

1994-02

# Incremental Art: A Neural Network System for Recognition by Incremental Feature Extraction

---

<https://hdl.handle.net/2144/2144>

*"Downloaded from OpenBU. Boston University's institutional repository."*

**INCREMENTAL ART: A NEURAL NETWORK  
SYSTEM FOR RECOGNITION BY  
INCREMENTAL FEATURE EXTRACTION**

J. Mario Aguilar and William D. Ross

February 1994

**Technical Report CAS/CNS-94-006**

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1994

Boston University Center for Adaptive Systems and  
Department of Cognitive and Neural Systems  
111 Cummington Street  
Boston, MA 02215

# Incremental ART: A Neural Network System for Recognition by Incremental Feature Extraction

J. Mario Aguilar<sup>1</sup> and William D. Ross<sup>2</sup>

Center for Adaptive Systems and Cognitive and Neural Systems Dept.  
111 Cummington St. Rm 240 , Boston University, Boston, MA 02215

## Abstract

Incremental ART extends adaptive resonance theory (ART) by incorporating mechanisms for efficient recognition through incremental feature extraction. The system achieves efficient confident prediction through the controlled acquisition of only those features necessary to discriminate an input pattern. These capabilities are achieved through three modifications to the fuzzy ART system: (1) A *partial feature vector complement coding rule* extends fuzzy ART logic to allow recognition based on partial feature vectors. (2) The addition of a  $F_2$  *decision criterion* to measure ART predictive confidence. (3) An *incremental feature extraction layer* computes the next feature to extract based on a measure of predictive value. Our system is demonstrated on a face recognition problem but has general applicability as a machine vision solution and as model for studying scanning patterns.

## Introduction

Classification algorithms including K-means, nearest neighbor, self-organizing feature maps, backpropagation and ART systems (Carpenter and Grossberg, 1987) all require that the entire input or feature vector be presented before confident prediction can be made. This requirement demands that preprocessing of a scene be completed prior to recognition despite that fact that much of that preprocessing may be unnecessary or irrelevant. For typical applications such preprocessing demands feature extraction by either expensive parallel or time-consuming sequential implementations. One common remedy is to determine and extract a set of task specific features off-line. By contrast, human visual recognition acquires scenic features incrementally through the intelligent deployment of the spatially limited foveal resource by automatically determining which scenic features are important for a range of tasks (Rojer and Schwartz, 1992; Seibert and Waxman, 1993). Our model extends fuzzy ART (Carpenter, Grossberg, and Rosen 1991) both as a pattern recognition device and as a model of human recognition through the incorporation of three modifications which interactively allow recognition through incremental feature evaluation and extraction.

Our control strategy results in an efficient system in which the status of the recognition process guides the deployment of acquisition and preprocessing resources offering the potential for real-time affordable operation even in very complex input environments.

The operation of Incremental ART is illustrated on a face recognition problem. In this example, Incremental ART was trained on a set of complete feature vectors extracted from

---

<sup>1</sup>Supported in part by ARPA (ONR N00014-92-J-4015) and the Office of Naval Research (ONR N00014-92-J-1309).

<sup>2</sup>Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0083), the National Science Foundation (NSF IRI 90-00530), and the Office of Naval Research (ONR N00014-91-J-4100).

12 faces taken from the MIT Eigenfaces database (Turk and Pentland 1991). Each complete feature vector consisted of a coarse spatial coding of local boundary activations concatenated with a much longer fine spatial coding of local boundary activations.

Testing was then carried out using the faces corrupted by noise. During testing, only the coarsely coded portion of the feature vector was initially presented. The system then extracted finely coded features incrementally from predictive regions of the face until confident recognition was achieved.

Feature extraction was accomplished by a minimal scheme for discounting the effects of variable illumination and detecting boundaries (Contreras-Vidal and Aguilar, 1993). Gabor filters were used to detect boundaries of 4 orientations, the resulting boundaries were sharpened by techniques inspired by the Boundary Contour System (BCS) (Grossberg and Mingolla 1985a,b), and then orientation activations were summed and normalized within both coarse and fine regions. The details of preprocessing are not relevant since the system is compatible with any preprocessing scheme.

## Incremental ART

Incremental ART is a departure from conventional feedforward approaches in which the system only processes that information which is initially supplied. A more efficient approach is to reduce the amount of information preprocessed at any one time provided that the system can actively control from which areas in the environment it will next extract features. This sequential acquisition process allows for a reduction in the complexity of initial preprocessing and for confident recognition based on minimal feature extraction.

Figure 1 presents a diagram of the Incremental ART system. The Incremental ART system dynamically guides the feature extraction process to accomplish face recognition. Algorithmically, the operation of the system can be summarized as a three step process in which each step incorporates one of the extensions to fuzzy ART. The three steps repeat until the decision criterion is met and confident recognition occurs.

Step 1: The incrementally completed feature vector is input to the  $F_1$  layer. This vector activates the  $F_2$  layer according to the fuzzy ART activation equation augmented by the *partial feature vector complement coding rule*. Step 2: The distributed pattern of  $F_2$  activation is then contrast enhanced and a winner is chosen if a *decision criterion* is met. Step 3: Otherwise the *incremental feature extraction* layer determines which feature to next extract.

### Partial feature vector complement coding rule

In order for a partial feature vector to activate  $F_2$  categories in a way that accurately reflects the partial information represented some distinction must be made between the absence of feature due to the fact that it has not yet been extracted and the absence of a feature due to an actual absence in the scene.

The complement coding preprocessing strategy provides a means of expressing this distinction. In complement code, each feature is represented by both on and off channels so that the on channel codes a measure of the feature's presence and the off channel codes a measure of the feature's absence (Carpenter, Grossberg and Rosen 1991). Thus, in both categorization and learning both the presence and the absence of a feature can be predictive.

The partial feature vector complement coding rule states that in a complement coded feature vector, the absence of a feature due to an actual absence in the scene (after extraction) is coded as a 0 in the on channel and a 1 in the off channel, whereas the absence of a feature due to the fact that it has not yet been extracted is coded as a 0 in the on channel and a 0 in the off channel.

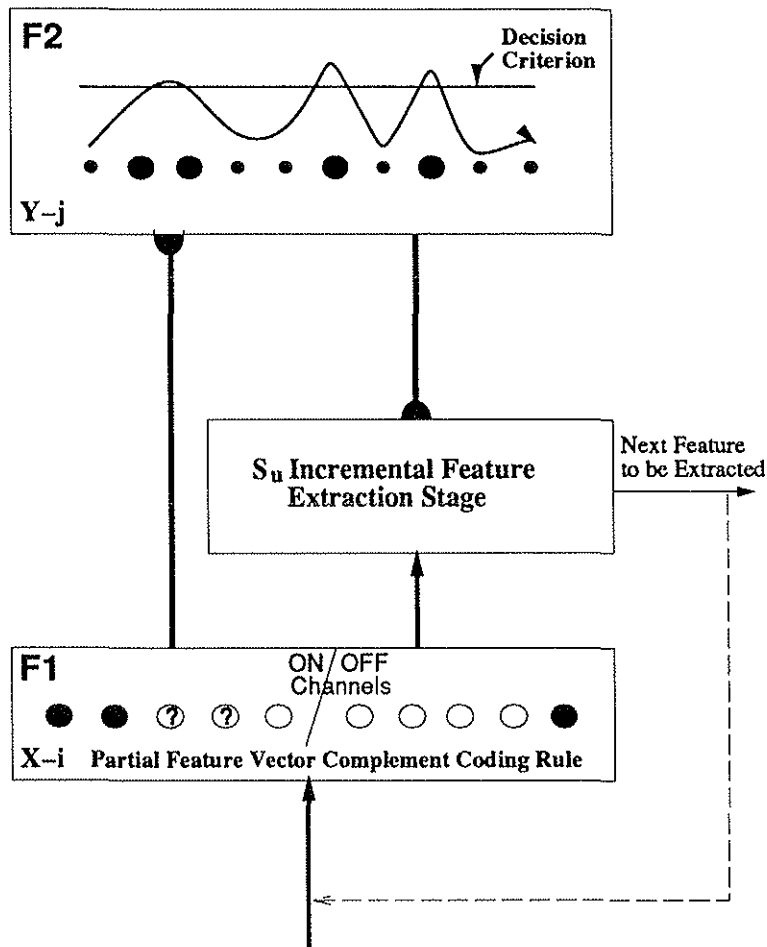


Figure 1: System diagram of incremental ART.

This rule allows informative activation of recognition categories at  $F_2$  by a partial feature vector. Since in the current implementation, learning is prevented when only a partial feature vector is present, the rule does not result in spurious decrease in weights for features that have not yet been extracted. However, a given partial feature vector may not yield an unambiguous winner among  $F_2$  categories.

### Decision criterion

An alternative to the standard ART  $F_2$  winner-take-all choice rule uses a *decision criterion* (DC) at the field  $F_2$ . The decision criterion is introduced in ART-EMAP to achieve efficient 3-D object recognition (Carpenter, and Ross 1993a,1993b). The decision criterion permits ART choice only when the most active  $F_2$  category  $J$  becomes a minimum proportion more active than the next most active  $F_2$  category. Thus if  $z_j$  codes system prediction

and  $y_j$  codes  $F_2$  node activation:

$$z_J = \begin{cases} 1 & \text{if } y_J > (DC)y_j \text{ for all } j \neq J \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where  $DC \geq 1$ . When  $DC = 1$ , the decision criterion rule reduces to the  $F_2$  winner-take-all choice rule used in previous ART systems. When  $DC > 1$ , the decision criterion prevents prediction in cases in which multiple  $F_2$  categories are about equally activated, representing ambiguous predictive evidence. For computational convenience, activity at  $F_2$  can be contrast enhanced prior to application of the DC by a normalized power rule:

$$y_j = \frac{(y_j)^q}{\sum_{n=1}^{N_{F_2}} (y_n)^q}. \quad (2)$$

When the decision criterion fails more features must be extracted to disambiguate.

#### Incremental feature extraction

The determination of which features of the input environment, in this example a face, to process next to best disambiguate between potential recognition categories requires some measure of the discriminating power of one portion of the feature vector versus others. The variation across the weights between a single feature and likely categories provides a good indication of the predictive value of extracting that feature, since it is a measure of the variation in  $y_j$  activation that would result. Therefore, in this implementation, we use,  $S_u$ , where  $u$  is the index of an unaccessed  $F_1$  feature, to be the sum of the biased differences between the template weights and the template weight of the maximally activated category. Each difference is biased according to the likelihood of that category as given by  $y_j$ . Thus, if the most active  $F_2$  category is  $J$ , the  $F_1$  features are  $x_i$ , and the  $F_1 \rightarrow F_2$  weights are  $w_{ij}$  then the feature to be extracted next,  $U$ , will be given when;

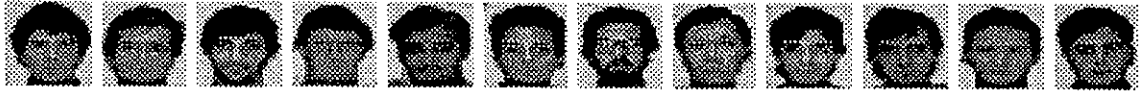
$$S_U > S_u \text{ for all } u \neq U \quad (3)$$

where,

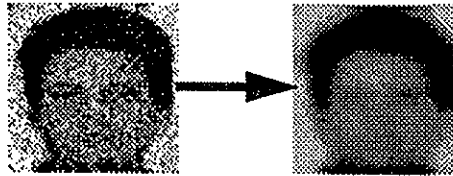
$$S_u = \sum_{j=1}^{N_{F_2}} y_j |w_{uj} - w_{uJ}|. \quad (4)$$

This incremental feature extraction layer determines each unaccessed features predictive value dynamically since after each extraction  $y_j$  changes. It should also be noted that the

Training set

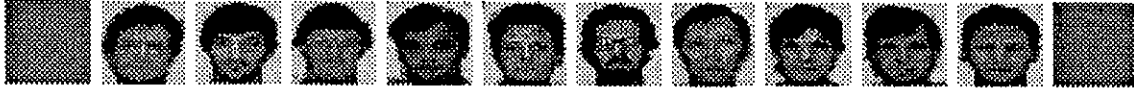


Noisy Test Image

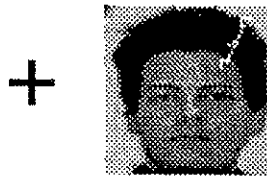


Coarse Coded Version  
of face #6

Candidate categories



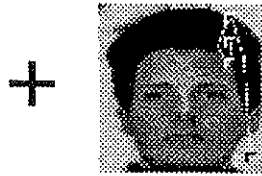
Saccade sequence 1



Candidate categories



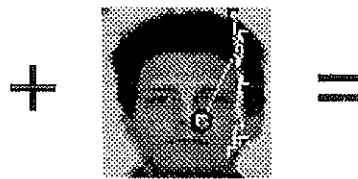
Saccade sequence 2



Candidate categories



Saccade sequence 3



Chosen category



Figure 2. Simulation of recognition sequence for noisy version of face 6. The system is trained with the faces shown in the first row. The sequence of feature as extracted by the system are shown after each new reduction in the number of candidate categories. Clean images are used to show the scanning patterns. The actual test input was noisy as shown.

value of  $q$  in the normalized power rule (2) scales the degree of bias according to category likelihood.

### Simulation results

In practice, we found that using  $q = 4$ , and  $DC = 1.25$  yielded perfect performance on the noise-corrupted face test set while requiring a minimal number of incremental feature extractions, on average 10% (15) of the finely coded features. Figure 2 illustrates performance for one of the test faces. The sequential extraction of predictive features rules out candidate face categories gradually as the marked scanning pattern emerges.

### Conclusion

A new neural network architecture is presented which extends fuzzy ART to accomplish efficient recognition through intelligent incremental feature extraction. Incremental ART is demonstrated to offer computational savings without loss of predictive accuracy while offering a model of saccadic scanning pattern.

### References

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.

G.A. Carpenter, S. Grossberg, and D.B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, Vol. 4, pp. 759-771, 1991.

G.A. Carpenter and W.D. Ross, "ART-EMAP: A neural network architecture for learning and prediction by evidence accumulation," Proceedings of the World Congress on Neural Networks (WCNN-93), III-649-656. Technical Report CAS/CNS-TR-93-015, Boston, MA: Boston University, 1993a.

G.A. Carpenter and W.D. Ross, "ART-EMAP: A neural network architecture for object recognition by evidence accumulation," Technical Report CAS/CNS-TR-93-035, Boston, MA: Boston University, 1993b.

J.L. Contreras-Vidal and M. Aguilar, "A fast BCS/FCS algorithm for image segmentation." In *Proceedings of the International Conference on Artificial Neural Networks (ICANN-93)*, Amsterdam, September 13-16, 1993.

S. Grossberg and E. Mingolla, "Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading," *Psychological Review*, Vol. 92, pp. 173-211, 1985.

A.S. Rojer and Schwartz, "A quotient space hough transform for space-variant visual attention," *Neural networks for vision and image processing*, pp. 408-436, 1992.

M. Seibert and A.M. Waxman, "An Approach to Face Recognition Using Saliency Maps and Caricatures," In Proceedings of the World Congress on Neural Networks (WCNN-93), III-661-664. M.A. Turk and A.P. Pentland, "Face Recognition using Eigenfaces," *Journal of Cognitive Neuroscience*, March 1991.