An efficient mapping of Fuzzy ART onto a neural architecture

Matthias Blume* and Sadik C. Esener
Department of Electrical and Computer Engineering
University of California at San Diego, La Jolla, CA 92093

Abstract

A novel mapping of the Fuzzy ART algorithm onto a neural network architecture is described. The architecture does not utilize bi-directional synapses, weight transport, or weight duplication, and requires one fewer layer of processing elements than the architecture originally proposed by Carpenter, Grossberg, & Rosen (1991a). In the new architecture, execution of the algorithm takes constant time per input vector regardless of the relationship between the input and existing templates, and several control signals are eliminated. This mapping facilitates hardware implementation of Fuzzy ART and furthermore serves as a tool for envisioning and understanding the algorithm.

Keywords: Fuzzy ART, Fuzzy ARTMAP, parallel hardware, neural architecture.

Fuzzy ART is a clustering algorithm that operates on vectors with analog-valued elements (Carpenter, Grossberg, & Rosen, 1991a). Adding a further layer of processing to Fuzzy ART yields a *supervised* clustering algorithm, Fuzzy ARTMAP (Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992). These algorithms have been successfully applied to numerous tasks, including speech recognition (Young, 1995), handwritten character recognition (Markuzon, 1994), and target recognition from radar range profiles (Rubin, 1994a). Such applications often require the formation of a very large number of clusters in a high-dimensional feature space, necessitating parallel implementation of the algorithm for high-speed or real-time classification. Several properties of Fuzzy ART facilitate implementation in hardware. Notably, because it uses the L_1 distance metric, multiplication is not required at each synapse. Furthermore, the algorithm can perform well with as few as four bits of weight precision (Rubin, 1994b). Consequently, very little circuit area is required at each synapse in a neural implementation. However, as originally specified, Fuzzy ART requires bi-directional synapses, weight transport, or weight duplication, any of which is troublesome for a parallel implementation. We describe a novel mapping of Fuzzy ART onto a neural architecture that eliminates this requirement. This mapping relies on a different sequencing of the operations of the algorithm, but the classification of the input vectors remains unchanged.

Fuzzy ART clusters vectors based on two separate distance criteria, *match* and *choice*. For input vector $\boldsymbol{I}$ and category j, the match function is defined by

$$S_j(\boldsymbol{I}) \equiv \frac{|\boldsymbol{I} \wedge \boldsymbol{w}_j|}{|\boldsymbol{I}|} \ ,$$

where $\boldsymbol{w}_j$ is an analog-valued weight vector associated with cluster j. $\wedge$ denotes the fuzzy AND operator, $(\boldsymbol{p} \wedge \boldsymbol{q})_i \equiv min(p_i, q_i)$, and the norm $|\cdots|$ is defined by $|\boldsymbol{p}| \equiv \sum_i |p_i|$. The choice function

*matthias@ucsd.edu
Submitted to *Neural Networks* August, 1995

is defined by

$$T_j(\boldsymbol{I}) \equiv \frac{|\boldsymbol{I} \wedge \boldsymbol{w}_j|}{\alpha + |\boldsymbol{w}_j|} ,$$

where α is a small constant. Increasing α biases the search more towards clusters with large $\boldsymbol{w}_j$. Each input vector is assigned to the category that maximizes $T_j(\boldsymbol{I})$ while satisfying $S_j(\boldsymbol{I}) \geq \rho$, where the *vigilance*, ρ, is a constant, $0 \leq \rho \leq 1$.

Carpenter, et al. (1991a) propose searching for the category J which maximizes $T_j(\boldsymbol{I})$ and then checking whether $S_J(\boldsymbol{I}) \geq \rho$. If not, category J is marked as ineligible (*reset*) and the search is repeated until a satisfactory category is found (*resonance*). The neural network realization described by Carpenter, Grossberg, & Rosen (1991a, 1991b) is a three-layer architecture, illustrated in figure 1. Calculation of the choice function takes place in layer F_2, whereas calculation of the match function and comparison to the vigilance parameter takes place in layer F_1 and at the reset node. Since the weight values, w_{ji}, are required for both calculations, copies are stored and independently updated in each of the two layers. The length of time between input presentation and selection of the corresponding cluster is variable, depending on how many search cycles are required.

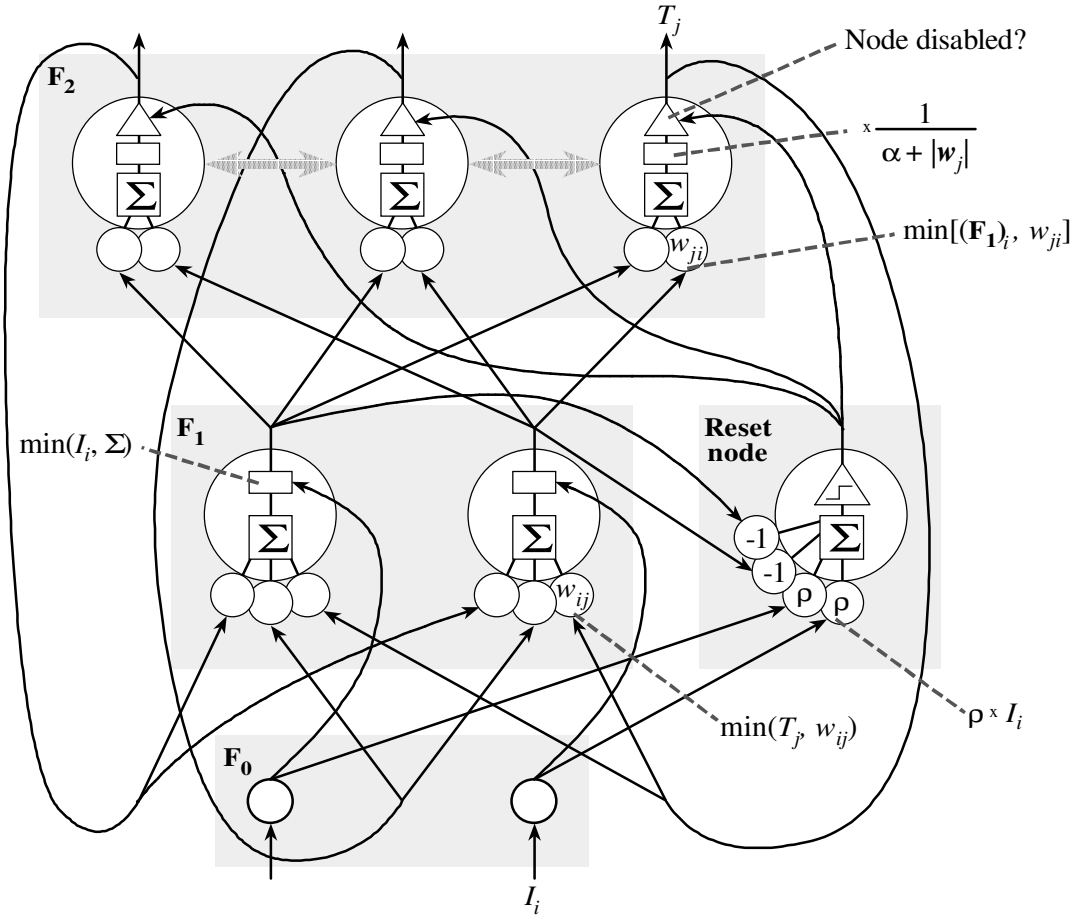

Figure 1: The Fuzzy ART architecture proposed by Carpenter, Grossberg, & Rosen (1991a) consists of three layers of processing elements and utilizes two identical sets of weights. Some control signals are not shown.

We propose an alternate approach, namely checking the match criterion for every cluster and then comparing the choice functions of only those clusters satisfying the match criterion. Figure 2 represents an architecture based on this order of the operations. The choice and match functions are both computed in the same layer, thus only one set of weights is stored and updated. Since both the match and choice functions are computed in parallel for all clusters, execution of the algorithm takes a constant length of time for any input vector. Finally, each processing element (PE) executes the same operations at every time step, eliminating the control signals associated with the reset / resonance process and facilitating pipelining.

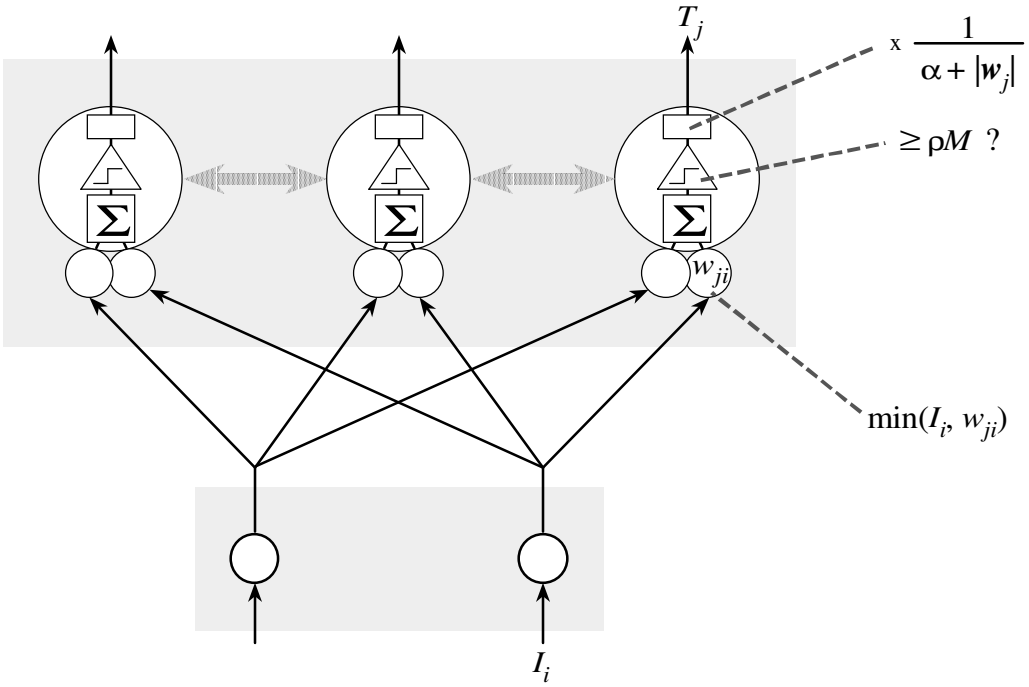

Figure 2: Novel Fuzzy ART architecture, consisting of two layers and requiring only one set of weights. Complement coding of the input vectors, eliminating the need for computation of $|\boldsymbol{I}|$, is assumed. In hardware, the winner-take-all interconnections (shaded arrows) may be replaced by one additional processing element that determines which layer 2 PE has the maximal output.

In the proposed architecture, hardware for computing the match criterion must be included in every processing element. Thus, complexity of this operation would be a drawback for the scheme. Fortunately, computation of the match criterion requires minimal additional effort. The relationship $S_j(\boldsymbol{I}) \geq \rho$ is equivalent to $|\boldsymbol{I} \wedge \boldsymbol{w}_j| \geq \rho|\boldsymbol{I}|$. The quantity on the left hand side of this equation is needed for the computation of $T_j(\boldsymbol{I})$ in any case, and thus requires no additional calculations. Typically, a complement coding scheme is utilized which guarantees that the L_1 lengths of all input vectors are identical (Carpenter et al., 1991a; Huang, Georgiopoulos, & Heileman, 1995). In this case, $\rho|\boldsymbol{I}|$ is a constant, M, and the evaluation of the match criterion is reduced to a single comparison, a simple operation for analog or digital electronics. If complement coding is not used, the magnitude of every input vector, $|\boldsymbol{I}|$, must be calculated. This may be accomplished by a single additional processing element that transmits its output value to the PEs of the second layer. In either case, the additional hardware required for the parallel computation of the match criterion is nearly negligible, and is certainly outweighed by the elimination of an entire layer of processing elements.

The fuzzy ART learning rule is given by

$$w_{Ji}^{new} = \begin{cases} w_{Ji}^{old} & w_{Ji} \le I_i \\ w_{Ji}^{old} - \beta\left(w_{Ji}^{old} - I_i\right) & w_{Ji} > I_i \end{cases},$$

where $0 < \beta \le 1$. In either architecture, the weights are updated using only locally available information. However, in the architecture proposed here, there is only one copy of the weights, so only one set of circuitry for modifying the weights is required. The Fuzzy ART algorithm specifies that only the weights of the cluster to which $\boldsymbol{I}$ has been assigned are updated. This characteristic is particularly advantageous for implementations in which secondary storage is utilized and the memory read bandwidth is greater than the write bandwidth (Blume & Esener, 1995).

The benefits of this architecture in terms of computation time and hardware complexity are clear for parallel implementations. However, the effect of computing the match criterion before the choice function (rather than visa versa) on serial execution time depends on the details of the implementation of the algorithm and on the distribution of the input vectors. If N clusters are formed, evaluation of the match criterion for all clusters requires N comparisons. After the computation of the choice functions for a given input vector, determining which $T_j(\boldsymbol{I})$ is maximal also requires approximately N comparisons. Thus, if the original scheme requires two or more search cycles on average, calculating the match criterion before the choice function would save time even in a serial implementation.

In summary, we have proposed an architecture that takes advantage of the simple operations and low numerical precision required by Fuzzy ART while eliminating a number of control signals and the weight replication that were previously suggested. It is especially well suited to implementation in parallel hardware. The classification of the input vectors remains identical to that produced by the original algorithm. We find that the new mapping onto a less complicated architecture promotes an intuitive understanding of the Fuzzy ART clustering process, making it easy to envision how each input vector is assigned to the cluster that maximizes the choice function while satisfying the match criterion.

Acknowledgment

This research was supported by the Advanced Research Projects Agency
(ARPA F49620-92-J-0520).

References

Blume, M. & Esener, S. C. (1995). Optoelectronic Fuzzy ARTMAP processor. *1995 OSA Technical Digest Series*, **10** (pp. 213-215). Washington, D.C.: Optical Society of America.

Carpenter; G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, **3**(5), 698-713.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991a). Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, **4**(6), 759-771.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991b). *A neural network realization of Fuzzy ART* (Technical Report CAS/CNS-91-021). Boston, MA: Boston University, Center for Adaptive Systems.

Markuzon, N. (1994). Handwritten digit recognition using Fuzzy ARTMAP network. In *World Congress on Neural Networks – San Diego* (vol. 3, pp. 117-122). Hillsdale, NJ: Lawrence Erlbaum Associates.

Huang, J., Georgiopoulos, M., & Heileman, G. L. (1995). Fuzzy ART Properties. *Neural Networks*, **8**(2), 203-213.

Rubin, M. A. (1994a). Fuzzy ARTMAP automatic target recognition from single and multiple wavelength radar range profiles. In *The 1994 IEEE International Conference on Neural Networks* (vol. 3, pp. 1613-1618). Piscataway, NJ: IEEE Service Center.

Rubin, M. A. (1994b). *Issues in automatic target recognition from radar range profiles using Fuzzy ARTMAP* (Naval Air Warfare Center preprint). China Lake, CA: Naval Air Warfare Center.

Young, D. P. (1995). *A Fuzzy ART MAP neural network speech recognition system based on formant ratios*. Master's thesis, Royal Military College of Canada, Kingston, Ontario.