has the ability of classifying input patterns, based on the criterion of the Hamming distance between previously stored patterns, and the actual input vectors [1]. The first layer – called the quantifier network - is composed of a number of neurons, which perform the Hamming distance computation. The second layer – called

the discrimination layer – is traditionally composed of a feed-forward network, which performs the winner-take-all (WTA) operation, i.e. selects the first-layer neuron of smallest Hamming distance as the winner. Pattern classification applications based on the charge-based operation of the Hamming VLSI circuit have been previously demonstrated [2]-[3], where the quantifier networks is composed of capacitive-threshold logic (CTL) gates [4], and the discrimination network consists of an n-input version of a sense-amplifier, all processing being thus performed in the analog domain.

The quantification network consisting of CTL gates is depicted in Figure 1, where a number of capacitances make the bridge between a common node called row and the set of digital inputs and pre-charge circuitry. A perturbation column is added to the regular CTL gate as an analog input to modify the row voltage state [3]. The comparator and buffer stages build the decision network, which provides the outside world with a binary decision.

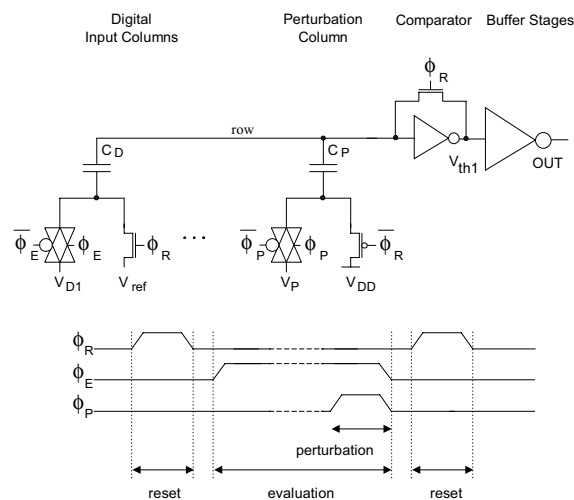


Fig. 1: The adapted CTL gate.

The circuit operation is based on a three cycles scheme, throughout which charge conservation applies to the row node. The quantification operation is realized with a very simple two phase non-overlapping clock scheme consisting of a pre-charge (Φ_R) and evaluation phase (Φ_E). All nodes are imposed a voltage during the pre-

charge phase. The row voltage is set as the comparator circuit threshold voltage V_{thl} while the capacitances other node voltages are imposed as reference voltages set to V_{DD} or GND. The amount of charge transferred to the row is given in Equation 1. All nodes are set back into high impedance after completion of the pre-charge phase. The subsequent evaluation phase starts then, throughout which several vectors may be applied without performing any extra reset. The charge on the row node is considered as constant, the time constant of the leakage parasitic process being significantly larger than that of the system operation. However, the charge on the row capacitors nodes is affected as given in Equation 2. Assuming the equality of these two charges, the row voltage is forced to vary to ΔV_{row} as given in Equation 3. Eventually, the comparator circuit restores a binary voltage depending on the sign of the row voltage variation as given in Equation 4.

$$Q_{reset} = (V_{thl} - V_{ref})C_D^{TOT} + (V_{thl} - V_{DD})C_P \quad (1)$$

with $C_D^{TOT} = \sum_n C_{D_n}$

$$Q_{eval} = \sum_n (V_{row} - V_{D_n})C_{D_n} + (V_{row} - V_P)C_P \quad (2)$$

$$\Delta V_{row} = \frac{\sum_n (V_{D_n} - V_{ref})C_D^{TOT} + (V_P - V_{DD})C_P}{C_D^{TOT} + C_P} \quad (3)$$

$$\begin{cases} \Delta V_{row} < 0 \Rightarrow V_{OUT} = V_{DD} \\ \Delta V_{row} > 0 \Rightarrow V_{OUT} = GND \\ \Delta V_{row} = 0 \Rightarrow \text{limit of the circuit precision} \end{cases} \quad (4)$$

A detailed overview of the perturbation process is given in [5] to which the reader is referred to for further details.

3. OPERATION MODES

The proposed circuit architecture is capable of implementing a number of distinct operation modes, which allows the synthesis of multiple functions. Each mode corresponds to a set of circuit operation characteristics that depend on the targeted application, and dictates a particular hardware configuration.

The circuit can be operated to detect relative or absolute Hamming distances. The neuron configurations for these two modes are explained in detail in [6].

In order to operate on relative distances, only the capacitances representing a stored ‘‘Logic 1’’ value (to be used for comparison) are integrated. Thus, the operation actually implemented is an AND function between the input and the pattern stored in the capacitances. Note that the operation of this kind of neuron results in a distance computation that is not absolute, i.e. it must be compared

to the respective results of other neuron computations in order to produce meaningful information. Using a unit capacitance for each input, while processing a digital Hamming distance operation prior to the capacitive stage, allows the computation of the absolute Hamming distance. The logic operation to be performed is an XOR function between the input and stored data. The row voltage after evaluation of this kind of neuron is proportional to the Hamming distance, thus it is possible to use it either way, with or without any WTA unit.

The second proposed working mode is related to the kind of perturbation signal that is applied to the perturbation columns, and hence affect the row voltages. In this paper we consider the perturbation signal to be equal for each neuron, which is not restrictive and can be easily modified for further applications. Both a pulse and a ramp perturbation signal prove to be interesting candidates, targeting very different potential applications. Their respective effects on row voltage and/or output voltage can be seen on Figure 2 as SPICE simulations, and Figure 8-b as circuit measurements.

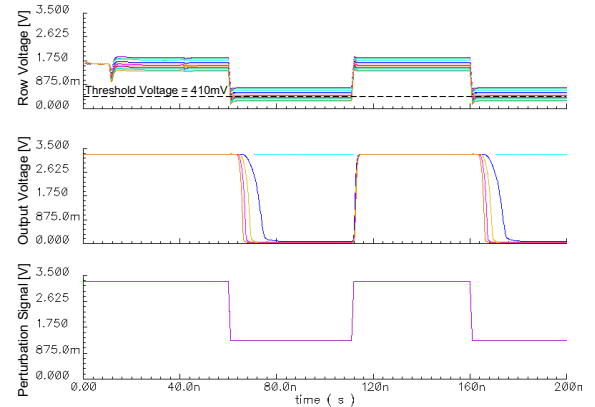


Fig. 2: Pulse perturbation of the neuron row voltage.

4. DIGITAL POST-PROCESSING AND APPLICATIONS

A number of functions can be addressed using the proposed circuit techniques. Post-processing can be added in order to enhance the possible functionality. In the following we restrict the range of applications to digital post-processing based systems.

Ranked order distance measurement based on the Hamming distance computation is depicted in Figure 3.

The distance computation is performed as described in the previous Section (3) by a Hamming ANN core. The output information triggers a pulse generator, which in turn allows a counter to switch into its next state. The rank is given as the value of the counter at the time the Hamming ANN output of a neuron switches. This

information is stored into an array latches, shown in Figure 3.

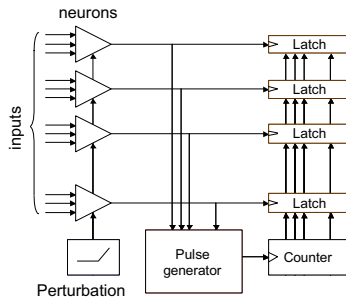


Fig. 3: Ranked order distance computation.

True Hamming distance can be extracted as the latched value of a counter sampled at the switching time of a neuron as depicted in Figure 4.

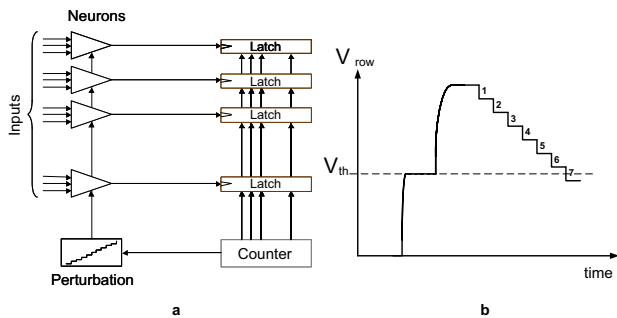


Fig. 4: True Hamming distance extraction. a) System description and b) example row voltage in the case of a Hamming distance equal to seven.

The perturbation signal is chosen as a pseudo-ramp signal, which is made of calibrated steps. Each of these steps is related to the unit voltage quantum required to discriminate to neurons with unit distance. The synchronization of the counter with the perturbation generation circuit guarantees the extraction of the correct value.

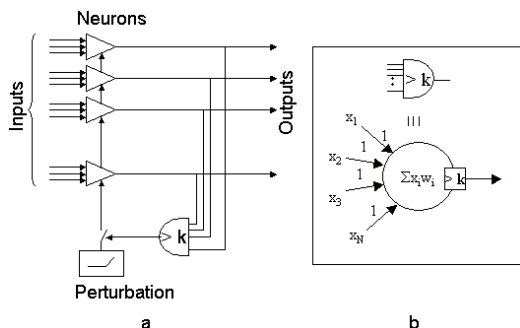


Fig. 5: a) Circuit-level k-winner-take-all description, and b) majority gate.

K-winner-take-all and k-loser-take-all circuits can be synthesized using a calibrated pulse system as described previously. A ramp perturbation signal can also be applied, in which case one additional majority gate has to

be added as post-processing element. A CTL gate is proposed in Figure 5, where the output of the gate controls the end of the process. The winner neurons can be read out of the Hamming core output lines directly.

5. VLSI REALIZATION

An integrated circuit was realized in CMOS 0.35 μm double-polysilicon technology, including a complete 16x16 array Hamming ANN core. The circuit layout of one individual cell is depicted in Figure 6. Two CMOS memory bits have been included into each cell in order to increase the versatility of the circuit. Each of these memory points can be replaced by the application specific connection to the desired sensor input; as such the Hamming core can be modified to accommodate on-chip light sensors in order to construct an intelligent CMOS sensor with early processing of Hamming-based algorithms in the analog domain.

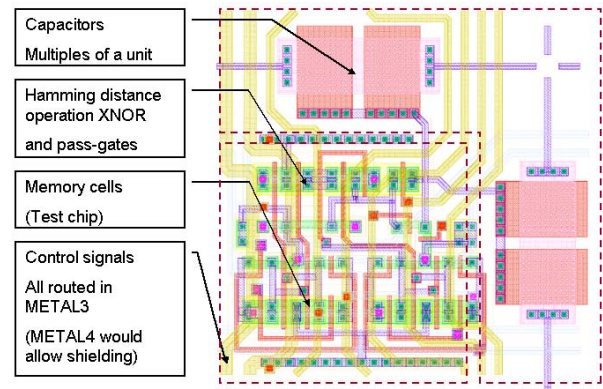


Fig. 6: Layout of one individual ANN cell.

The complete core consists of an array arrangement of the individual cells as depicted on Figure 7. The calibrated perturbation capacitors have been placed on two sides of the core, whereas the offset cells and the output driving circuitry has been placed on two opposed sides of the core, thus resulting in a $(600 \times 600) \mu\text{m}^2$ circuit. The number of signals routed to I/O pads is 80, allowing easy direct and parallel access to most of the signals to be tested. The minimal number of pads allowing parallel access is only five more than the number of data I/Os.

The circuit was measured using a precision ramp generator and a high-speed oscilloscope. Figure 8-a shows the ramp perturbation test applied to sixteen neurons. Each neuron has a different Hamming distance to the stored pattern, ranging from zero to fifteen. Full operability in this mode is achieved, which is witnessed by the successive switching of all neurons.

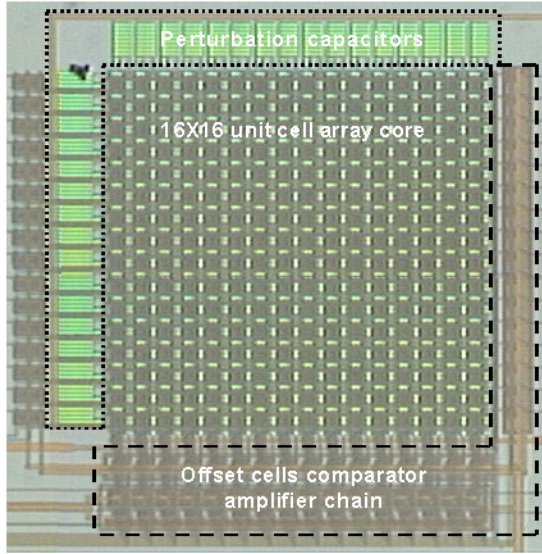


Fig. 7: Microphotograph of the complete Hamming ANN.

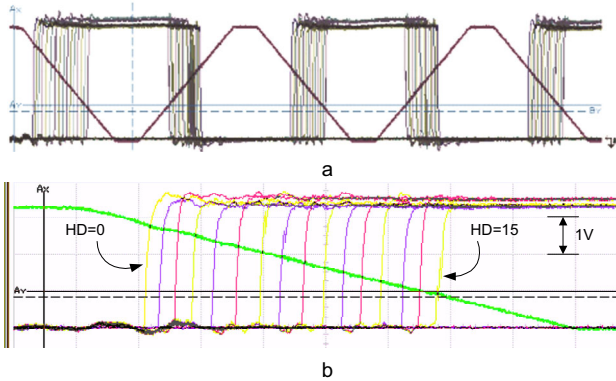


Fig. 8: Oscilloscope waveforms of the ramp test using persistence mode – three channels for each measurement. a) Several cycles, and b) detail of the measurement.

Unexpected coupling effects could be identified to cause a deviation of the absolute perturbation switching voltage from its expected value, as can be seen on Figure 9, where the dependence of the extracted Hamming distance on the stored pattern is evidenced to cause a distance error varying between zero and four, thus reducing the absolute circuit accuracy in pulse-perturbation mode. Two extreme cases are considered, where the reference image loaded into the network consists of “Logic 1” and “Logic 0.”

Figure 10 shows the response linearity. Monotonic operation is confirmed by the lower limit of the acceptance interval being above 0V. In order to guarantee predictable step size of the perturbation, the upper limit of the acceptance interval should remain below a value equal to twice the lower limit. This is generally the case; however potential errors may result from extreme values, attributed to measurement artifacts.

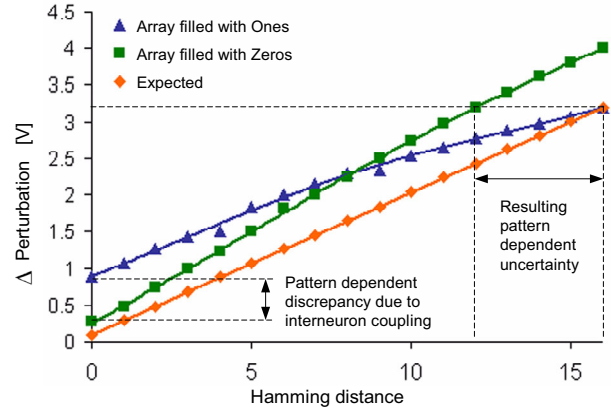


Fig. 9: Measured Hamming distance vs. perturbation voltage: expected, and two extreme patterns.

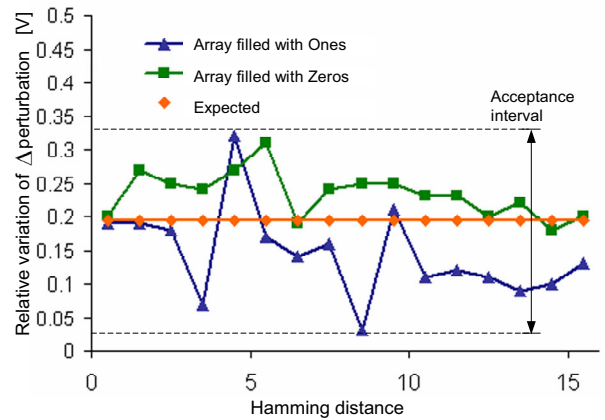


Fig. 10: Measured nonlinearity in the cases of Fig. 9.

The circuit fully qualifies for use with a ramp, or stepped ramp perturbation. A careful redesign is necessary for use in high-accuracy pulse-perturbation mode. Closed-loop operation and digital post-processing are available solutions to compensate for the observed excessive coupling.

REFERENCES

- [1] Richard P. Lippmann, “An Introduction to Computing with Neural Nets,” IEEE ASSP Magazine, April 1987.
- [2] U. Cilingiroglu, “A Charge-Based Neural Hamming Classifier,” IEEE Journal of Solid-State Circuits, Vol. 28, No. 1, January 1993.
- [3] Alexandre Schmid, Yusuf Leblebici and Daniel Mlynek, “Hardware Realization of a Hamming Neural Network with On-Chip Learning,” Proceedings of the 1998 IEEE International Symposium on Circuits and Systems ISCAS’98, Monterey, CA, 1998.
- [4] H. Özdemir, A. Kepkep, B. Pamir, Y. Leblebici, U. Cilingiroglu, “A Capacitive Threshold-Logic Gate,” IEEE Journal of Solid-State Circuits, Vol. 31, No. 8, August 1996.
- [5] S. Badel, A. Schmid, Y. Leblebici, “VLSI Realization of a Two-Dimensional Hamming Distance Comparator ANN for Image Processing Applications,” ESANN’03, Bruges, Belgium, April 2003.
- [6] S. Badel, A. Schmid, Y. Leblebici, “A VLSI Artificial Neural Network with k-Winner-Take-All and k-Loser-Take-All Capability,” IJCNN’03, Portland OR, July 2003.