1994 Conference on

Neural Information Processing Systems
(NIPS'94)

Denver, Colorado, November 28 - December 1

A Real Time Clustering CMOS Neural Engine

T. Serrano-Gotarredona, B. Linares-Barranco, and J. L. Huertas

Dept. of Analog Design, National Microelectronics Center (CNM), Ed. CICA, Av. Reina Mercedes s/n, 41012 Sevilla, SPAIN. Phone: (34)-5-4239923, Fax: (34)-5-4624506, E-mail: `bernabe@cnm.us.es`

Abstract

We describe an analog VLSI implementation of the ART1 algorithm (Carpenter, 1987). A prototype chip has been fabricated in a standard low cost 1.5µm double-metal single-poly CMOS process. It has a die area of $1cm^2$ and is mounted in a 120-pins PGA package. The chip realizes a modified version of the original ART1 architecture. Such modification has been shown to preserve all computational properties of the original algorithm (Serrano, 1994a), while being more appropriate for VLSI realizations. The chip implements an ART1 network with 100 *F1* nodes and 18 *F2* nodes. It can therefore cluster 100 binary pixels input patterns into up to 18 different categories. Modular expansibility of the system is possible by assembling an $N \times M$ array of chips without any extra interfacing circuitry, resulting in an *F1* layer with $100 \times N$ nodes, and an *F2* layer with $18 \times M$ nodes. Pattern classification is performed in less than 1.8µs, which means an equivalent computing power of 2.2×10^9 connections and connection-updates per second. Although internally the chip is analog in nature, it interfaces to the outside world through digital signals, thus having a true asynchrounous digital behavior. Experimental chip test results are available, which have been obtained through test equipments for digital chips.

1 INTRODUCTION

The original ART1 algorithm (Carpenter, 1987) proposed in 1987 is a massively parallel architecture for a self-organizing neural binary-pattern recognition machine. In response to arbitrary orderings of arbitrarily many and complex binary input patterns, ART1 is

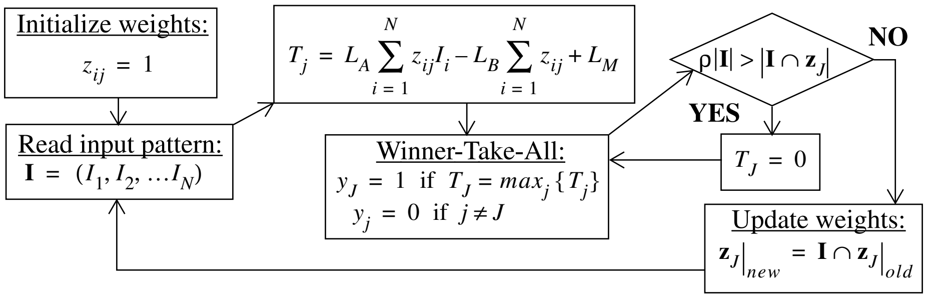


Fig. 1: Modified *Fast Learning* or *Type-3* ART1 implementation algorithm

capable of learning, in an unsupervised way, stable recognition codes. The ART1 architecture is described by a set of Short Term Memory (STM) and another set of Long Term Memory (LTM) time domain nonlinear differential equations. It is valid to assume that the STM equations settle much faster (instantaneously) than the LTM equations, so that the STM differential equations can be substituted by nonlinear algebraic equations that describe the steady-state of the STM differential equations. Furthermore, in the fast-learning mode (Carpenter, 1987), the LTM differential equations are as well substituted by their corresponding steady-state nonlinear algebraic equations. This way, the ART1 architecture can be behaviorally modelled by the sequential application of nonlinear algebraic equations. Three different levels of ART1 implementations (both in software and in hardware) can therefore be distinguished:

Type-1: Full Model Implementation: both STM and LTM time-domain differential equations are realized. This implementation is the most expensive (both in software and in hardware), and requires a large amount of computational power.

Type-2: STM steady-state Implementation: only the LTM time-domain differential equations are implemented. The STM behavior is governed by nonlinear algebraic equations. This implementation requires less resources than the previous one. However, a proper sequencing of STM events has to be introduced artificially, which is architecturally implicit in the *Type-1* implementation.

Type-3: Fast Learning Implementation: STM and LTM is implemented with algebraic equations. This implementation is computationally the less expensive one. In this case an artificial sequencing of STM and LTM events has to be done.

The implementation presented in this paper realizes a modified version of the original ART1 *Type-3* algorithm, more suitable for VLSI implementations. Such modified ART1 system has been shown to preserve all computational properties of the original ART1 architecture (Serrano, 1994a). The flow diagram that describes the modified ART1 architecture is shown in Fig. 1. Note that there is only one binary-valued weight template (z_{ij}), instead of the two weight templates (one binary-valued and the other real-valued) of the original ART1. For a more detailed discussion of the modified ART1 algorithm refer to (Serrano, 1994a, 1994b).

In the next Section we will provide an analog current-mode based circuit that implements in hardware the flow diagram of Fig. 1. Note that, although internally this circuit is analog in nature, from its input and output signals point of view it is a true asynchronous digital circuit, easy to interface with any conventional digital machine. Finally, in Section 3 we will provide experimental results measured from the chip using a digital data acquisition test equipment.

2 CIRCUIT DESCRIPTION

The ART1 chip reported in this paper has an *F1* layer with 100 neurons and an *F2* layer with 18 neurons. This means that it can handle binary input patterns of 100 pixels each, and cluster them into up to 18 different categories, according to a digitally adjustable vigilance parameter ρ. The circuit architecture of the chip is shown in Fig. 2(a). It consists of an array of 18×100 synapses, a 1×100 array of "vigilance synapses", a unity gain 18-outputs current mirror, an adjustable gain 18-outputs current mirror (with ρ=0.0, 0.1, ... 0.9)[1], 18 current-comparator-controlled switches and an 18-input-currents Winner-Take-All (WTA) (Serrano, 1994b). The inputs to the circuit are the 100 binary digital input voltages I_i, and the outputs of the circuit are the 18 digital output voltages y_j. External control signals allow to change parameters ρ, L_A, L_B, and L_M. Also, extra circuitry has been added for reading the internal weights z_{ij} while the system is learning.

Each row of synapses generates two currents,

$$\begin{aligned} T_j &= L_A \sum_{i=1}^{100} z_{ij} I_i - L_B \sum_{i=1}^{100} z_{ij} + L_M \\ V_j &= L_A \sum_{i=1}^{100} z_{ij} I_i \end{aligned} \tag{1}$$

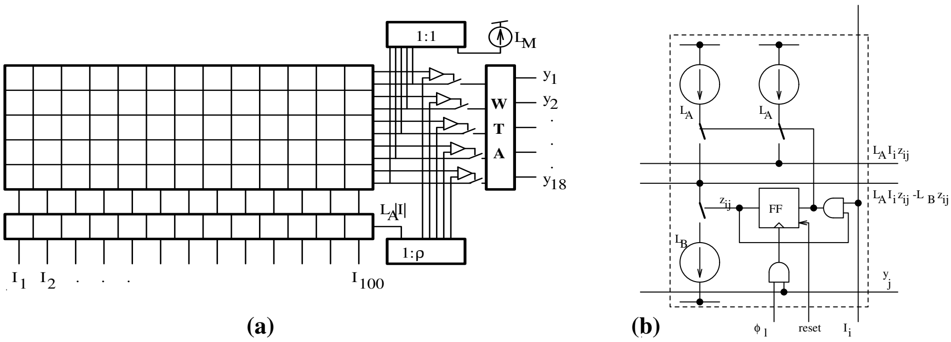
while the row of the "vigilance synapses" generates the current

$$V_\rho = L_A \sum_{i=1}^{100} I_i \tag{2}$$

Each of the current comparators compares the current V_j versus ρV_ρ, and allows current T_j to reach the WTA only if $\rho V_\rho \le V_j$. This way competition and vigilance occur simultaneously and in parallel, speeding up significantly the search process.

Fig. 2(b) shows the content of a synapse in the 18x100 array. It consists of three current sources with switches, two digital AND gates and a flip-flop. Each synapse receives two input voltages I_i and y_j, and two global control voltages ϕ_l (to enable/disable learning) and *reset* (to initialize all weights z_{ij} to '1'). Each synapse generates two currents $L_A I_i z_{ij} - L_B z_{ij}$ and $L_A I_i z_{ij}$,which will be summed up for all the synapses in the same row to generate the currents T_j and V_j. If learning is enabled ($\phi_l = 1$) the value of z_{ij} will change to $I_i z_{ij}$ if $y_j = 1$. The "vigilance synapses" consist each of a current-source of value L_A with a switch controlled by the input voltage I_i. The current comparators are those proposed in (Domínguez-Castro, 1992), the WTA used is reported in (Lazzaro, 1989), and the digitally adjustable current mirror is based on (Loh, 1989), while its continuous gain fine tuning mechanism has been taken from (Adams, 1991).

1. An additional pin of the chip can fine-tune ρ between 0.9 and 1.0.



(a) **(b)**

Fig. 2: (a) System Diagram of Current-Mode ART1 Chip, (b) Circuit Diagram of Synapse

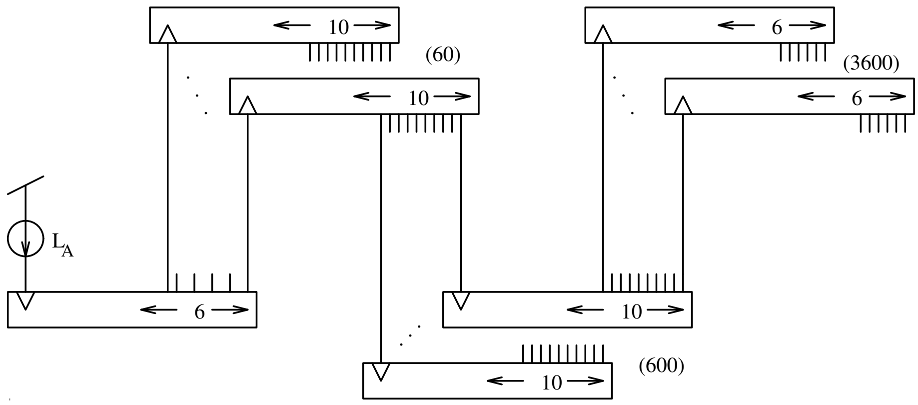


Fig. 3: Tree based current-mirror scheme for matched current sources

The circuit has been designed in such a way that the WTA operates with a precision around 1.5% (~6 bits). This means that all L_A and L_B current sources have to match within an error of less than that. From a circuit implementation point of view this is not easy to achieve, since there are 5500 current sources spread over a die area of $1cm^2$. Typical mismatch between reasonable size MOS transistors inside such an area extension can be expected to be above 10% (Pelgrom, 1989). To overcome this problem we implemented a tree-based current mirror scheme as is shown in Fig. 3. Starting from a unique current reference, and using high-precision 10(or less)-outputs current mirrors (each yielding a precision around 0.2%), only up to four cascades are needed. This way, the current mismatch attained at the synapse current sources was around 1% for currents between $L_{A/B} = 5\mu A$ and $L_{A/B} = 10\mu A$. This is shown in Fig. 4, where the measured dc output current-error (in %) versus input current of the tree based configuration for 18 of the 3600 L_A synapse sources is depicted.

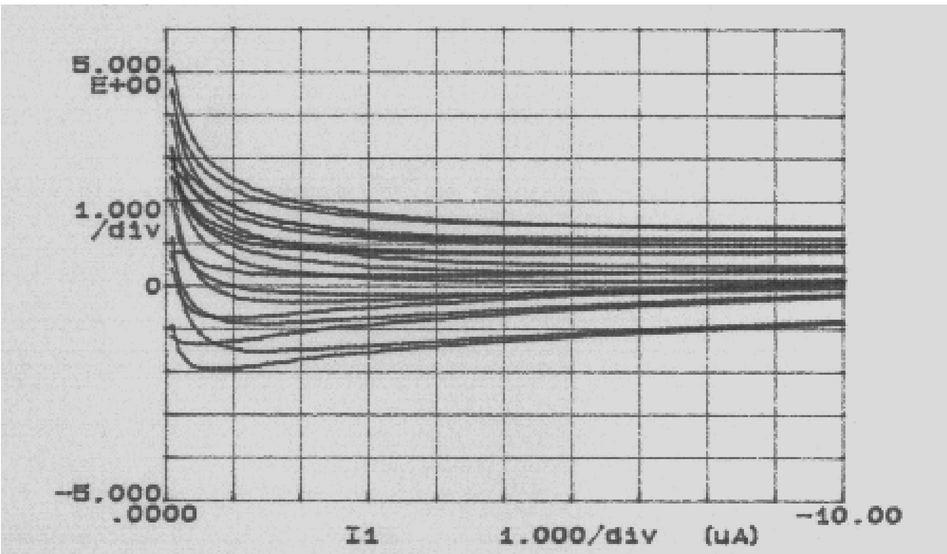


Fig. 4: Measured current mirror cascade missmatch (1%/div) for L_A for currents below 10µA

3 EXPERIMENTAL RESULTS

Fig. 5 shows a microphotograph of a prototype chip fabricated in a standard digital double-metal, single-poly 1.5µm low cost CMOS process. The chip die area is $1cm^2$, and it is mounted in a 120-pins PGA package. Fig. 6 shows a typical training sequence accomplished by the chip and obtained experimentally using a test equipment for digital chips. The only task performed by the test equipment was to provide the input data patterns **I** (first column in Fig. 6), detect which of the output nodes became '1' (pattern with a vertical bar to its right), and extract the learned weights. Each 10×10 square in Fig. 6 represents either a 100-pixels input vector **I**, or one row of 100-pixels synaptic weights $\mathbf{z}_j \equiv (z_{1j}, z_{2j}, \ldots z_{100j})$. Each row of squares in Fig. 6 represents the input pattern (first square) and the 18 vectors $\mathbf{z}_j$ after learning has been performed for this input pattern. The sequence shown in Fig. 6 has been obtained for $\rho = 0.7$, $L_A = 10\mu A$, $L_B = 9.5\mu A$, and $L_M = 950\mu A$. Only two iterations of input patterns presentations were necessary, in this case, for the system to learn and self-organize in response to these 18 input patterns.

The last row in Fig. 6 shows the final learned templates. Fig. 7 shows final learned templates for different values of ρ. The numbers below each square indicate the input patterns that have been clustered into each $\mathbf{z}_j$ category.

Delay time measurements have been performed for the feedforward action of the chip (establishment of currents T_j, V_j, and V_ρ, and their competitions until the WTA settles), and for the updating of weights. The feedforward delay is pattern and bias currents (L_A, L_B, L_M) dependent, but has been measured to be always below $1.6\mu s$. The learning time is constant and is around $180ns$. Therefore, throughput time is less than $1.8\mu s$. A digital neuroprocessor able to perform *a* connections/s, *b* connection-updates/s, and with a dedicated WTA section with a *c* seconds delay, must satisfy

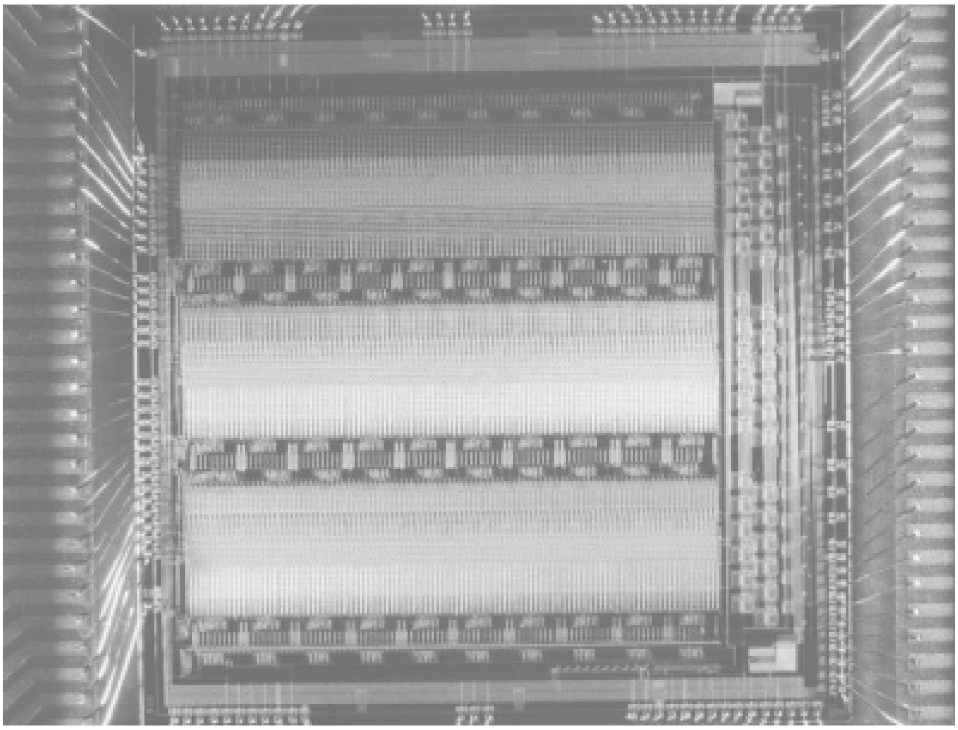


Fig. 5: Microphotograph of ART1 chip

$$\frac{3700}{a} + \frac{100}{b} + c = 1.8\mu s \qquad (3)$$

to meet the performance of our prototype chip. If $a = b$ and $c = 100ns$, the equivalent speed would be $a = b = 2.2 \times 10^9$ connections and connection-updates per second.

4 CONCLUSIONS

A high speed analog current-mode categorizer chip has been built using a standard low cost digital CMOS process. The high performance of the chip is achieved thanks to a simplification of the original ART1 algorithm. The simplifications introduced are such that all the original computational properties are preserved. Experimental chip test results are provided.

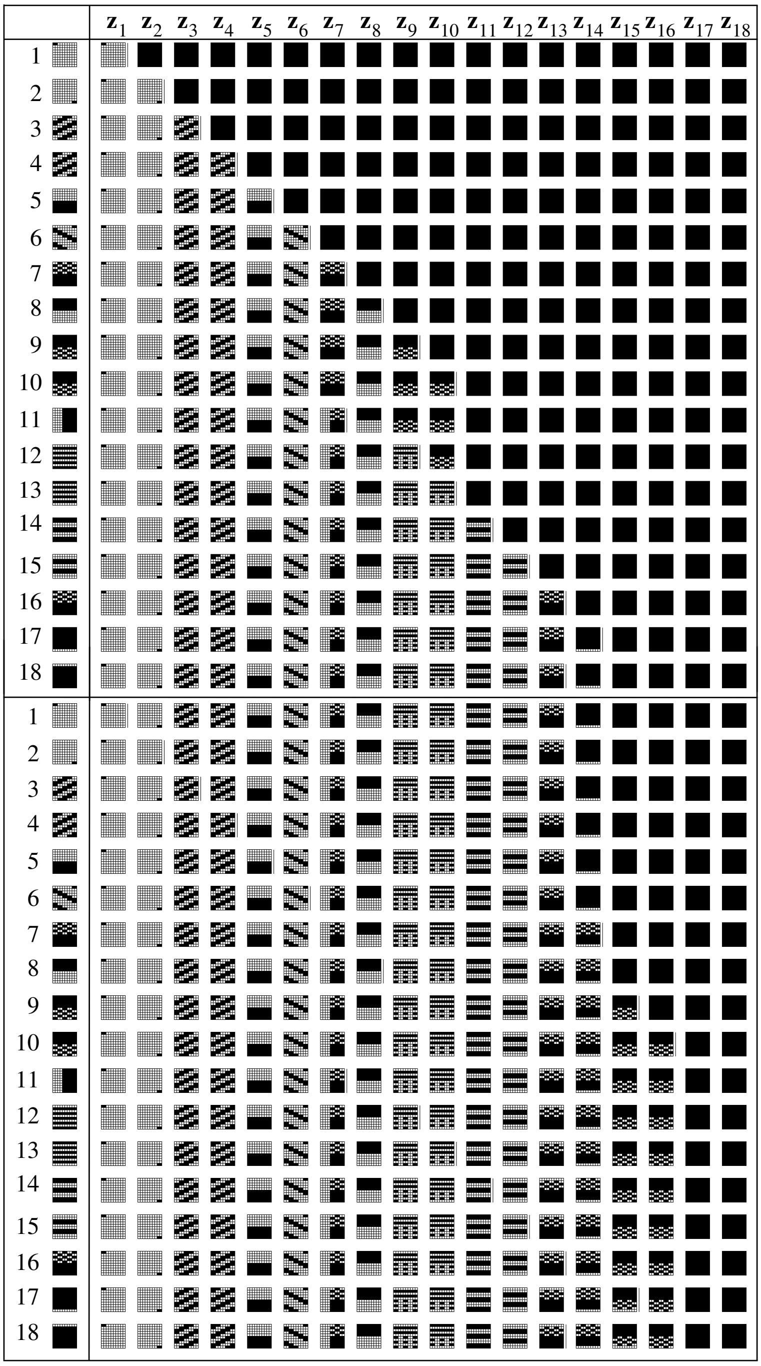


Fig. 6: Test sequence obtained experimentally for ρ=0.7, L_A=10μA, L_B=9.5μA, and L_M=950μA

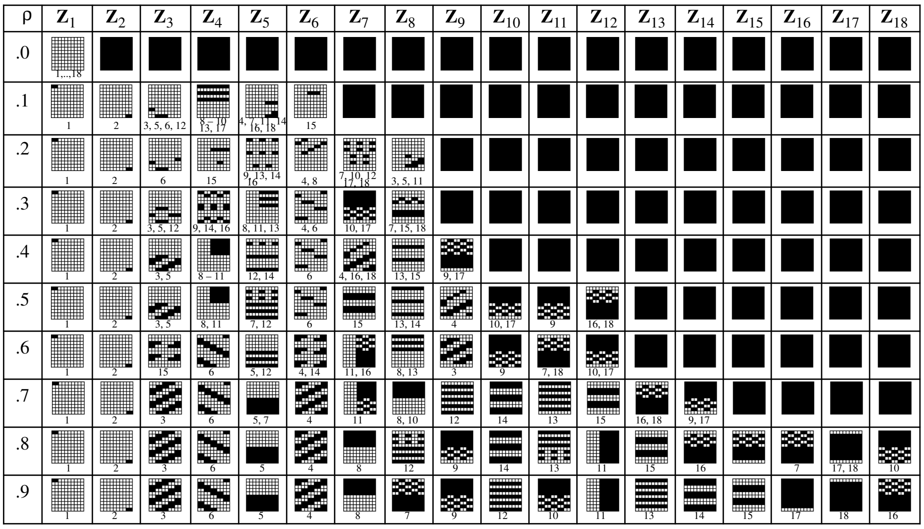


Fig. 7: Categorization of the input patterns for L_A=3.2μA, L_B=3.0μA, L_M=400μA, and different values of ρ

References

W. J. Adams and J. Ramímez-Angulo. (1991) "Extended Transconductance Adjustment/Linearisation Technique," *Electronics Letters,* vol. 27, No. 10, pp. 842-844, May 1991.

G. A. Carpenter and S. Grossberg. (1987) "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54-115, 1987.

R. Domínguez-Castro, A. Rodríguez-Vázquez, F. Medeiro, and J. L. Huertas. (1992) "High Resolution CMOS Current Comparators," *Proc. of the 1992 European Solid-State Circuits Conference (ESSCIRC'92),* pp. 242-245, 1992.

J. Lazzaro, R. Ryckebusch, M. A. Mahowald, and C. Mead. (1989) "Winner-Take-All Networks of O(n) Complexity," in *Advances in Neural Information Processing Systems*, vol. 1, D. S. Touretzky (Ed.), Los Altos, CA: Morgan Kaufmann, 1989, pp. 703-711.

K. Loh, D. L. Hiser, W. J. Adams, and R. L. Geiger. (1989) "A Robust Digitally Programmable and Reconfigurable Monolithic Filter Structure," *Proc. of the 1989 Int. Symp. on Circuits and Systems (ISCAS'89),* Portland, Oregon, vol. 1, pp. 110-113, 1989.

M. J. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers. (1989) "Matching Properties of MOS Transistors," *IEEE Journal of Solid-State Circuits,* vol. 24, No. 5, pp. 1433-1440, October 1989.

T. Serrano-Gotarredona and B. Linares-Barranco. (1994a) "A Modified ART1 Algorithm more suitable for VLSI Implementations," submitted for publication (journal paper).

T. Serrano-Gotarredona and B. Linares-Barranco. (1994b) "A Real-Time Clustering Microchip Neural Engine," submitted for publication (journal paper).

1995 IEEE International Symposium on Circuits and Systems (ISCAS'95)

Seattle, Washington, April 29 - May 3, 1995

Experimental Results of An Analog Current-Mode ART1 Chip

T. Serrano-Gotarredona and B. Linares-Barranco

National Microelectronics Center (CNM), Ed. CICA, Av. Reina Mercedes s/n, 41012 SPAIN. Phone: 34-5-4239923
Fax: 34-5-4624506, E-mail: bernabe@cnm.us.es

Abstract

A hardware implementation of a modified version of the ART1 algorithm has been built. The modified algorithm preserves all the computational properties of the original one but allows a more efficient hardware implementation [1]. A prototype chip has been fabricated in a standard 1.5 μm low cost single poly CMOS process. The chip implements an ART1 network with 100 *F1* nodes and 18 *F2* nodes. Therefore, it can cluster input patterns with 100 binary pixels into up to 18 categories. The estimated time needed to classify and learn an input pattern is less than 1 μs, which means a computing power of 1.8 Gconnetions/s. The chip has a die area of $1cm^2$ and is mounted on a 120-pins PGA package. The operation of the chip is analog, mainly based on currents additions and a Winner-Take-All action, but all the inputs and outputs are digital voltages, making the chip digitally compatible. Experimental results of time measurements, current mismatches among the synapses and the global operation of the system are available.

I. Introduction

This paper is a continuation of previously published work [2] and focuses on providing experimental results.

ART1 is the first of a series [3] of algorithms published by Carpenter and Grossberg. An ART1 neural network is able to learn in an unsupervised way to cluster binary input patterns into different categories. It forms stable recognition codes in response to arbitrarily ordered and arbitrary many and complex binary input patterns. Carpenter and Grossberg also introduced the fast learning mode of the ART1 algorithm.

The circuit presented in this paper performs a modified version of the fast learning mode of the ART1 algorithm, as is depicted in Fig. 1 [1]-[2]. The modification introduces several simplifications that allow its hardware realization using more efficient building blocks while preserving all the computational properties of the original algorithm [1].

The architecture consists of two layers. Layer *F1* or input layer has N nodes, one for each binary pixel of the input pattern. The second layer *F2* is the category layer and has M nodes. Both layers are fully connected by a matrix of binary weights z_{ij} ($1 \le i \le N$ and $1 \le j \le M$).

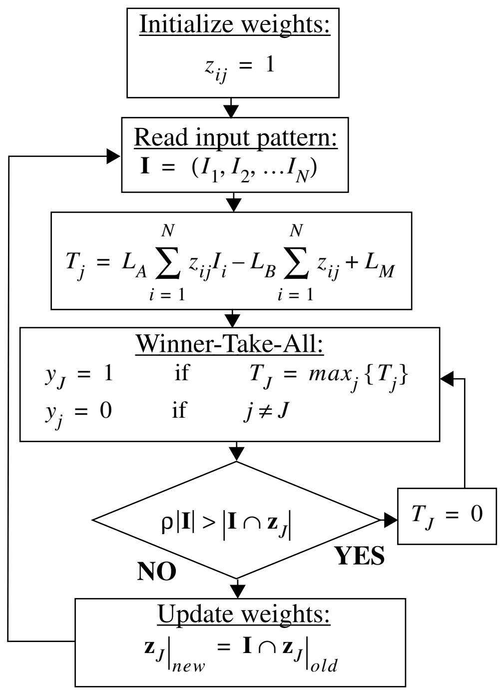


Fig. 1: Modified *Fast Learning* ART1 algorithm

Each time an input pattern $\mathbf{I}=(I_1,...,I_N)$ is presented the result of the operation

$$T_j = L_A \sum_{i=1}^{N} I_i z_{ij} - L_B \sum_{i=1}^{N} z_{ij} + L_M \tag{1}$$

is computed as the input to each *F2* category node j.

Layer *F2* is organized as a Winner-Take-All network, so that the category node J which receives the largest input (which is the one whose stored pattern $\mathbf{z}_J$ [1] most closely resembles the input pattern $\mathbf{I}$), will output a "1", while the other category nodes j will output a "0".

Once a winning category has been selected the matching between the input pattern $\mathbf{I}$ and the active template $\mathbf{z}_J$ is checked. A comparison takes place between $|\mathbf{I} \cap \mathbf{z}_J|$ and the input pattern $|\mathbf{I}|$ according to a "vigilance criterion".

a) If $|\mathbf{I} \cap \mathbf{z}_J| \ge \rho|\mathbf{I}|$, the vigilance criterion is satisfied and the input pattern $\mathbf{I}$ is classified in category J.

1. $\mathbf{z}_j \equiv (z_{1j}, z_{2j}, \ldots, z_{Nj})$

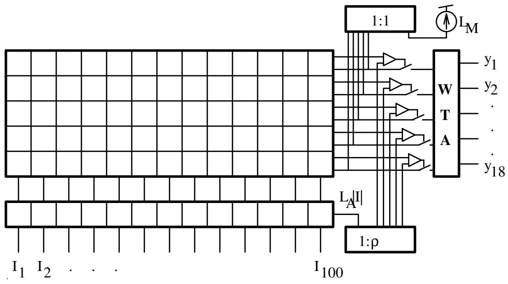


Fig. 2: System Diagram of Current-Mode ART1 Chip

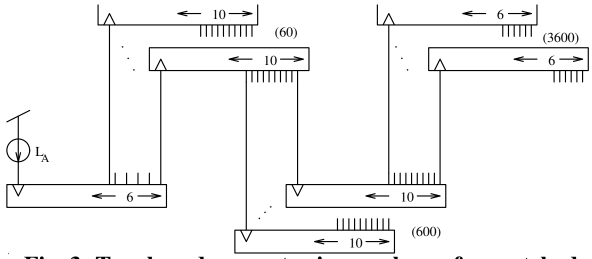


Fig. 3: Tree based current-mirror scheme for matched current sources

b) If $|\mathbf{I} \cap \mathbf{z}_J| < \rho|\mathbf{I}|$, the vigilance criterion is not satisfied, category J is inhibited and the network looks for another winner. The searching process continues until a winner that satisfies the vigilance criterion is found or a new category is formed.

Here ρ is the "vigilance parameter" which is a tunable parameter that takes values between 0 and 1.

Once the input pattern has been accepted as belonging to a category J, the weights of this category z_{iJ} are updated to incorporate the new knowledge, following the rule

$$\mathbf{z}_J\big|_{new} = \mathbf{I} \cap \mathbf{z}_J\big|_{old} \tag{2}$$

In Section II the diagram of the circuit that implements this algorithm is explained and in Section III experimental results of the tests performed on the circuit are summarized.

II. Circuit Description

A system diagram of the ART1 chip reported in this paper is depicted in Fig. 2. The system has a 100 nodes *F1* layer and an 18 nodes *F2* layer. This means that it is able to cluster input patterns with 100 binary pixels into up to 18 different categories. It is also possible to increase the number of input pixels to $100 \times N$ and the number of categories to $18 \times M$ by interconnecting a matrix of $N \times M$ chips.

A single chip consists of a matrix of 100×18 synapses, an array of 100 "vigilance synapses", a current mirror of tunable gain ρ, an array of 18 current comparators, 18 switches controlled by the current comparators outputs, and an 18-inputs WTA circuit.

The array of "vigilance synapses" generates a current $L_A \sum I_i = L_A|\mathbf{I}|$. This current enters the adjustable gain 18-outputs current mirror, which distributes an equal current $L_A \rho |\mathbf{I}|$ to the input of each current comparator.

Each synapse outputs two currents:

a) A current $L_A I_i z_{ij}$ that flows to a common node for all the synapses in the same row resulting in a total current $L_A \sum I_i z_{ij} = L_A|\mathbf{I} \cap \mathbf{z}_j|$ that enters in the *jth* current comparator.

b) And a current $L_A I_i z_{ij} - L_B z_{ij}$ that results in a total current $T_j = L_A \sum I_i z_{ij} - L_B \sum z_{ij} + L_M$ for each row of synapses *j* which is the input to the *jth* branch of the WTA circuit.

Each current comparator receives an input $L_A|\mathbf{I} \cap \mathbf{z}_j| - L_A \rho |\mathbf{I}|$ so that if this input current is negative the vigilance criterion is not satisfied for category *j* and the switch controlled by the current comparator output prevents current T_j from reaching the WTA, this way the searching process is implemented in a parallel way, instead of serially, speeding up considerably the complete process.

Each weight z_{ij} is updated when a learn signal Φ_l is enabled, and if $y_j = 1$, changing its stored value to $z_{ij} I_i$.

III. Experimental results

The chip has been fabricated in a 1.5µm standard low cost single poly CMOS process. It has a die area of 1cm^2 and has been mounted on a 120-pins PGA package.

Several elements of the chip have separately been measured and the whole chip operation has been tested on a digital testing equipment. The results of the different measurements are provided in this Section.

A. Current mismatches

The loss of precision of the system operation is due to two main error sources: current replication error in the synapses, and mismatches between transistors in the different branches of the WTA.

Current L_A has to be replicated $2 \times 1800 + 100 = 3700$ times all over the chip, while current L_B is needed 1800 times. If the current sources were spread over the chip a typical mismatch for MOS transistors of a reasonable size in such an area would be above 10%, which would result in a very poor precision in the system operation [4]. To overcome this limitation the current sources have been implemented in a tree structure as is shown in Fig. 3 for the L_A current sources. The current mirrors in each stage are implemented with their transistors organized in a common centroid structure. Fig. 4 depicts the measured input-output errors for 18 different L_A currents. It can be noted how the error is around 1% for currents between 5µA and 10µA.

B. WTA Precision

The WTA structure is based on the one reported by Lazzaro [5]. The deviation among the different branches of the WTA has been measured for several levels of input currents. Table 1 shows the measurements performed for an input current T_j of 10µA, 100µA and 1mA. One input is maintained constant at T_j, a second input is swept from $0.95 \times T_j$ to $1.05 \times T_j$, and the rest are 0µA.

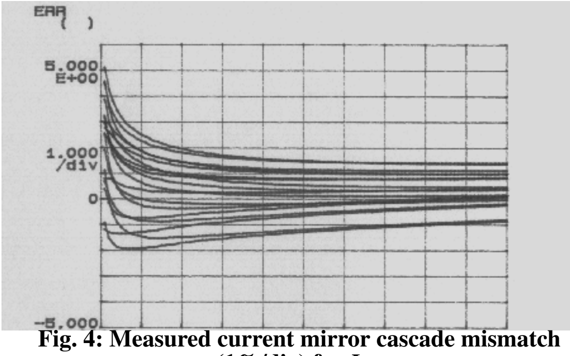


Fig. 4: Measured current mirror cascade mismatch (1%/div) for L_A

T_j	10µA	100µA	1mA
σT_j	1.73%	.86%	.99%

Table 1. Precision of the WTA

C. Global System Performance

The system has been trained with a sequence of 18 input patterns of 10×10 binary pixels, which are depicted in Fig. 5.The data were loaded externally by a digital test equipment. Then, for each presented pattern, the winning category and the updated weights were read. Generally, several presentations of the input patterns are needed until learning self-stabilizes. The training sequence was applied for different values of parameter $\alpha = L_A / L_B$ and vigilance parameter ρ.

Fig. 6 depicts the categories formed when learning has stabilized after the presentation of the 18 input patterns for $\alpha = 1.07$ and ρ varying from 0 to 0.9. The numbers below each category indicate the input patterns that have been clustered in each category[2]. It can be noted how increasing the vigilance parameter has the effect of increasing the number of categories.

Fig. 7 shows the test results for $\rho = 0$, so that no reset by the vigilance subsystem occurs and the different clustering of patterns is controlled by parameter α, which varies from 1.07 to 50.

D. Yield and Fault Tolerance

All the prototypes have been exhaustively tested, concluding that only 2 out of 30 samples were totally fault free, which corresponds to a yield of 6.7%. However, 12 of them (40%) were operative although having important defects that affected mainly the values of the currents in the synapses. Fig. 8 contains the results of the training test applied to the 12 operative samples, for $\alpha = 1.07$ and $\rho = 0.5$.

For each sample, it is indicated how many iterations of the input patterns are needed to stabilize the learning, as well as the resulting categories. The input patterns clustered in each category are indicated by the numbers below each category.

Only samples 1 and 2 are fault free but, as Fig. 8 shows, 6 of the 12 samples do exactly the same classification as the fault free samples, while the others perform a different but valid clustering.

E. Operation Speed

Throughput times are the sum of the WTA delay plus the learning time.

The delay between the activation of the learn signal and the updating of the weights is 190ns. Table 2 contains measurements of the WTA delays. The input current T_1 switches between T_1^a and T_1^b, causing the winning node to change from category y_1 to y_2. T_{d1} is the time needed by category y_1 to win when T_1 switches from T_1^a to T_1^b, and T_{d2} is the time spent by category y_2 in winning when T_1 decreases from T_1^b to T_1^a.

T_1^a	T_1^b	T_2	$T_3,..,T_{18}$	T_{d1}	T_{d2}
0µA	200µA	100µA	0µA	659ns	466ns
0µA	1mA	500µA	0µA	550ns	570ns
100µA	150µA	125µA	100µA	440ns	399ns
400µA	600µA	500µA	400µA	209ns	460ns
500µA	1.5mA	1mA	500µA	231ns	316ns

Table 2. Classification times of the WTA

IV. Conclusions

An analog current-mode high-speed real-time clustering neural network has been built in a standard CMOS process. The chip is based on a simplification of the ART1 algorithm that allows a more efficient hardware implementation while preserving all the computational properties. Experimental results based on exhaustive tests performed on the chips are provided.

V. References

[1] T. Serrano and B. Linares-Barranco, “A Modified ART1 Algorithm more suitable for VLSI implementations,” *IEEE Transactions on Neural Networks*, submitted for publication.

[2] T. Serrano, B. Linares-Barranco, and J. L. Huertas, “A VLSI-friendly ART1 Neural Algorithm and its CMOS Analog Current-Mode Implementation,” *Proc. of the 1994 Int. Conf. on Microelectronics for Neural Networks and Fuzzy Systems,* pp. 405-411, Torino, Italy, 1994.

[3] G. A. Carpenter and S. Grossberg, “A Massively Parallel Architecture for a self-organizing Neural Pattern Recognition Machine,” *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54-115, 1987.

[4] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, “Matching Properties of MOS Transistor,” *IEEE Journal of Solid-State Circuits*, vol 24, October 1989, pp. 1433-1440.

[5] J. Lazzaro, R. Ryckebush, M. A. Mahowald, and C. Mead, “Winner-Take-All Networks of O(n) Complexity,” in *Advances in Neural Information Processing Systems*, vol. 1, D. S. Touretzky(Ed.), Los Altos, CA: Morgan Kaufmann, 1989, pp. 703-711.

2. If there is no number, none of the input patterns belongs to this category. This is a known ART1 behavior: during the clustering process some categories might be created that will not represent any of the training patterns.

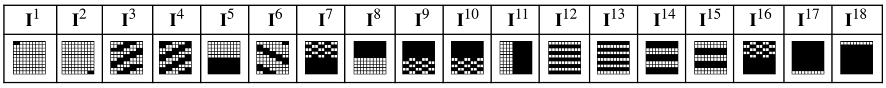


Fig. 5: Set of input patterns

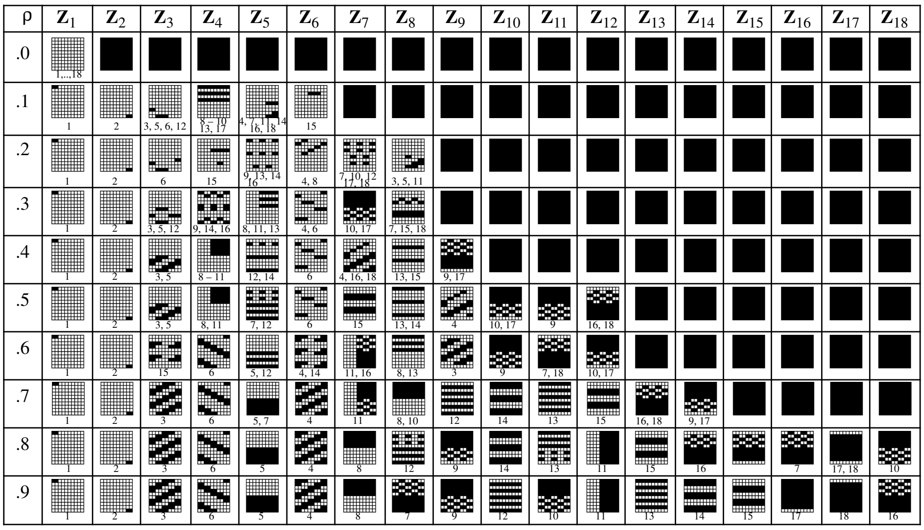


Fig. 6: Categorization of the input patterns for L_A=3.2μA, L_B=3μA, L_M=400μA, and different values of ρ

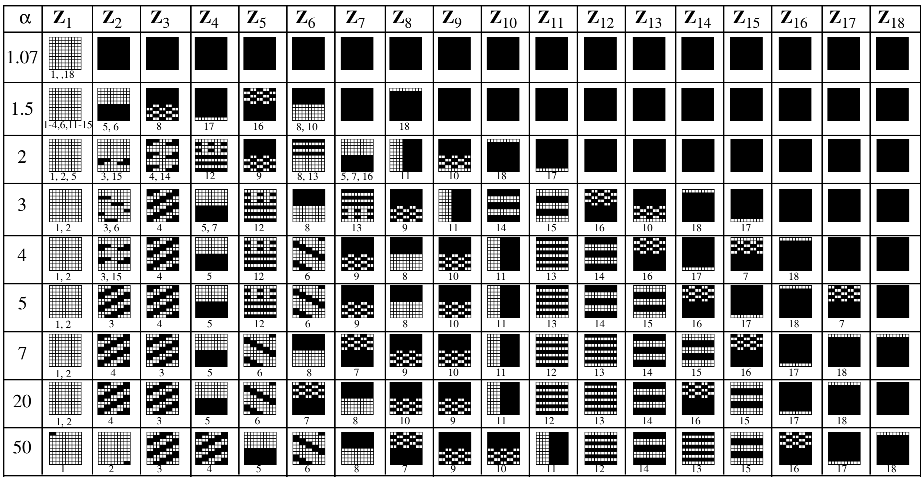


Fig. 7: Categorization of the input patterns for ρ=0 and different values of α

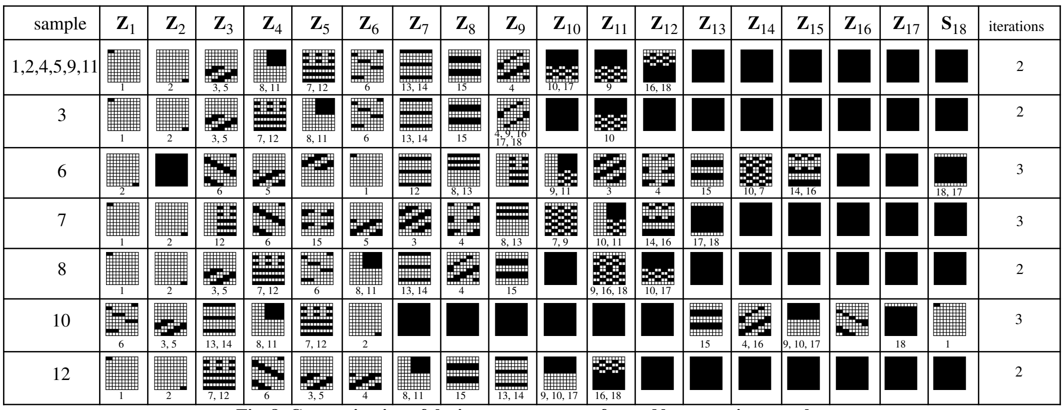


Fig. 8: Categorization of the input patterns performed by operative samples