INVARIANT PATTERN RECOGNITION AND RECALL BY AN ATTENTIVE SELF-ORGANIZING ART ARCHITECTURE IN A NONSTATIONARY WORLD

Gail A. Carpenter*+ and Stephen Grossberg*

Ctr. Adaptive Systems (*)
Boston University
111 Cummington Street
Boston, MA 02215

Department of Mathematics (+)
Northeastern University
360 Huntington Avenue
Boston, MA 02115

1. Introduction

A neural network architecture is described which can

(1) stably self-organize an invariant pattern recognition code in response to a sequence of analog or digital input patterns;

(2) be attentionally primed to ignore all but a designated category of input patterns;

(3) automatically shift its prime as it satisfies internal criteria in response to the occurrence of a previously primed category of input patterns;

(4) learn to generate an arbitrary spatiotemporal output pattern in response to any input pattern exemplar of an activated recognition category.

This architecture (Figure 1) exploits properties of the ART 1 and ART 2 adaptive resonance theory architectures which have been developed in Carpenter and Grossberg (1985, 1987a, 1987b, this volume); the Boundary Contour System for boundary segmentation and the Feature Contour System for figural filling-in which have been developed in Cohen and Grossberg (1984), Grossberg (1987a, 1987b), Grossberg and Mingolla (1985a, 1985b, 1987a, this volume), and Grossberg and Todorović (1987a, this volume); theorems on associative pattern learning and associative map learning (Grossberg, 1969, 1970, 1982); and circuit designs to focus attention on desired goal objects by using learned feedback interactions between external sensory events and internal homeostatic events (Grossberg, 1972, 1982, 1987c; Grossberg and Levine, 1987a, this volume; Grossberg and Schmajuk, 1987a, this volume). The overall circuit design embodies, in a primitive way, an intentional learning machine in which distinct cognitive, homeostatic, and motor representations are self-organized in a coordinated fashion.

2. Outline of the Architecture

The ART 2 architecture stably self-organizes disjoint recognition categories in response to temporal sequences of analog or digital input patterns. A vigilance parameter can be set to determine the coarseness of this classification, thereby compensating for source of variability, including noise, in the input exemplars of the emergent recognition categories. These input patterns may be the output patterns of a preprocessor stage;

This research was supported in part by the Air Force Office of Scientific Research (AFOSR F49620-86-C-0037), the Army Research Office (ARO DAAG-29-85-K-0095), and the National Science Foundation (NSF DMS-86-11959 (G.A.C.) and NSF IRI-84-17756 (S.G.)).

Acknowledgements: We wish to thank Cynthia Suchta and Carol Yanakakis for their valuable assistance in the preparation of the manuscript and illustrations.

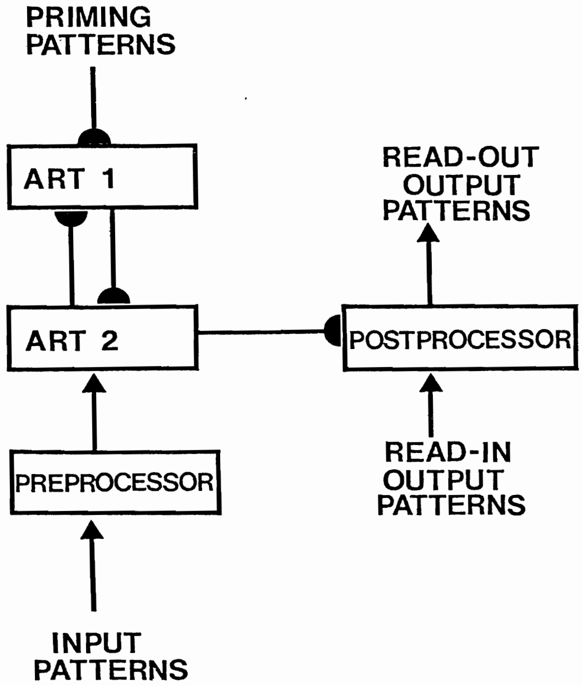

Figure 1. Outline of a self-organizing invariant pattern recognition and recall architecture.

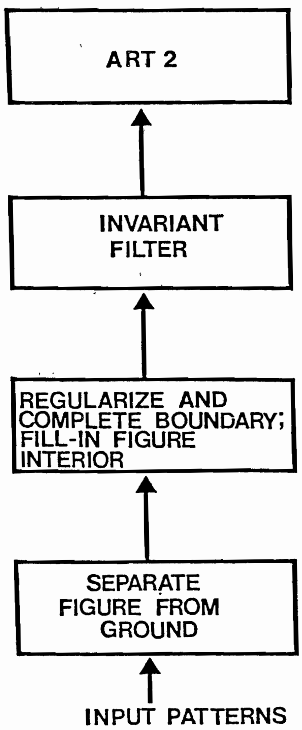

Figure 2. A possible preprocessor for the architecture.

in particular, the preprocessor outputs may represent invariant spectra computed by one or more prior stages of input filtering (Figure 2). The capacity of ART 2 to be activated by any number of arbitrarily chosen analog or digital input patterns without destabilizing its emergent classification provides great freedom in designing such preprocessors for specialized applications.

In the present application, the preprocessor has been designed by using available techniques to present ART 2 with spectra that are invariant under 2-D spatial translation, dilation, and rotation (Casasent and Psaltis, 1976; Cavanagh, 1978, 1984; Szu, 1986). Using an invariant input filter such as a Fourier-Mellon filter is a familiar and direct approach to achieving 2-D spatial invariance in pattern recognition. Alternative approaches use several hierarchically organized processing stages to gradually free image processing from its spatial coordinates (Fukushima, 1980; Fukushima and Miyake, 1984; Grossberg, 1978, Section 19; reprinted in Grossberg, 1982); higher-order threshold logic units (Maxwell, Giles, and Chen, 1986); or match-gated associative mechanisms (Grossberg, 1987d; Grossberg and Kuperstein, 1986, Chapter 6).

Before the input pattern is processed by the invariant filter, the image figure to be recognized must be detached form the image background. This can be accomplished, for example, by using a range detector focussed at the distance of the figure, and detaching the figure from contiguous ground by spatially intersecting the range pattern with a pattern from another detector that is capable of differentiating figure from ground. A doppler image can be intersected when the figure is moving. The intensity of laser return can be intersected when the figure is stationary (Gschwendtner, Harney, and Hull, 1983; Harney, 1980, 1981; Harney and Hull, 1980; Hull and Marcus, 1980; Kolodzy, this volume; Sullivan, 1980, Sullivan, 1980, 1981; Sullivan, Harney, and Martin, 1979).

If the figure derived in this way is noisy and irregular, the vigilance parameter may be set at a value that adjusts for the expected level of noise fluctuations generated by the imaging devices. In addition, as in Figure 2, the figure boundary may be extracted, completed, and regularized using the emergent boundary segmentation process of a Boundary Contour System (Grossberg and Mingolla, 1985a, 1985b, 1987a, this volume; Grossberg, 1987a, 1987b). Or some portion of the Boundary Contour System can be used, such as its second competitive stage, which can choose the maximally activated oriented filter at each spatial location and inhibit responses from statistically unoriented regions. This completed boundary can then serve as the input pattern to ART 2 (Figure 3). The figure interior within the emergent segmentation may also be smoothly completed using the filling-in process of a Feature Contour System (Cohen and Grossberg, 1984; Grossberg and Todorović, 1987a, this volume), and this filled-in representation used as an input source to ART 2. Combinations of the completed boundary and filled-in figural representations may also be chosen as the inputs to ART 2, thereby fully exploiting the architecture's self-scaling properties.

Figures 4–6 describe the results of preliminary computer simulations of noisy boundary images by ART 2. Figure 4 describes correct classification of 40 noisy exemplars of 4 trucks into 4 categories. Pairs of trucks were chosen to be very similar to one another. Figures 5 and 6 describe simulations of invariant pattern recognition. Here trucks were translated, rotated, and shrunk with respect to each of 4 prototypes, and then subjected to noise. Figure 5 describes a consistent classification in 5% noise of 32 exemplars (8 from each prototype) into 5 consistent categories. Exemplars of one truck type were split into two categories (3,5). Figure 6 describes a consistent classification in 10% noise of 32 exemplars into 7 consistent categories. Exemplars of one truck type were split into four categories (3,4,5,7). These results were derived in our initial runs using very coarse boundary filtering techniques and preliminary choice of parameters.

Even at this preliminary stage, the postprocessor mechanism described below can be used to quickly learn a compression into 4 category outputs. Inspection of the last three rows of Figure 5 and the last five rows of Figure 6 illustrates how well the architecture has managed to separate subtle differences, obscured by image transformation and noise, of this pair of trucks.

As ART 2 self-organizes recognition categories in response to these preprocessed inputs, its categorical choices at the F_2 classifying level self-stabilize through time. In examples wherein F_2 makes a choice, it can be used as the first level of an ART 1 architecture, or yet another ART 2 architecture, should one prefer. Let us call the classifying level of this latter architecture F_3. Level F_3 can be used as a source of pre-wired priming inputs to F_2. Alternatively, self-stabilizing choices by F_3 can quickly be learned in response to the choices made at F_2. Then F_3 can be used as a source of self-organized priming inputs to F_2, and a source of priming patterns can be associated with each of the F_3 choices via mechanisms of associative pattern learning (Grossberg, 1969, 1982). For this to work well, the normalizing property of F_3 is important. After learning of these primes takes place, turning on a particular prime can activate a learned $F_3 \rightarrow F_2$ top-down expectation. Then F_2 can be supraliminally activated only by an input exemplar which is a member of the recognition category of the primed F_2 node. The architecture ignores all but the primed set of input patterns. In other words, the prime causes the architecture to pay attention only to expected sources of input information. Due to the spatial invariance properties of the preprocessor, the expected input patterns can be translated, dilated, or rotated in 2-D without damaging recognition. Due to the similarity grouping properties of ART 2 at a fixed level of vigilance, suitable deformations of these input patterns, including deformations due to no more than the anticipated levels of noise, can also be recognized.

The output pathways from level F_2 of ART 2 to the postprocessor (Figure 1) can learn to read out any spatial pattern or spatiotemporal pattern of outputs by applying theorems about associative spatial pattern learning in avalanche-type circuits (Grossberg, 1969, 1970, 1982). Thus the architecture as a whole can stably self-organize an invariant recognition code and an associative map to an arbitrary format of output patterns.

The model of priming patterns can be both modified and refined. The interactions (priming $\rightarrow$ ART) and (ART $\rightarrow$ postprocessor) can be modified so that output patterns are read-out only if the input patterns have yielded rewards in the past and if the machine's internal needs for these rewards have not yet been satisfied (Grossberg, 1972, 1982, 1987c; Grossberg and Levine, 1987a, 1987b; Grossberg and Schmajuk, 1987a, 1987b). In this variation of the architecture, the priming patterns supply motivational signals for releasing outputs only if an input exemplar from a desired recognition category is detected.

References

Carpenter, G.A. and Grossberg, S., Category learning and adaptive pattern recognition: A neural network model. *Proceedings of the Third Army Conference on Applied Mathematics and Computing*, 1985, ARO Report 86-1, 37–56.

Carpenter, G.A. and Grossberg, S., A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, 1987, **37**, 54–115 (a).

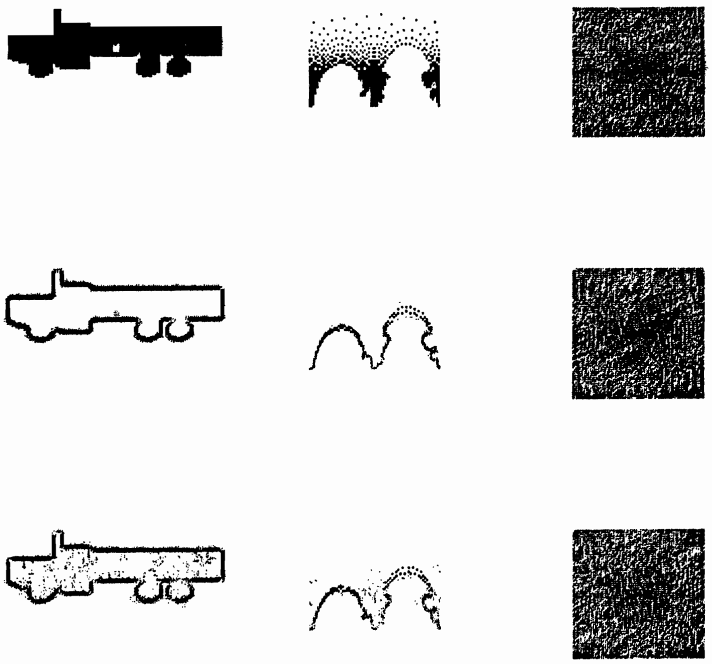

Figure 3. The left column depicts input images, the middle column log-polar maps of the image, and the third column Fourier transform amplitudes of the maps, which input to ART 2. The first image in row 1 is transformed into that in row 2 by choosing the maximally activated oriented mask at each position. The first image in row 3 is derived in the same way from a noisy version of the original figure.

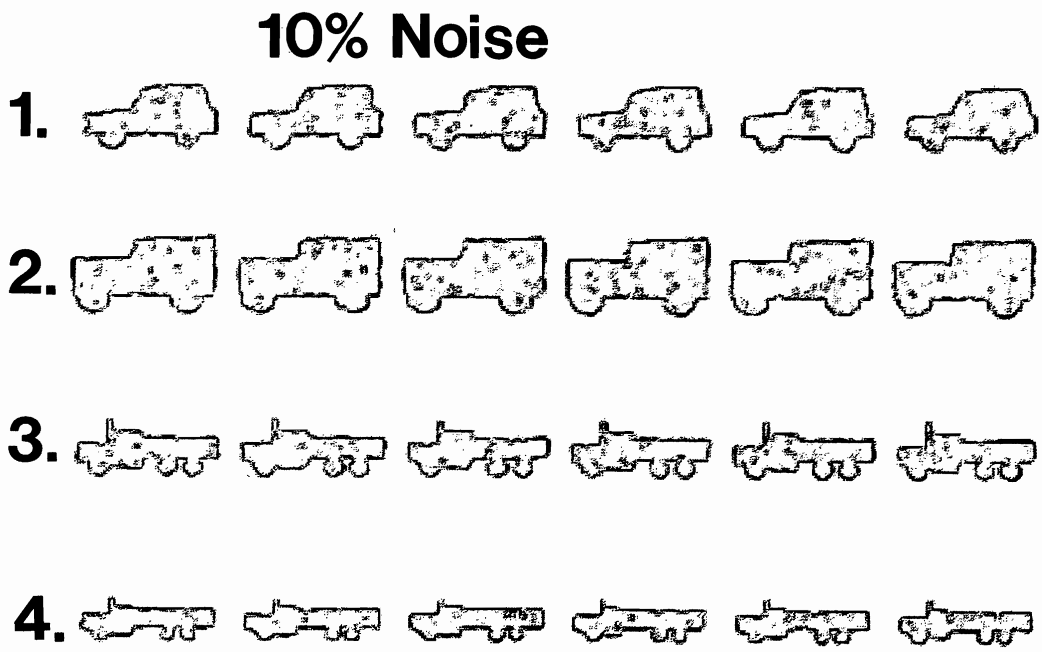

Figure 4. Classification of 40 noisy exemplars of 4 trucks by ART 2 into 4 correct categories. Noise level was 10%. (Six of the 10 exemplars in each category are displayed.)

Carpenter, G.A. and Grossberg, S., ART 2: Stable self-organization of pattern recognition codes for analog input patterns. *Applied Optics*, special issue on neural networks, in press, 1987 (b).

Carpenter, G.A. and Grossberg, S., ART 2: Self-organization of stable category recognition codes for analog input patterns. This volume, 1987 (c).

Casasent, D. and Psaltis, D., Position, rotation, and scale invariant optical correlations. *Applied Optics*, 1976, **15**, 1793–1799.

Cavanagh, P., Size and position invariance in the visual system. *Perception*, 1978, **7**, 167–177.

Cavanagh, P., Image transforms in the visual system. In P.C. Dodwell and T. Caelli (Eds.), **Figural synthesis**. Hillsdale, NJ: Erlbaum, 1984.

Cohen, M.A. and Grossberg, S., Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, 1984, **36**, 428–456.

Fukushima, K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by a shift in position. *Biological Cybernetics*, 1980, **36**, 193–202.

Fukushima, K. and Miyake, S., Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, 1984, **15**, 455–469.

Grossberg, S., On learning and energy-entropy dependent in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, 1969, **1**, 319–350.

Grossberg, S., Some networks that can learn, remember, and reproduce any number of complicated space-time patterns II. *Studies in Applied Mathematics*, 1970, **49**, 135–166.

Grossberg, S., A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences*, 1972, **15**, 253–285.

Grossberg, S., A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.), **Progress in theoretical biology**, Vol. 5. New York: Academic Press, 1978.

Grossberg, S., **Studies of mind and brain: Neural principles of learning, perceptionn, development, cognition, and motor control**. Boston: Reidel Press, 1982.

Grossberg, S., Cortical dynamics of three-dimensional form, color, and brightness perception: I: Monocular theory. *Perception and Psychophysics*, 1987, **41**, 87–116 (a).

Grossberg, S., Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory. *Perception and Psychophysics*, 1987, **41**, 117–158 (b).

Grossberg, S. (Ed.), **The adaptive brain, I: Cognition, learning, reinforcement, and rhythm**. Amsterdam: Elsevier/North-Holland, 1987 (c).

Grossberg, S., Nonlinear neural networks: Principles, mechanisms, and architectures. *Applied Optics*, in press, 1987 (d).

Grossberg, S. and Kuperstein, M., **Neural dynamics of adaptive sensory-motor control: Ballistic eye movements**. Amsterdam: Elsevier/North-Holland, 1986.

Grossberg, S. and Levine, D.S., Neural dynamics of attentionally-modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics*, in press, 1987 (a).

Grossberg, S. and Levine, D.S., Attentional mechanisms in neural information processing: Examples from Pavlovian conditioning. This volume, 1987 (b).

Grossberg, S. and Mingolla, E., Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, 1985, **92**, 173–211 (a).

Grossberg, S. and Mingolla, E., Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, 1985, **38**, 141–171 (b).

Grossberg, S. and Mingolla, E., Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and Image Processing*, 1987, **37**, 116–165 (a).

Grossberg, S. and Mingolla, E., A neural network architecture for preattentive vision: Multiple scale segmentation and regularization. This volume, 1987 (b).

Grossberg, S. and Schmajuk, N.A., A neural network architecture for attentionally-modulated Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. Submitted for publication, 1987 (a).

Grossberg, S. and Schmajuk, N.A., A neural network architecture for attentionally-modulated Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. This volume, 1987 (b).

Grossberg, S. and Todorović, D., Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. Submitted for publication, 1987 (a).

Grossberg, S. and Todorović, D., A neural network architecture for brightness perception under constant and variable illumination conditions. This volume, 1987 (b).

Gschwendtner, A.B., Harney, R.C., and Hull,, R.J., Coherent IR radar technology. In D.K. Killinger and A. Mooradian (Eds.), **Optical and lasar remote sensing**. New York: Springer-Verlag, 1983.

Harney, R.C., Infrared airborne radar. *Proceedings of the IEEE 1980 Electronic and Aerospace Systems Conference*, 1980, 462–471.

Harney, R.C., Military applications of coherent infrared radar. In **Physics and technology of coherent infrared radar**. Proceedings of the SPIE, **300**, 2–11, 1981.

Harney, R.C. and Hull, R.J., Compact infrared radar technology. In **CO_2 lasar devices and applications**. Proceedings of the SPIE, **227**, 162–170, 1980.

Hull, R.J. and Marcus, S., A tactical 10.6 micrometer imaging radar. *Proceedings of the IEEE 1980 Electronic and Aerospace Systems Conference*, 1980, **2**, 662–668.

Kolodzy, P., Multidimensional machine vision using neural networks. This volume, 1987.

Maxwell, T., Giles, G.L., and Lee, Y.C., Nonlinear dynamics of artificial neural systems. In J.S. Denker (Ed.), **Neural networks for computing**. New York: American Institute of Physics, 1986.

Sullivan, D.R., Active 10.6 micrometer image processing. **Image processing and missile guidance**. Proceedings of the SPIE, **238**, 103–117, 1980.

Sullivan, D.R., Harney, R.C., and Martin, J.B., Real-time quasi-3-dimensional display of infrared radar images. **Real-time signal processing II**. Proceedings of the SPIE, **180**, 57–65, 1979.

Sullivan, L.J., Infrared coherent radar. In **CO_2 lasar devices and applications**. Proceedings of the SPIE, **227**, 148 161, 1980.

Sullivan, L.J., Firepond lasar radar. *Electro/81 Conference Record*, Session 34, 1981.

Szu, H., Holographic coordinate transformations and optical computing. In H. Szu (Ed.), **Optical and hybrid computing**. Bellingham, WA: International Society for Optical Engineering, SPIE Volume 634, 1986.