

**REPORT DOCUMENTATION PAGE**

Form Approved  
OMB No. 0704-0188

**AD-A243 052**



cost to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering the collection of information, Send comments regarding this burden estimate or any other aspect of this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

2

1. REPORT DATE: ANNUAL 15 Mar 90 to 14 Mar 91

4. TITLE AND SUBTITLE  
THE COGNITIVE, PERCEPTUAL, AND NEURAL BASES OF SKILLED PERFORMANCE

5. FUNDING NUMBERS  
GR - AFOSR-90-0175  
PE - 61103D  
PR - 3484  
TA - A4

6. AUTHOR(S)  
Dr Stephen B Grossberg

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  
Boston University  
College of Liberal Arts  
111 Cummington Street, Second Floor  
Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER  
AFOSR TR- 91-0013

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  
Dr John F. Tangney  
AFOSR/NL  
Building 410  
Bolling AFB, DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT  
Approved for public release;  
distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)  
This report reviews progress from the Boston University, Northeastern University, and Harvard University/Cambridge University research groups of our AFOSR University Research Initiative grant. The report lists books and articles, summaries of research, and selected abstracts of key articles. The report also encloses the program (see next page) and the abstract book of an AFOSR-supported conference that was held at the Wang Institute of Boston University on May 11-13, 1990. The topic of the conference was Neural Networks for Automatic Target Recognition. There were fifteen invited speakers and thirty-eight contributed posters. Three hundred scientists and students attended from twenty countries and thirty-five states of the United States.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT  
UNCLASSIFIED

18. SECURITY CLASSIFICATION OF THIS PAGE  
UNCLASSIFIED

19. SECURITY CLASSIFICATION OF ABSTRACT  
UNCLASSIFIED

20. LIMITATION OF ABSTRACT  
UNLIMITED

**INTERIM SCIENTIFIC REPORT**


**Contract AFOSR 90-0175**

**THE COGNITIVE, PERCEPTUAL, AND NEURAL BASES  
OF SKILLED PERFORMANCE**

**March 15, 1990—March 14, 1991**

**Principal Investigator: Stephen Grossberg  
Wang Professor of Cognitive and Neural Systems  
Professor of Mathematics, Psychology,  
and Biomedical Engineering  
Director, Center for Adaptive Systems  
Chairman, Department of Cognitive and Neural Systems  
Boston University  
111 Cummington Street  
Boston, MA 02215  
(617) 353-7857**

Approved For	
Distribution	
By	
Distribution	
Availability Codes	
Avail and/or	
Dist	Special
A-1	

**91-16524**  


**91 1126 012**

## TABLE OF CONTENTS

<b>Project Summaries</b>	1
<b>Publications Partially Supported by AFOSR (Boston University)</b>	2
<b>Boston University Project Summaries</b>	5
1. Cortical Dynamics of Motion Perception	5
2. The Organization of Cortical Systems for Form and Motion Perception	7
3. Emergent Segmentation of Moving Images	8
4. Psychophysical Studies of Motion Segmentation	8
5. Synchronized Oscillations during Cooperative Feature Linking in in Visual Cortex	9
6. Automatic Figure-Ground Separation of Connected Scenic Components	9
7. An Improved Boundary Segmentation Network for Processing of of Static Images	10
8. Autonomous Learning, Pattern Recognition, and Prediction	10
9. Vector Associative Maps: Self-Organizing Spatial Representations and Motor Controllers	11
<b>Selected Abstracts</b>	13
<b>Northeastern University Project Summaries</b>	25
Color Appearance and Color Mechanisms	25
Publications and Conference Presentations	26
<b>Harvard University/Cambridge University Project Summaries</b>	28
(a) Relaxation Computation of Non-Orthogonal Image Transforms	28
(b) Self-Similar 2-D Gabor Wavelet Representations	29
(c) Neural Mechanisms for Interpreting the Dynamic Visual World: Figure/Ground Segregation Based on Differential Motion Cues	30
(d) Interpretation of Differential Visual Motion: Experimental and Theoretical Issues	31

## PROJECT SUMMARIES

This report reviews progress from the Boston University, Northeastern University, and Harvard University/Cambridge University research groups of our AFOSR University Research Initiative grant. The report lists books and articles, summaries of research, and selected abstracts of key articles.

The report also encloses the program (see next page) and the abstract book of an AFOSR-supported conference that was held at the Wang Institute of Boston University on May 11-13, 1990. The topic of the conference was Neural Networks for Automatic Target Recognition. There were fifteen invited speakers and thirty-eight contributed posters. Three hundred scientists and students attended from twenty countries and thirty-five states of the United States.

## PROGRAM

May 11, 1990

REGISTRATION 1:00PM-5:00PM

RECEPTION 3:00PM-5:00PM

EVENING SESSION 5:00PM-7:30PM

- 5:00-5:50PM: Dr. Barbara L. Yoon, DARPA  
"The DARPA Artificial Neural Network Technology Program"
- 5:50-6:40PM: Dr. Joe R. Brown, Martin Marietta  
"Multi-Sensor ATR Using Neural Nets"
- 6:40-7:30PM: Dr. Robert Hecht-Nielsen, HNC  
"Spatiotemporal Pattern Recognizer Neural Network"
- 

May 12, 1990

MORNING SESSION 8:00AM-11:50AM

- 8:00-8:50AM: Professor Alex Waibel, Carnegie Mellon University  
"Patterns, Sequences, and Variability:  
Advances in Connectionist Speech Recognition"
- 8:50-9:40AM: Dr. Christopher L. Scofield, Nestor Inc. & Brown University  
"Neural Network Automatic Target Recognition by Active  
and Passive Sonar Signals"
- 9:40-10:10AM: Coffee Break
- 10:10-11:00AM: Professor Stephen Grossberg, Boston University  
"A Neural Network Theory of 3-D Vision and Figure-Ground  
Separation"
- 11:00-11:50AM: Professor Gail Carpenter, Boston University  
"Target Recognition by Adaptive Resonance: ART for ATR"
- 11:50AM-1:00PM: Lunch

AFTERNOON SESSION 1:00PM-7:00PM

- 1:00-1:50PM: Dr. Ken Johnson, Rockwell International  
"The Application of Neural Networks to the Acquisition and Tracking  
of Maneuvering Tactical Targets in High Clutter IR Imagery"
- 1:50-2:40PM: Dr. Allen M. Waxman, MIT Lincoln Lab  
"Invariant Learning and Recognition of 3-D Objects  
from Temporal View Sequences"
- 2:40-3:30PM: Dr. Paul Kolodzy, MIT Lincoln Lab  
"A Multi-Dimensional ATR System"
- 3:30-4:00PM: Coffee Break
- 4:00-7:00PM: Poster Session
- 7:00-8:30PM: Dinner at Wang Institute
-

May 13, 1990

**MORNING SESSION 8:00AM-1:00PM**

- 8:00-8:50AM: **Professor Nabil H. Farhat, University of Pennsylvania**  
"Bifurcating Networks for Target Recognition"
- 8:50-9:40AM: **Dr. William Thoet, Booz-Allen & Hamilton**  
"Current Status and Results of Three Government Programs  
in Neural Network-Based ATR"
- 9:40-10:10AM: **Coffee Break**
- 10:10-11:00AM: **Dr. Michael Kuperstein, Neurogen**  
"INFANT Neural Controller for Adaptive Sensory-Motor Coordination  
and Dynamic Movements"
- 11:00-11:50AM: **Dr. Yann LeCun, AT&T Laboratories**  
"Structured Back Propagation Networks for Handwriting Recognition"
- 11:50-12:40PM: **Dr. Stephen D. Simmes, Science Applications International**  
"Massively Parallel Approaches to Automatic Target Recognition"
- 12:40-1:00PM: **Discussion**
-

PUBLICATIONS PARTIALLY SUPPORTED BY  
THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

CENTER FOR ADAPTIVE SYSTEMS  
BOSTON UNIVERSITY

CONTRACT AFOSR 90-0175

MARCH 15, 1990-MARCH 14, 1991

BOOKS

1. Carpenter, G.A. and Grossberg, S. (Eds.) (1991). **Pattern recognition by self-organizing neural networks**. Cambridge, MA: MIT Press, in preparation. (\*%#+)

ARTICLES

1. Bullock, D. and Grossberg, S. (1991). Adaptive neural networks for control of movement trajectories invariant under speed and force rescaling. *Human Movement Science*, 10, 3-53. (+)
2. Bullock, D. and Grossberg, S. (1991). Emergence of tri-phasic muscle activation from the nonlinear interactions of central and spinal neural network circuits. *Human Movement Science*, in press. (# +)
3. Carpenter, G.A. and Grossberg, S. (1991). Self-organizing cortical networks for distributed hypothesis testing and recognition learning. In J. G. Taylor and C. L. T. Mannion (Eds.), **Theory and applications of neural networks**. London: Springer-Verlag. (\*%#+)
4. Carpenter, G.A. and Grossberg, S. (1991). Distributed hypothesis testing, attention shifts, and transmitter dynamics during the self-organization of brain recognition codes. In H. G. Schuster and W. Singer (Eds.), **Nonlinear dynamics and neuronal networks**. New York: Springer-Verlag. (\*%#+)
5. Carpenter, G.A. and Grossberg, S. (1991). Adaptive resonance theory. **Encyclopedia of artificial intelligence, Second edition**. New York: Wiley and Sons. (%\*#+)
6. Carpenter, G.A., Grossberg, S., and Reynolds, J. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, in press. (\*%#+)
7. Carpenter, G.A., Grossberg, S., and Reynolds, J. (1991). A self-organizing ARTMAP neural architecture for supervised learning and pattern recognition. In R. Mammone and Y. Zeevi (Eds.), **Neural networks: Theory and applications**. New York: Academic Press. (\*%#+)
8. Carpenter, G.A., Grossberg, S., and Reynolds, J. (1991). A neural network architecture for fast on-line supervised learning and pattern recognition. In H. Wechsler (Ed.), **Neural networks for human and machine perception**. New York: Academic Press. (\*%#+)
9. Carpenter, G.A., Grossberg, S., and Reynolds, J. (1991). ARTMAP: A self-organizing neural network architecture for fast supervised learning and pattern recognition. In O. Simula (Ed.), **Proceedings of the international conference on artificial neural networks**. Amsterdam: Elsevier. (\*%#+)
10. Carpenter, G.A., Grossberg, S., and Rosen, D.B. (1991). ART2-A: An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks*, in press. (\*%#+)
11. Cruthirds, D., Gove, A., Grossberg, S., and Mingolla, E. (1991). Preattentive texture segmentation and grouping by the Boundary Contour System. To appear in **Proceedings**

- of the international joint conference on neural networks, Seattle, Washington, July. (#=)
12. Gaudiano, P. and Grossberg, S. (1991). Adaptive vector integration to end point: Self-organizing neural circuits for control of planned movement trajectories. *Human Movement Science*, in press. (#+)
  13. Gaudiano, P. and Grossberg, S. (1991). Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks*, 4, 147-183. (#+)
  14. Gaudiano, P. and Grossberg, S. (1991). Self-organization of spatial representations and arm trajectory controllers by vector associative maps energized by cyclic random generators. In A. Babloyantz (Ed.), **Self-organization, emerging properties and learning**. London: Plenum Press. (#+)
  15. Grossberg, S. (1991). Why do parallel cortical systems exist for the perception of static form and moving form? *Perception and Psychophysics*, 49, 117-141. (\*#=)
  16. Grossberg, S. (1991). The symmetric organization of parallel cortical systems for form and motion perception. In H. Wechsler (Ed.), **Neural networks for human and machine perception**. New York: Academic Press. (#=)
  17. Grossberg, S. and Mingolla, E. (1990). Neural dynamics of motion segmentation. Invited article in **Proceedings of graphics interface/vision interface 90**, Halifax, Nova Scotia, May 17, 1990. Toronto: Canadian Information Processing Society, 112-119. (\*)
  18. Grossberg, S. and Mingolla, E. (1990). Neural dynamics of motion segmentation: Direction fields, apertures, and resonant grouping. In M. Caudill (Ed.), **Proceedings of the international joint conference on neural networks**, January, I, 11-14. Hillsdale, NJ: Erlbaum Associates. (\*#)
  19. Grossberg, S. and Mingolla, E. (1991). Neural dynamics of motion segmentation: Direction fields, apertures, and resonant grouping. Submitted for publication. (\*#)
  20. Grossberg, S. and Rudd, M.E. (1991). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, in press. (\*#=)
  21. Grossberg, S. and Rudd, M.E. (1991). A neural theory of visual motion perception. In B. Blum (Ed.), **Channels in the visual nervous system: Neurophysiology, psychophysics, and models**. Tel Aviv: Freund Publishing Ltd. (\*#)
  22. Grossberg, S. and Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, 4, 453-466. (\*#&)
  23. Grossberg, S. and Somers, D. (1991). Synchronized oscillations during cooperative feature linking in visual cortex. In O. Simula (Ed.), **Proceedings of the international conference on artificial neural networks**. Amsterdam: Elsevier. (\*#&)
  24. Grossberg, S. and Somers, D. (1991). Synchronized oscillations during cooperative feature linking in visual cortex. Submitted for publication. (\*#&)
  25. Grossberg, S. and Wyse, L. (1991). Invariant recognition of cluttered scenes by a self-organizing ART architecture: Figure-ground separation. *Neural Networks*, in press. (\*#=)
  26. Grossberg, S. and Wyse, L. (1991). A neural model for figure-ground separation of connected visual figures. Submitted for publication. (\*#&)
  27. Mingolla, E. (1991). Neural dynamics of static and motion visual segmentation. In R. Plamondon and E. H. D. Chang (Eds.), **Pattern recognition: Architectures, algorithms, and applications**. Singapore: World Scientific Publishers.
  28. Mingolla, E. (1991). Neural dynamics of motion segmentation and grouping. Invited article in R. Lippmann, J. Moody, and D. S. Touretsky (Eds.), **Advances in neural information processing systems**. San Diego: Morgan Kaufman.



29. Mingolla, E., Todd, J.T., and Norman, J.F. (1991). The perception of globally coherent motion. Submitted for publication.

- \* Also supported in part by the Army Research Office (terminating 08/31/91).
- % Also supported in part by British Petroleum (terminating 08/31/92).
- # Also supported in part by DARPA (terminating 02/14/92).
- = Also supported in part by the Hughes Research Labs (terminated 11/30/90).
- & Also supported in part by NASA (terminating 08/31/92).
- + Also supported in part by the National Science Foundation (terminating 10/31/91).

## BOSTON UNIVERSITY PROJECT SUMMARIES

The following summaries highlight some aspects of the Boston University research that are not already included in the enclosed Abstracts.

### 1. Cortical Dynamics of Motion Perception [articles 17, 18, 20, 21, 27, 28]

This work develops a Motion Boundary Contour System (Motion BCS) to explain a large body of data concerning how we see things move. Our everyday percepts of moving objects are so immediate and compelling that the synthetic nature of the perceptual processes which generate these percepts are not easily understood. The task of rapidly detecting a leopard leaping from a jungle branch under a sun-dappled forest canopy illustrates the subtlety and vigor of these processes. Consider how spots on the leopard's coat move as its limbs and muscles surge. Imagine how patterns of light and shade play upon the leopard's coat as it leaps through the air. These luminance and color contours move across the leopard's body in a variety of directions that do not necessarily point in the direction of the leopard's leap. Indeed, the leopard's body generates a scintillating mosaic of moving contours that could easily prevent its detection. Remarkably, our perceptual processes are able to actively reorganize such a scintillating mosaic into a coherent object percept with a unitary direction-of-motion. The leopard as a whole then seems to quickly "pop out" from the jungle background and to draw our attention. Such a perceptual process clearly has a high survival value for animals who possess it.

This description of the leaping leopard emphasizes that the process of motion perception is an active one. It is capable of transforming a motion signal that is generated by a luminance contour into a different motion percept. In this sense, our percepts of moving objects are often percepts of *apparent motion*, albeit an adaptive and useful form of apparent motion. The task of understanding how we see "real" motion thus requires that we also understand "apparent" motion. The present article explains a large body of classical and recent data about apparent motion to further support a new theory of motion perception that was described in articles 20 and 21. Most of these data have not yet been explained by alternative theories of motion perception (see Abstract of article 20).

This new theory of motion perception grew out of an earlier theory of static form perception that was introduced by the Boston University group. A key new insight of the static form theory can be summarized by the paradoxical phrase that "all boundaries are invisible". An illustration of this property is provided by the percept of a reverse-contrast Kanizsa square. In this percept, a square boundary emerges between the four pac man inducers. The vertical components of this boundary join together dark-light vertical contrasts with light-dark vertical contrasts. Thus the boundaries can form between opposite directions-of-contrast. Another way of saying this is that the output of the boundary completion process is insensitive to direction-of-contrast, even though it is sensitive to amount-of-contrast. A process whose output does not distinguish between dark-light and light-dark cannot carry a visible signal. Hence "all boundaries are invisible."

This boundary completion process has been called the Boundary Contour System, or BCS, in order to emphasize that its boundaries emerge from contrast-sensitive processes. The boundaries formed by the BCS are not created only in response to edges. Rather, they may be generated in response to combinations of edge, texture, shading, and stereo information at multiple size scales. That is why the term "boundary completion" rather than "edge detection" is used. These form-sensitive boundary structures have been called *boundary webs* by Grossberg and Mingolla.

Since the BCS does not represent visible percepts, another process than boundary completion must also exist that does generate visible percepts. This process has been suggested

to discount the illuminant, or compensate for variable illumination conditions, and to fill-in surface properties of brightness, color, and depth using the discounted signals. It has been called the Feature Contour System, or FCS, because it generates the visible percepts that scientists had earlier attributed to "feature detectors," and it does so using a contrast-sensitive process.

What is the relationship between the contrast-sensitive processes of the BCS and the FCS? Remarkably, these processes obey laws that are computationally complementary. The BCS and FCS overcome the limitations of their complementary processes by interacting with one another through both serial and parallel pathways undergoing both feedforward and feedback interactions. These interactions give rise to a visual representation that is called a FACADE representation because it suggests how properties of Form-And-Color-And-DEpth are combined in a visual percept. The theory that explains how BCS and FCS interactions generate these representations is called FACADE Theory.

FACADE representations are predicted to occur in prestriate area V4 of the visual cortex. More generally, BCS and FCS processes have been used to explain and predict perceptual and neurobiological data about the regions V1, V2, and V4 of visual cortex, notably the cortical stream  $V1 \rightarrow V2 \rightarrow V4$  that has been linked to perceptual properties of static form, color, and depth. In keeping with these properties, the BCS is called the *Static BCS* in order to differentiate it from the *Motion BCS*.

Indeed, a parallel cortical stream  $V1 \rightarrow MT$  exists from cortical area V1 to area MT. Cells in area MT are sensitive to properties of motion. Why has Nature needed to evolve parallel cortical streams  $V1 \rightarrow V2$  and  $V1 \rightarrow MT$  for the processing of static form and moving form? This is a nontrivial question if only because the first processing stage in V1, the simple cells, are already sensitive to direction-of-motion and to changes in stimulus intensity. Why has evolution needed to generate region MT when even the simple cells of V1 are already direction-sensitive and change-sensitive? What computational properties are achieved by MT that are not already available in V1 and its prestriate projections V2 and V4?

A precise answer to this question has come into view through an analysis of why the Static BCS is not adequate for motion processing. This inadequacy of the Static BCS is a consequence of the fact that "all boundaries are invisible." The scientific explication of this paradoxical statement has, in fact, forced a pervasive shift in theoretical perspective that underlies much of the enhanced explanatory power of FACADE Theory.

In order to understand why the Static BCS is inadequate for motion processing, we show how the process which makes the output signals of the Static BCS insensitive to direction-of-contrast also makes them insensitive to direction-of-motion. A perceptual system whose output is insensitive to direction-of-motion is certainly not well suited to be a motion processor. In particular, simple cells combine their output signals to activate complex cells that are sensitive to the *amount* of image contrast, but not to the *direction* of image contrast. The complex cell response is insensitive to direction-of-contrast because it adds output signals from a pair of simple cells which are sensitive to opposite directions-of-contrast. This construction also renders the complex cells insensitive to direction-of-motion. For example, a vertically oriented model complex cell could respond, say, to a black-white vertical edge moving to the right or left *and* to a white-black vertical edge moving to the right or left. Thus the process whereby complex cells become insensitive to direction-of-contrast has rendered them insensitive to direction-of-motion. This combination of properties of cortical complex cells has been reported by several laboratories, notably the lab of Dan Pollen.

This observation led us to the following theoretical question: What is the minimal change of the Static BCS with which to fashion a Motion BCS whose output signals are insensitive to direction-of-contrast—which is just as important for processing static images as moving images—yet are sensitive to direction-of-motion? The Motion BCS that was hereby derived has been used here to explain a large data base about motion perception. As a result of this

approach, the Static BCS and the Motion BCS can be viewed as variations of one another. Prior to this observation, data about static perception and motion perception had typically been studied as parts of separate scientific enterprises. The present synthesis allows them to be explained as variations of a common design for the architecture of visual cortex.

## 2. The Organization of Cortical Systems for Form and Motion Perception [articles 15, 16]

Grossberg has further developed this theme by predicting that the Static BCS and Motion BCS are parallel subsystems of a single total BCS system. This prediction suggests that this total BCS system arises during cortical development as an expression of a global symmetry principle, called FM Symmetry (F=form, M=motion). Manifestations of this symmetry principle are familiar to us in our daily perceptual experiences, as noted below. The FM Symmetry principle rationalizes how the cortical streams  $V1 - V2$  and  $V1 - MT$  develop as parallel, homologous subsystems of a global design for visual cortex. Considering the large data base that can be explained by the theory, its rationale is unusually simple. FM Symmetry is predicted to control the simultaneous satisfaction of three constraints: (1) multiplicative interaction, or gating, of all combinations of sustained cell and transient cell output signals to form four sustained-transient cell types; (2) symmetric organization of these sustained-transient cell types into two opponent on-cell and off-cell pairs, such that (3) output signals from all the opponent cell types are independent of direction-of-contrast.

Multiplicative gating of sustained cells and transient cells is shown to generate change-sensitive receptive field properties of oriented on-cells and off-cells within the Static BCS, and direction-sensitive cells within the Motion BCS. The constraint that output signals be independent of direction-of-contrast enables both the Static BCS and the Motion BCS to generate emergent boundary segmentations along image contrast reversals.

The summary above suggests how the static-form and motion-form systems may both arise. This discussion does not, however, disclose how these systems control different perceptual properties whose behavioral usefulness has preserved their integrity throughout the evolutionary process. The following behavioral implications of the symmetry principle are explained by the theory.

We are all so familiar with the different geometries for processing static *orientations* and motion *directions* that we often take them for granted. For example, we all take for granted that the opposite *orientation* of "vertical" is "horizontal," a difference of  $90^\circ$ ; yet the opposite *direction* of "up" is "down," a difference of  $180^\circ$ . Why are the perceptual symmetries of static form and motion form different?

A clue is provided by considering how the  $90^\circ$  and  $180^\circ$  symmetries are reflected in percepts of negative afterimages. These symmetries suggest an opponent organization whereby orientations that differ by  $90^\circ$  are grouped together, whereas directions that differ by  $180^\circ$  are grouped together. Negative aftereffects illustrate a key property of this opponent organization. For example, after sustained viewing of a radial input pattern, looking at a uniform field triggers a percept of a circular MacKay afterimage. The orientations within the input and the circular afterimage differ from each other by  $90^\circ$ . After sustained viewing of a downwardly moving image, looking at a uniform field triggers a percept of an upwardly moving afterimage, as in the waterfall illusion. The directions within the downward input and the upward afterimage differ from each other by  $180^\circ$ .

In summary, the geometries of both static form perception and motion form perception include an opponent organization in which offset of the input pattern after sustained viewing triggers onset of a transient *antagonistic rebound*, or activation of the opponent channel.

Antagonistic rebound within opponent channels is needed to control the complementary perceptual processes of *resonance* and *reset*. Within the BCS, positive feedback signals between the hypercomplex cells and bipole cells can cooperatively link similarly oriented

features at approximately colinear locations into emergent boundary segmentations. These positive feedback interactions selectively amplify and sharpen the globally "best" cooperative grouping and provide the activation for inhibiting less favored groupings. The positive feedback interactions also subserve the coherence, hysteresis, and structural properties of the emergent segmentations.

The positive feedback can, however, maintain itself for a long time after visual inputs terminate. Thus the very existence of cooperative linking could seriously degrade perception by maintaining long-lasting positive afterimages, or smearing, of every percept.

Although some smearing can occur, it is known to be actively limited by inhibitory processes that are triggered by changing images. The new theory suggests how antagonistic rebounds between oppositely organized on-cells and off-cells can actively inhibit CC Loop resonances when the input pattern changes. This inhibitory process *resets* the resonance and enables the CC Loop to flexibly establish new resonances in response to rapidly changing scenes.

In summary, the symmetry principle that is predicted to control the parallel development of the static form and motion form systems enables these systems to rapidly reset their resonant segmentations in response to rapidly changing inputs.

### **3. Emergent Segmentation of Moving Images [articles 17, 18, 19, 27, 28]**

These articles model the process whereby local motion signals are grouped into a global boundary segmentation within the CC Loop of the Motion BCS. The first point of interest is that there does exist an analog of the CC Loop from the Static BCS within the Motion BCS. The second point is that the motion CC Loop is specialized to group motion *directions* that are pooled from many different oriented contrasts in an image, not merely the static *orientations* that are processed by the Static CC Loop.

These results suggest how ambiguous local movements on a complex moving shape are actively reorganized into a coherent global motion signal. Unlike many previous researchers, we analyse how a coherent motion signal is imparted to all regions of a moving figure, not only to regions at which unambiguous motion signals exist. The model hereby suggests a solution to the global aperture problem. Along the way, the model suggests an explanation of key motion segmentation and grouping phenomena, including the aperture problem, barberpole illusion, and motion capture. These analyses include hypotheses concerning the role of end-stopped simple cells, the spatial layout of simple cell receptive fields, and competition among signals sensitive to different directions-of-motion. These concepts are illustrated through computer simulations which study how the Motion BCS responds to changes in the bounding orientations, shapes, and motion directions of an object.

### **4. Psychophysical Studies of Motion Segmentation [article 29]**

In collaboration with James T. Todd and J. Farley Norman of Brandeis University, Ennio Mingolla has been conducting psychophysical investigations of the perception of globally coherent motion. Their research has examined how ambiguous velocity measures along smooth contours are spatially integrated to obtain a globally coherent perception of motion. Observers viewed displays containing a large number of apertures, with each aperture containing one or more contours, whose orientation and velocity could be independently specified. The total pattern of the contour trajectories across the individual apertures was manipulated to produce globally coherent motions, such as rotations, expansions, or translations. When the displays contained only straight contours extending to the circumferences of the apertures, observers' reports of global motion direction were biased whenever the sampling of contour orientations was asymmetric relative to the direction of motion. Performance was improved by the presence of identifiable features, such as line ends or crossings, whose trajectories could be tracked over time. The reports of the observers were consistent with a pooling

process involving a vector average of measures of the component of velocity normal to contour orientation, rather than with the predictions of the intersection-of-constraints analysis of "velocity space." This work has been presented in preliminary form at a poster session and is the subject of a talk at an international conference on vision. A manuscript has also been accepted for publication by *Vision Research*.

#### 5. Synchronizing Oscillations during Cooperative Feature Linking in Visual Cortex [articles 22, 23, 24]

The laboratories of Eckhorn and of Singer have recently reported that spatially distant cells in visual cortex that are tuned to similar visual features may oscillate in phase when stimulated by moving bar stimuli. The present results show how synchronous oscillations can occur as part of the CC Loop segmentation process. The main observation is that synchronous oscillations may occur when inhibitory interneurons of the CC Loop react more slowly than their excitatory counterparts. One of the hardest data properties to explain has been the speed with which synchrony develops. In the model, synchrony locks in very rapidly—in fact, within a single processing cycle—due to the cooperative grouping action of bipole cells. These cells were predicted to exist in 1984 by Grossberg, Cohen, and Mingolla, and were subsequently reported in neurophysiological experiments of von der Heydt *et al.* They are a key design feature of the CC Loop.

#### 6. Automatic Figure-Ground Separation of Connected Scenic Components [articles 25, 26]

An important stage in the perception and recognition of objects is the process whereby a figure, or object, in a scene is separated from other figures and background clutter. This is called the stage of *figure-ground separation*. Whereas knowledge about a figure may facilitate its separation, such knowledge is clearly not *necessary* for biological vision systems to carry out figure-ground separation. Experiences abound of unfamiliar figures that "pop out" from their backgrounds before they ever enter our corpus of learned knowledge about the world. The fact that figure-ground separation can occur even for unfamiliar figures contributes to the general-purpose nature of biological vision, which can process both unfamiliar and familiar scenes, and does not require prior instruction about an environment in order to operate effectively.

This work describes a new type of system that is capable of automatic figure-ground separation. This process separates scenic figures whose emergent boundary segmentations surround a connected region. As a result of this property, such a system can automatically distinguish between connected and disconnected spirals, a benchmark that gained fame through its emphasis in the book by Minsky and Papert on perceptrons.

Some themes of particular interest include the following: A new feedforward boundary segmentation model, called the CORT-X 2 model, is developed. It uses combinations of on-cells and off-cells, and of large and small receptive field sizes, to overcome deficiencies of these elements taken separately. On-cells and off-cells respond with *complementary* deficiencies to various noisy image properties. The same is true for small and large receptive field sizes. The CORT-X 2 model shows how to join these processing elements to overcome their complementary deficiencies.

A second theme is that the process of *filling-in* may be used to separate figure-from-ground.

A third theme is that *double opponent* cells, which usually are used for color processing, are here shown useful for figure-ground separation of regions whose boundaries contain gaps.

A fourth theme is that the work clarifies why humans *cannot* quickly distinguish the Minsky-Papert spirals, yet can rapidly detect conjunctions of disparity and color, or of disparity and motion, thereby clarifying results about visual search and attention from the labs of Triesman and of Nakayama.

## 7. An Improved Boundary Segmentation Network for Processing of Static Images [article 11]

This work develops an improved Boundary Contour System for the processing of static images. It uses insights from the Grossberg-Wyse model and a refined receptive field structure of bipole cells to demonstrate good texture segmentation of  $200 \times 400$  pixel images from the laboratory of Jacob Beck.

## 8. Autonomous Learning, Pattern Recognition, and Prediction [articles 6, 7, 8, 9]

An important open problem in biological science and technology is to design autonomous systems capable of learning to recognize and predict nonstationary data in which mixtures of rare, frequent, and unexpected events may occur. In order to cope with rare events, fast learning is needed. Fast learning can, however, destabilize many learning schemes. In order to cope with nonstationary combinations of rare and frequent events, different degrees of generalization, or code compression, must be learnable by a single system. Many learning schemes cannot simultaneously operate at multiple scales of coarseness. In order to rapidly learn different predictions in response to rare events than to a cloud of similar frequent events in which they are embedded, predictive feedback about success or failure needs to operate in real-time using only local operations to separate the rare exemplar from the frequent cloud. Many learning schemes that use predictive feedback can only operate in an off-line mode, or need to use slow learning, or are computed using non-local operations.

The present work introduces a new class of real-time neural networks that overcome all of these problems. These neural networks are defined by high-dimensional nonlinear dynamical systems that operate at multiple time scales. They are designed to carry out fast, stable, autonomous learning of recognition codes and multidimensional maps in response to arbitrary sequences of input patterns. In order to learn quickly and stably in response to a nonstationary input stream, the networks incorporate operations that were derived from an analysis of human cognition, and that have been used to explain and predict many behavioral and neural data. These operations include the learning of abstractions and expectations, paying attention, hypothesis testing, memory search, novelty detection, and confidence. Dynamical systems that embody these operations are often called Adaptive Resonance Theory, or ART, networks because such a network enters a resonant state when it pays attention to data about which it will learn.

The new neural network architecture, called ARTMAP, autonomously learns to classify arbitrarily many, arbitrarily ordered vectors into recognition categories based on predictive success. This supervised learning system is built up from a pair of ART modules ( $ART_a$  and  $ART_b$ ) that are capable of self-organizing stable recognition categories in response to arbitrary sequences of input patterns. During training trials, the  $ART_a$  module receives a stream  $\{a^{(p)}\}$  of input patterns, and  $ART_b$  receives a stream  $\{b^{(p)}\}$  of input patterns, where  $b^{(p)}$  is the correct prediction given  $a^{(p)}$ . These ART modules are linked by an associative learning network and an internal controller that ensures autonomous system operation in real time. During test trials, the remaining patterns  $a^{(p)}$  are presented without  $b^{(p)}$ , and their predictions at  $ART_b$  are compared with  $b^{(p)}$ . Tested on a benchmark machine learning database in both on-line and off-line simulations, the ARTMAP system learns orders of magnitude more quickly, efficiently, and accurately than alternative algorithms, and achieves 100% accuracy after training on less than half the input patterns in the database.

ARTMAP achieves these properties by using an internal controller that realizes a new Minimax Learning Rule, which conjointly maximizes predictive generalization and minimizes predictive error by linking predictive success to category size on a trial-by-trial-basis, using only local operations. This computation increases the vigilance parameter  $\rho_a$  of  $ART_a$  by the minimal amount needed to correct a predictive error at  $ART_b$ . Parameter  $\rho_a$  calibrates

the minimum confidence that  $ART_a$  must have in a category, or hypothesis, activated by an input  $a^{(p)}$  in order for  $ART_a$  to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing. Parameter  $\rho_a$  is compared with the degree of match between  $a^{(p)}$  and the top-down learned expectation, or prototype, that is read-out subsequent to activation of an  $ART_a$  category. Search occurs if the degree of match is less than  $\rho_a$ . ARTMAP is thus a type of self-organizing expert system that calibrates the selectivity of its hypotheses based upon predictive success. As a result, rare but important events can be quickly and sharply distinguished even if they are similar to frequent events with different consequences.

Between input trials  $\rho_a$  relaxes to a baseline vigilance  $\bar{\rho}_a$ . When  $\bar{\rho}_a$  is large, the system runs in a conservative mode, wherein predictions are made only if the system is confident of the outcome. Very few false-alarm errors then occur at any stage of learning, yet the system reaches asymptote with no loss of speed. Because ARTMAP learning is self-stabilizing, it can continue learning one or more databases, without degrading its corpus of memories, until its full memory capacity is utilized.

## 9. Vector Associative Maps: Self-Organizing Spatial Representations and Motor Controllers [articles 12, 13, 14]

This work develops self-organizing neural circuits for the control of planned arm movements during visually guided reaching. More generally, it introduces a modelling framework for unsupervised, real-time, error-based learning. The problem that motivated these results concerns the issue of how a child learns to reach for objects that it sees. This problem requires an understanding of the interactions between two distinct modalities: vision (seeing an object) and motor control (moving a limb). In particular, we need to characterize the self-regulating mechanisms whereby an individual can stably learn transformations within and between the different modalities that provide accurate control of goal-oriented movements.

The Swiss psychologist Jean Piaget has suggested that learning of this type can take place through a *circular reaction*. As a child performs random, spontaneously generated movements of his arm, its eyes follow the arm's motion, thereby enabling learning of a transformation from a visual representation of arm position to a motor representation of the same arm position. As more and more arm positions are sampled through time, the transformation eventually enables the child to reach for objects that it sees.

A similar kind of circular reaction is found in the "babbling phase" of speech acquisition in infants. Here interactions take place between the speech perception (hearing) and production (speaking) systems. When the child babbles a sound, an auditory feedback representation of the sound is activated and coexists with the motor representation that gave rise to the sound. As the child learns a transformation from the auditory representation to the motor representation, it can begin to imitate heard sounds that are produced by other speakers.

The above examples introduce the circular reaction as an autonomously controlled behavioral cycle with two components: *production* and *perception*. Learning links the two modalities to enable sensory-guided action to occur. Such a circular reaction is *intermodal*; that is, it consists of the coupling of two systems operating in different modalities.

In order for the intermodal circular reaction to generate stable learning of the parameters that couple the two systems, the control parameters within each system must already be capable of accurate performance. Otherwise, performance may not be consistent across trials and a stable mapping could not be learned between different modalities. Thus it is necessary to self-organize the correct *intramodal* control parameters before a stable *intermodal* mapping can be learned.

It is here shown how the arm movement system can endogenously generate movements during a "motor babbling" phase. These movements create the data needed to learn cor-



rect arm movement control parameters. These movements also activate the target position representations that are used to learn the visuo-motor transformation that controls visually guided reaching.

The results grew out of the VITE model for variable-speed adaptive control of multi-joint limb trajectories. Using this model, Bullock and Grossberg have suggested how motor synergies can be dynamically bound and unbound in real-time. Once bound, the multiple muscles within a synergy can move a limb at variable speeds by synchronously contracting variable amounts in equal time. In this view, trajectory formation is an emergent invariant that arises through interactions among two broad types of control mechanisms: planned control and automatic control. Planned control variables include (1) target position, or where we want to move; and (2) speed of movement, or how fast we want to move to the desired position, and the "will" to move at all. Automatic control variables compensate for (3) the present position of the arm; (4) unexpected inertial forces and external loads; and (5) changes in the physiognomy of the motor plant, due for example to growth, injury, exercise, and aging.

The VAM model was discovered through an analysis of how parameters of the VITE model are learned during motor behavior. A major surprise was that the Difference Vector (DV) that controls *synchronous performance* of a motor synergy can also be used as a vector error signal to control *learning* of coordinate transformations.

A second surprise was that this learning scheme generalizes. For example, it has since been used to clarify how the 3-D body-centered spatial representations that control motor trajectories are themselves learned in a way that is sensitive to bodily parameters. This latter insight has opened up a large new field of investigation in the areas of spatial orientation and flexible sensory-motor control.

## SELECTED ABSTRACTS

## ADAPTIVE NEURAL NETWORKS FOR CONTROL OF MOVEMENT TRAJECTORIES INVARIANT UNDER SPEED AND FORCE RESCALING

Daniel Bullock† and Stephen Grossberg‡

*Human Movement Science*, 1991, 10, 3-53

This article describes two neural network modules that form part of an emerging theory of how adaptive control of goal-directed sensory-motor skills is achieved by humans and other animals. The Vector-Integration-To-Endpoint (VITE) model suggests how synchronous multi-joint trajectories are generated and performed at variable speeds. The Factorization-of-LEngth-and-TEnSion (FLETE) model suggests how outflow movement commands from a VITE model may be performed at variable force levels without a loss of positional accuracy. The invariance of positional control under speed and force rescaling sheds new light upon a familiar strategy of motor skill development: Skill learning begins with performance at low speed and low limb compliance and proceeds to higher speeds and compliances. The VITE model helps to explain many neural and behavioral data about trajectory formation, including data about neural coding within the posterior parietal cortex, motor cortex, and globus pallidus, and behavioral properties such as Woodworth's Law, Fitts Law, peak acceleration as a function of movement amplitude and duration, isotonic arm movement properties before and after arm-deafferentation, central error correction properties of isometric contractions, motor priming without overt action, velocity amplification during target switching, velocity profile invariance across different movement distances, changes in velocity profile asymmetry across different movement durations, staggered onset times for controlling linear trajectories with synchronous offset times, changes in the ratio of maximum to average velocity during discrete versus serial movements, and shared properties of arm and speech articulator movements. The FLETE model provides new insights into how spino-muscular circuits process variable forces without a loss of positional control. These results explicate the size principle of motor neuron recruitment, descending co-contractive compliance signals, Renshaw cells, Ia interneurons, fast automatic reactive control by ascending feedback from muscle spindles, slow adaptive predictive control via cerebellar learning using muscle spindle error signals to train adaptive movement gains, fractured somatotopy in the opponent organization of cerebellar learning, adaptive compensation for variable moment-arms, and force feedback from Golgi tendon organs. More generally, the models provide a computational rationale for the use of nonspecific control signals in volitional control, or "acts of will", and of efference copies and opponent processing in both reactive and adaptive motor control tasks.

---

† Supported in part by the National Science Foundation (NSF IRI-87-16960).

‡ Supported in part by the National Science Foundation (NSF IRI-87-16960) and the Air Force Office of Scientific Research (AFOSR 90-0128 and AFOSR 90-0175).

**EMERGENCE OF TRI-PHASIC MUSCLE ACTIVATION FROM  
THE NONLINEAR INTERACTIONS OF CENTRAL AND  
SPINAL NEURAL NETWORK CIRCUITS**

Daniel Bullock† and Stephen Grossberg‡

*Human Movement Science*, in press, 1991

The origin of the tri-phasic burst pattern, observed in the EMGs of opponent muscles during rapid self-terminated movements, has been controversial. Here we show by computer simulation that the pattern emerges from interactions between a central neural trajectory controller (VITE circuit) and a peripheral neuromuscular force controller (FLETE circuit). Both neural models have been derived from simple functional constraints that have led to principled explanations of a wide variety of behavioral and neurobiological data, including, as shown here, the generation of tri-phasic bursts.

---

† Supported in part by the National Science Foundation (NSF IRI-87-16960).

‡ Supported in part by the National Science Foundation (NSF IRI-87-16960), the Air Force Office of Scientific Research (AFOSR URI 90-0175), and Defense Advanced Research Projects Agency (DARPA) (AFOSR-90-0083).

**ARTMAP:  
SUPERVISED REAL-TIME LEARNING AND CLASSIFICATION  
OF NONSTATIONARY DATA BY A SELF-ORGANIZING  
NEURAL NETWORK**

Gail A. Carpenter†, Stephen Grossberg‡, and John H. Reynolds§

*Neural Networks*, in press, 1991

This article introduces a new neural network architecture, called ARTMAP, that autonomously learns to classify arbitrarily many, arbitrarily ordered vectors into recognition categories based on predictive success. This supervised learning system is built up from a pair of Adaptive Resonance Theory modules ( $ART_a$  and  $ART_b$ ) that are capable of self-organizing stable recognition categories in response to arbitrary sequences of input patterns. During training trials, the  $ART_a$  module receives a stream  $\{a^{(p)}\}$  of input patterns, and  $ART_b$  receives a stream  $\{b^{(p)}\}$  of input patterns, where  $b^{(p)}$  is the correct prediction given  $a^{(p)}$ . These ART modules are linked by an associative learning network and an internal controller that ensures autonomous system operation in real time. During test trials, the remaining patterns  $a^{(p)}$  are presented without  $b^{(p)}$ , and their predictions at  $ART_b$  are compared with  $b^{(p)}$ . Tested on a benchmark machine learning database in both on-line and off-line simulations, the ARTMAP system learns orders of magnitude more quickly, efficiently, and accurately than alternative algorithms, and achieves 100% accuracy after training on less than half the input patterns in the database. It achieves these properties by using an internal controller that conjointly maximizes predictive generalization and minimizes predictive error by linking predictive success to category size on a trial-by-trial basis, using only local operations. This computation increases the vigilance parameter  $\rho_a$  of  $ART_a$  by the minimal amount needed to correct a predictive error at  $ART_b$ . Parameter  $\rho_a$  calibrates the minimum confidence that  $ART_a$  must have in a category, or hypothesis, activated by an input  $a^{(p)}$  in order for  $ART_a$  to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing. Parameter  $\rho_a$  is compared with the degree of match between  $a^{(p)}$  and the top-down learned expectation, or prototype, that is read-out subsequent to activation of an  $ART_a$  category. Search occurs if the degree of match is less than  $\rho_a$ . ARTMAP is hereby a type of self-organizing expert system that calibrates the selectivity of its hypotheses based upon predictive success. As a result, rare but important events can be quickly and sharply distinguished even if they are similar to frequent events with different consequences. Between input trials  $\rho_a$  relaxes to a baseline vigilance  $\bar{\rho}_a$ . When  $\bar{\rho}_a$  is large, the system runs in a conservative mode, wherein predictions are made only if the system is confident of the outcome. Very few false-alarm errors then occur at any stage of learning, yet the system reaches asymptote with no loss of speed. Because ARTMAP learning is self-stabilizing, it can continue learning one or more databases, without degrading its corpus of memories, until its full memory capacity is utilized.

---

† Supported in part by BP (98-A-1204), DARPA (AFOSR 90-0083), and the National Science Foundation (NSF IRI-90-00539).

‡ Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175 and AFOSR 90-0128), the Army Research Office (ARO DAAL-03-88-K0088), and DARPA (AFOSR 90-0083).

§ Supported in part by DARPA (AFOSR 90-0083).

## ART2-A: AN ADAPTIVE RESONANCE ALGORITHM FOR RAPID CATEGORY LEARNING AND RECOGNITION

Gail A. Carpenter†, Stephen Grossberg‡, and David B. Rosen§

*Neural Networks*, in press, 1991

This article introduces ART2-A, an efficient algorithm that emulates the self-organizing pattern recognition and hypothesis testing properties of the ART 2 neural network architecture, but at a speed two to three orders of magnitude faster. Analysis and simulations show how the ART2-A systems correspond to ART 2 dynamics at both the fast-learn limit and at intermediate learning rates. Intermediate learning rates permit fast commitment of category nodes but slow recoding, analogous to properties of word frequency effects, encoding specificity effects, and episodic memory. Better noise tolerance is hereby achieved without a loss of learning stability. The ART 2 and ART2-A systems are contrasted with the leader algorithm. The speed of ART2-A makes practical the use of ART 2 modules in large-scale neural computation.

---

† Supported in part by British Petroleum (89-A-1204), DARPA (AFOSR 90-0083), and the National Science Foundation (NSF IRI-90-00530).

‡ Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175 and AFOSR 90-0128), the Army Research Office (ARO DAAL-03-88-K0088), and DARPA (AFOSR 90-0083).

§ Supported in part by DARPA (AFOSR 90-0083).

**PREATTENTIVE TEXTURE SEGMENTATION AND GROUPING  
BY THE BOUNDARY CONTOUR SYSTEM**

Dan Cruthirds, Alan Gove, Stephen Grossberg, and Ennio Mingolla

*Proceedings of the International Joint Conference on Neural Networks, 1991*

An improved Boundary Contour System (BCS) neural network model of preattentive vision is applied to two images that produce strong "pop-out" of emergent groupings in humans. In humans these images generate groupings collinear with or perpendicular to image contrasts. Analogous groupings occur in computer simulations of the model. Long-range cooperative and short-range competitive processes of the BCS dynamically form the stable groupings of texture regions in response to the images.

---

**Acknowledgements:** This research was supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175) and Hughes Aircraft Company (SK-902369-SDB).

# VECTOR ASSOCIATIVE MAPS: UNSUPERVISED REAL-TIME ERROR-BASED LEARNING AND CONTROL OF MOVEMENT TRAJECTORIES

Paolo Gaudiano† and Stephen Grossberg‡

*Neural Networks*, 1991, 4, 147-183

This article describes neural network models for adaptive control of arm movement trajectories during visually guided reaching and, more generally, a framework for unsupervised real-time error-based learning. The models clarify how a child, or untrained robot, can learn to reach for objects that it sees. Piaget has provided basic insights with his concept of a *circular reaction*: As an infant makes internally generated movements of its hand, the eyes automatically follow this motion. A transformation is learned between the visual representation of hand position and the motor representation of hand position. Learning of this transformation eventually enables the child to accurately reach for visually detected targets. Grossberg and Kuperstein have shown how the eye movement system can use visual error signals to correct movement parameters via cerebellar learning. Here it is shown how endogenously generated arm movements lead to adaptive tuning of arm control parameters. These movements also activate the target position representations that are used to learn the visuo-motor transformation that controls visually guided reaching. The AVITE model presented here is an adaptive neural circuit based on the Vector Integration to Endpoint (VITE) model for arm and speech trajectory generation of Bullock and Grossberg. In the VITE model, a Target Position Command (TPC) represents the location of the desired target. The Present Position Command (PPC) encodes the present hand-arm configuration. The Difference Vector (DV) population continuously computes the difference between the PPC and the TPC. A speed-controlling GO signal multiplies DV output. The PPC integrates the (DV).(GO) product and generates an outflow command to the arm. Integration at the PPC continues at a rate dependent on GO signal size until the DV reaches zero, at which time the PPC equals the TPC. The AVITE model explains how self-consistent TPC and PPC coordinates are autonomously generated and learned. Learning of AVITE parameters is regulated by activation of a self-regulating Endogenous Random Generator (ERG) of training vectors. Each vector is integrated at the PPC, giving rise to a movement command. The generation of each vector induces a complementary postural phase during which ERG output stops and learning occurs. Then a new vector is generated and the cycle is repeated. This cyclic, biphasic behavior is controlled by a specialized gated dipole circuit. ERG output autonomously stops in such a way that, across trials, a broad sample of workspace target positions is generated. When the ERG shuts off, a modulator gate opens, copying the PPC into the TPC. Learning of a transformation from TPC to PPC occurs using the DV as an error signal that is zeroed due to learning. This learning scheme is called a Vector Associative Map, or VAM. The VAM model is a general-purpose device for autonomous real-time error-based learning and performance of associative maps. The DV stage serves the dual function of reading out new TPCs during performance and reading in new adaptive weights during learning, without a disruption of real-time operation. VAMs thus provide an on-line unsupervised alternative to the off-line properties of supervised error-correction learning algorithms. VAMs and VAM cascades for learning motor-to-motor and spatial-to-motor maps are described. VAM models and Adaptive Resonance Theory (ART) models exhibit complementary matching, learning, and performance properties that together provide a foundation for designing a total sensory-cognitive and cognitive-motor autonomous system.

---

† Supported in part by the National Science Foundation (NSF IRI-87-16960).

‡ Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), DARPA (AFOSR 90-0083), and the National Science Foundation (NSF IRI-87-6960).



## WHY DO PARALLEL CORTICAL SYSTEMS EXIST FOR THE PERCEPTION OF STATIC FORM AND MOVING FORM?

Stephen Grossberg†

*Perception and Psychophysics*, 1991, **49**, 117-141

This article analyses computational properties that clarify why the parallel cortical systems  $V1 \rightarrow V2$ ,  $V1 \rightarrow MT$ , and  $V1 \rightarrow V2 \rightarrow MT$  exist for the perceptual processing of static visual forms and moving visual forms. The article describes a symmetry principle, called FM Symmetry, that is predicted to govern the development of these parallel cortical systems by computing all possible ways of symmetrically gating sustained cells with transient cells and organizing these sustained-transient cells into opponent pairs of on-cells and off-cells whose output signals are insensitive to direction-of-contrast. This symmetric organization explains how the static form system (Static BCS) generates emergent boundary segmentations whose outputs are insensitive to direction-of-contrast and insensitive to direction-of-motion, whereas the motion form system (Motion BCS) generates emergent boundary segmentations whose outputs are insensitive to direction-of-contrast but sensitive to direction-of-motion. FM Symmetry clarifies why the geometries of static and motion form perception differ; for example, why the opposite orientation of vertical is horizontal ( $90^\circ$ ), but the opposite direction of up is down ( $180^\circ$ ). Opposite orientations and directions are embedded in gated dipole opponent processes that are capable of antagonistic rebound. Negative afterimages, such as the MacKay and waterfall illusions, are hereby explained, as are aftereffects of long-range apparent motion. These antagonistic rebounds help to control a dynamic balance between complementary perceptual states of resonance and reset. Resonance cooperatively links features into emergent boundary segmentations via positive feedback in a CC Loop, and reset terminates a resonance when the image changes, thereby preventing massive smearing of percepts. These complementary preattentive states of resonance and reset are related to analogous states that govern attentive feature integration, learning, and memory search in Adaptive Resonance Theory. The mechanism used in the  $V1 \rightarrow MT$  system to generate a wave of apparent motion between discrete flashes may also be used in other cortical systems to generate spatial shifts of attention. The theory suggests how the  $V1 \rightarrow V2 \rightarrow MT$  cortical stream helps to compute moving-form-in-depth and how long-range apparent motion of illusory contours occurs. These results collectively argue against vision theories that espouse independent processing modules. Instead, specialized subsystems interact to overcome computational uncertainties and complementary deficiencies, to cooperatively bind features into context-sensitive resonances, and to realize symmetry principles that are predicted to govern the development of visual cortex.

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K-0088), DARPA (AFOSR 90-0083), and Hughes Research Labs (S1-903136).

## CORTICAL DYNAMICS OF VISUAL MOTION PERCEPTION: SHORT-RANGE AND LONG-RANGE APPARENT MOTION

Stephen Grossberg† and Michael E. Rudd‡

*Psychological Review*, in press, 1991

This article describes further evidence for a new neural network theory of biological motion perception that is called a Motion Boundary Contour System. This theory clarifies why parallel streams  $V1 \rightarrow V2$  and  $V1 \rightarrow MT$  exist for static form and motion form processing among the areas  $V1$ ,  $V2$ , and  $MT$  of visual cortex. The Motion Boundary Contour System consists of several parallel copies, such that each copy is activated by a different range of receptive field sizes. Each copy is further subdivided into two hierarchically organized subsystems: a Motion Oriented Contrast Filter, or MOC Filter, for preprocessing moving images; and a Cooperative-Competitive Feedback Loop, or CC Loop, for generating emergent boundary segmentations of the filtered signals. The present article uses the MOC Filter to explain a variety of classical and recent data about short-range and long-range apparent motion percepts that have not yet been explained by alternative models. These data include split motion; reverse-contrast gamma motion; delta motion; visual inertia; group motion in response to a reverse-contrast Ternus display at short interstimulus intervals; speed-up of motion velocity as interflash distance increases or flash duration decreases; dependence of the transition from element motion to group motion on stimulus duration and size; various classical dependencies between flash duration, spatial separation, interstimulus interval, and motion threshold known as Korte's Laws; and dependence of motion strength on stimulus orientation and spatial frequency. These results supplement earlier explanations by the model of apparent motion data that other models have not explained; a recent proposed solution of the global aperture problem, including explanations of motion capture and induced motion; an explanation of how parallel cortical systems for static form perception and motion form perception may develop, including a demonstration that these parallel systems are variations on a common cortical design; an explanation of why the geometries of static form and motion form differ, in particular why opposite orientations differ by  $90^\circ$ , whereas opposite directions differ by  $180^\circ$ , and why a cortical stream  $V1 \rightarrow V2 \rightarrow MT$  is needed; and a summary of how the main properties of other motion perception models can be assimilated into different parts of the Motion Boundary Contour System design.

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K0088), DARPA (AFOSR-90-0083), and Hughes Aircraft Company (S1-903136).

‡ Supported in part by the Army Research Office (ARO DAAL-03-88-K0088).

# SYNCHRONIZED OSCILLATIONS DURING COOPERATIVE FEATURE LINKING IN A CORTICAL MODEL OF VISUAL PERCEPTION

Stephen Grossberg† and David Somerst‡

*Neural Networks*, 1991, 4, 453-466

A neural network model of synchronized oscillator activity in visual cortex is presented in order to account for recent neurophysiological findings that such synchronization may reflect global properties of the stimulus. In these recent experiments, it was reported that synchronization of oscillatory firing responses to moving bar stimuli occurred not only for nearby neurons, but also occurred between neurons separated by several cortical columns (several *mm* of cortex) when these neurons shared some receptive field preferences specific to the stimuli. These results were obtained not only for single bar stimuli but also across two disconnected, but colinear, bars moving in the same direction. Our model and computer simulations obtain these synchrony results across both single and double bar stimuli. For the double bar case, synchronous oscillations are induced in the region between the bars, but no oscillations are induced in the regions beyond the stimuli. These results were achieved with cellular units that exhibit limit cycle oscillations for a robust range of input values, but which approach an equilibrium state when undriven. Single and double bar synchronization of these oscillators was achieved by different, but formally related, models of preattentive visual boundary segmentation and attentive visual object recognition, as well as nearest-neighbor and randomly coupled models. In preattentive visual segmentation, synchronous oscillations may reflect the binding of local feature detectors into a globally coherent grouping. In object recognition, synchronous oscillations may occur during an attentive resonant state that triggers new learning. These modelling results support earlier theoretical predictions of synchronous visual cortical oscillations and demonstrate the robustness of the mechanisms capable of generating synchrony.

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K0088), and DARPA (AFOSR 90-0083).

‡ Supported in part by NASA (NGT-50497).

## A NEURAL NETWORK ARCHITECTURE FOR FIGURE-GROUND SEPARATION OF CONNECTED SCENIC FIGURES

Stephen Grossberg† and Lonce Wyset‡

*Neural Networks*, in press, 1991

A neural network model, called an FBF network, is proposed for automatic parallel separation of multiple image figures from each other and their backgrounds in noisy gray-scale or multi-colored images. The figures can then be processed in parallel by an array of self-organizing Adaptive Resonance Theory (ART) neural networks for automatic target recognition. An FBF network can automatically separate the disconnected but interleaved spirals that Minsky and Papert introduced in their book *Perceptrons*. The network's design also clarifies why humans cannot rapidly separate interleaved spirals, yet can rapidly detect conjunctions of disparity and color, or of disparity and motion, that distinguish target figures from surrounding distractors. Figure-ground separation is accomplished by iterating operations of a Feature Contour System (FCS) and a Boundary Contour System (BCS) in the order FCS-BCS-FCS, hence the term FBF, that have been derived from an analysis of biological vision. The FCS operations include the use of nonlinear shunting networks to compensate for variable illumination and nonlinear diffusion networks to control filling-in. A key new feature of an FBF network is the use of filling-in for figure-ground separation. The BCS operations include oriented filters joined to competitive and cooperative interactions designed to detect, regularize, and complete boundaries in up to 50 percent noise, while suppressing the noise. A modified CORT-X filter is described which uses both on-cells and off-cells to generate a boundary segmentation from a noisy image.

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), the Army Research Office (ARO DAAL-03-88-K0088), DARPA (AFOSR 90-0083), and Hughes Research Laboratories (S1-804481-D and S1-903136).

‡ Supported in part by the American Society for Engineering Education and Hughes Research Laboratories (S1-804481-D).

## THE PERCEPTION OF GLOBALLY COHERENT MOTION

Ennio Mingolla†, James T. Todd, and Farley Norman

How do human observers perceive a coherent pattern of motion from a disparate set of local velocity measures? Our research has examined how ambiguous velocity measures along smooth contours are spatially integrated to obtain a globally coherent perception of motion. Observers viewed displays containing a large number of apertures, with each aperture containing one or more contours whose orientations and velocities could be independently specified. The total pattern of the contour trajectories across the individual apertures was manipulated to produce globally coherent motions, such as rotations, expansions, or translations. When the displays contained only straight contours extending to the circumferences of the apertures, observers' reports of global motion direction were biased whenever the sampling of contour orientations was asymmetric relative to the direction of motion. Performance was improved by the presence of identifiable features, such as line ends or crossings, whose trajectories could be tracked over time. The reports of our observers were consistent with a pooling process involving a vector average of measures of the component of velocity normal to contour orientation, rather than with the predictions of the intersection-of-constraints analysis of "velocity space."

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175).

NORTHEASTERN UNIVERSITY  
PROJECT SUMMARIES

Adam Reeves, Principal Investigator

### Color Appearance and Color Mechanisms

We proposed to study (1) the perception of chromatic surface appearance in complex 2-D and 3-D visual displays; and (2) underlying chromatic mechanisms. We have also studied (3) a basis function for spatial vision.

Under heading (1), it was proposed to extend our (Arend and Reeves, *JOSA*, 3, 1986, 1743-1751) original matching paradigm. We found that once visual adaptation was controlled the extent of color constancy depended heavily on task instructions but hardly at all on stimulus configuration. We found nearly the same results as before when a long intervening dark period was used to try and disassociate the test and standard displays. This work is being prepared for publication. We also found similar results in a new paradigm, in which the subject was asked to adjust test patch chromaticity to a unique color, assuming that the memory representations of unique colors (such as yellow, grey, blue, etc.) are stable over time and, especially, are independent of the current display illumination. However, in this paradigm color constancy was incrementally better than before. Moreover, Troost and De Weert (1991, "Naming versus matching in color constancy," submitted ms) have replicated our matching work, but found near-constancy (actually, over constancy) in a color naming paradigm. Our current efforts are directed towards establishing the basis for this difference between our results and the naming results.

We also proposed to study the perception of an induced color, brown. As planned, we presented yellow test stripes alternating with white stripes in the form of sinusoidal gratings. We had planned to obtain a CSF by measuring the decrement in the intensity of the yellow stripe necessary to induce brown as a function of grating period. However, we have found very poor or no brown responses in this situation. We are currently varying the stimulus configuration in an attempt to understand why.

We (Yang, Peli, and Reeves) have also studied the perception of achromatic displays; in particular, we have measured perceived contrast as a function of luminance. Unlike previous reports, we found little evidence for constancy of perceived contrast once luminance varied by a factor of ten or more. Earlier work showed constancy over a very large range, but used haploscopic displays in which the eyes can adapt independently. Our work used natural vision and limited adaptation by asking subjects to look back and forward between displays, as in the work with Arend.

Under (2), we (Reeves, Wu, Armington) have obtained strong evidence for color opponency in the pattern-elicited electro-retinogram (PERG), an objective indicator of retinal function. Trying to use an adaptation of the Stiles paradigm to isolate middle-wave cone responses, a method which has worked for flash ERGs, we recorded instead a strongly opponent contribution to the PERG. That is, at wavelengths for which the red-green opponent system is most active, the PERG shows a decline in sensitivity relative to neutral wavelengths. This work is being presented at ISCEV and is being prepared for publication.

We (Schirillo and Reeves) have made a psychophysical study of the field additivity of the pathway responsible for detection of blue-green and green test flash increments, following up on earlier work which showed field additivity in Stiles' configuration ( $\pi_4$ ). We found field additivity in Stockman's configuration, which taps  $\pi_4^*$ , but sub-additivity in David Foster's, in which the test and field are co-incident and opponent responses are supposedly produced. These results show that the isolation of a field-additive pathway depends not only on wavelength composition of the stimuli, but also on spatial configuration.

We (Reeves, Rudd, and Grossberg) have also been successful in modelling the effects of flicker and duty cycle of an adapting light on the extent of transient tritanopia, using a

Grossberg dipole model, and we have obtained reasonable fits to the currently available data. The model is based on presumed retinal functioning, as transient tritanopia is visible in the ERG. Development of the complete model was slow, and so model predictions, which we had planned to test in the PI's Maxwellian-view system this year, have not yet been tested.

Under (3), we (Yang and Reeves) have also developed a model for spatial vision which was not originally proposed as part of the URI grant, but which is related to current efforts to model spatial vision at the Center. We have elicited visual-evoked potentials (VEPs) using one-dimensional gratings in the form of Hermitian functions  $H$  (these are Gaussian derivatives weighted by exponentials). These functions form a complete, orthonormal, spatially localized basis set; other basis sets, such as Gabors, standard Gaussian derivatives, and sinusoids, share some but not all of these properties. We used the power of the VEP as the response measure. Power is linear with contrast; linear superposition of power holds; and linear differencing also holds (i.e., the power evoked by a change from  $H_i$  to  $H_j$  equals the difference between the power evoked by  $H_i$  and that by  $H_j$ ). The effect of scale (size of the stimulus) can be normalized away. Using an approximation to the optical MTF of the eye, we have been able to predict these results with a one-parameter black-box model. We also used the same model to predict psychophysical discrimination between  $H_i$  and  $H_j$  (discrimination scores are lower when the spectral overlap produced by passing the stimuli through the MTF of the eye is greater). In future work, we hope to develop a neural model to fill in the black box.

**CONSULTANT:** L. Arend, Eye Research Institute of Retina Foundation, 20 Staniford Street, Boston Mass 02114.

## PAPERS PREPARED UNDER THE URI GRANT

Schirillo, J., Reeves, A., and Arend, L.E. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception and Psychophysics*, **48**, 82-90.

Wu, S., Armington, J.C., and Reeves, A. (1990). Electroretinogram and visual evoked responses to pattern displacement. *Visual Neuroscience* (accepted for publication).

Yang, J., Peli, E., Goldstein, R., and Reeves, A. (1991). The effect of luminance on suprathreshold contrast matching. *Journal of the Optical Society of America* (accepted for publication).

Reeves, A. (1991). The visual perception of time: Temporal resolution. In **Handbook of perception and action, Volume 1: Perception**, W. Prinz and B. Bridgeman (eds.). New York: Academic Press.

Reeves, A. and Barse, M. (1991). The luminance dependence of the pattern-evoked electroretinogram. In **Channels in the visual nervous system: Neurophysiology, psychophysics, models**, B. Blum (ed.). Tel Aviv: Freund.

## CONFERENCE PRESENTATIONS

Note: IOVS refers to a *Supplement to Investigative Ophthalmology and Visual Science*, which abstracts papers presented to the annual meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida.

Wu, S., Armington, J.C., and Reeves, A. (1990). Linearity and non-linearity of visual responses evoked by pattern displacement. *IOVS Supplement*, 31.

Yang, J., Peli, E., Goldstein, R., and Reeves, A. (1990). The effects of luminance on superthreshold contrast matching. IOVS Supplement, 31.

Reeves, A. and Tijus, C. (1990). Extremely rapid visual erasure. IOVS Supplement, 31.

Wu, S., Armington, J.C., and Reeves, A. (1990). Isolation of middle-wave cones in the ERG. *Perception*, 19, 4, A34c.

Yang, J. and Reeves, A. (1990). A polynomial basis function for vision: Tests with visual evoked potentials. *Perception*, 19, 4, A77a (ECV Supplement).

Reeves, A. and Yang, J. (1990). Contrast linearity and superposition of visual potentials evoked by pattern-reversing weighted Hermite Polynomials. Optical Society of America, Boston, November 1990.

Schirillo, J. and Reeves, A. (1990). Spatial factors determining field additivity of the middle-wavelength cone pathway. Optical Society of America, Boston, November 1990.

Reeves, A. and Tijus, C. (1990). The "pop-out" effect in a simple 3D visual matching task. Presented to Cognitiva, Madrid.



## HARVARD UNIVERSITY/CAMBRIDGE UNIVERSITY PROJECT SUMMARIES

John Daugman, Principal Investigator

- Daugman, J.G. (1991). Self-similar oriented wavelet pyramids with sequence regularity: Conjectures about neural non-orthogonality. In **Tacit assumptions in vision research** (A. Gorea, Editor). Cambridge: Cambridge University Press.
- Daugman, J.G., Downing, C.J. and Palafox, G. (1990). Two paradoxes of motion perception. *Perception*, 19(11), Proceedings of the ECVF-90, Paris.
- Daugman, J.G. (1990). **Signal processing by neural mechanisms**. Toshiba Endowed Lecture Series, Tokyo Institute of Technology, Department of Computer Science (Inaugural Volume), 249 pages. Tokyo: Tokyo Institute of Technology Press.
- Daugman, J.G. (1990). An information-theoretic view of analog representation in striate cortex. In **Computational neuroscience** (E. Schwartz, Editor). Cambridge, MA: MIT Press.
- Daugman, J.G. (1990). Brain metaphor and brain theory. In **Computational neuroscience** (E. Schwartz, Editor). Cambridge, MA: MIT Press.
- Daugman, J.G. (1990). Non-orthogonal wavelet representations in relaxation networks: Image encoding and analysis with biological visual primitives. In **Neurocomputers** (J. Taylor, Editor). London: Institute of Physics, IOP Press.

### Summary of Recent Research

Towards the end of 1990 J. Daugman moved his laboratory from Harvard University to Cambridge University, where he joined an active research group in computational neuroscience within the Faculty of Biology. Accordingly, the U.R.I. subcontract to Harvard University was terminated in September 1990, and the unspent Year-1 funds were folded into the Year-2 subcontract which commenced at Cambridge University on March 15, 1991. This move afforded closer interaction with both experimental neurobiologists, computational theorists, and outstanding vision scientists than was possible at Harvard. Since this subcontract portion of the AFOSR U.R.I. effort focuses on neural mechanisms of motion and texture vision, it will benefit from close interaction with the energetic and expert Cambridge research communities investigating the neural dimensions of visual coding. Among the related projects pursued in collaboration here are: (1) stochastic computational strategies revealed in the structure of neural times series, decomposed by Karhunen-Loeve expansion; (2) how biological neural networks distinguish between signal and noise; and (3) neural strategies for interpreting 3-D world structure from 2-D images, by use of differential motion information.

The following sections provide a summary of some recent research results, together with a description of ongoing activities.

#### (a) Relaxation Computation of Non-Orthogonal Image Transforms

It is often desirable in image processing to represent image structure in terms of a set of coefficients on a family of expansion functions. For example, familiar approaches to image coding, texture classification, feature extraction, image segmentation, statistical and spectral analysis, and compression, all involve such methods. It has usually been necessary that the

expansion functions employed comprise an orthogonal basis for the image space, because the problem of obtaining the correct coefficients on a large non-orthogonal set of expansion functions is usually arduous if not impossible. Image coding in biological visual systems clearly involves non-orthogonal representations. Indeed, from a genetic viewpoint it would be extremely costly to satisfy the precise constraints on kernel center positions, 2-D weighting structure, and overlap factors, that would be demanded for an orthogonal representation.

The receptive field profiles of visual neurons with linear response properties typically have large overlaps and large inner products, and are suggestive of a conjoint (spatial and spectral) "2-D Gabor representation." As originally proposed by Daugman in 1985, the 2-D Gabor transform has useful decorrelating properties and provides a conjoint image description resembling a speech spectrogram, in which local 2-D image regions are analyzed for orientation and spatial frequency content while preserving 2-D positional information. A fundamental difficulty in working with such representation, which had prevented their earlier exploration, is that their expansion functions are non-orthogonal (and hence their coefficients cannot be obtained by inner product projection). We have developed a general-purpose three-layered relaxation "neural network" that efficiently computes the correct coefficients for this and other, non-orthogonal, image transforms. Examples of applications in image analysis include: (1) image compression to around 0.3 bit/pixel; (2) textural image segmentation based upon the statistics of the 2-D Gabor coefficients found by the relaxation network; and (3) motion interpretation based on a 3-D Spectral Coplanarity Theorem, which uses 3-D spatio-temporal Gabor filters to solve the problem of measuring local velocity vector fields regardless of the spatial form or boundaries of the moving objects.

#### (b) Self-Similar 2-D Gabor Wavelet Representations

Building upon the relaxation network approach described above, several different schemes of image analysis and representation have been explored. Once the restrictive constraint of orthogonality has been lifted, many new approaches become possible which were previously prohibited by the lack of efficient means for obtaining the coefficients that constitute the image representation. Accordingly, we have explored several different image codes based on self-similar 2-D Gabor "wavelets," governed by a family of *generative equations* for self-similarity under dilation, rotation, and translation. All members of this family can be generated by dilations, rotation, and shifts of a single basic wavelet. A particular focus of work has concerned the trade-off between the numbers of discrete orientations employed in the representation, and the number of positions. We have shown that for a given number of linearly independent degrees-of-freedom, many good image representations can be obtained with many different variations in the sampling rules along these underlying dimensions. These developments provide new possibilities for matching an image coding strategy to the characteristic properties and statistics of any particular class of images.

All wavelet schemes, including the present non-orthogonal one, are parameterized by a geometric scale parameter  $m$  and position parameter  $n$  which relate members of the family to each other. In the classic one-dimensional case extensively studied by the mathematicians Meyer, Daubechies, Grossmann, and Mallat, all wavelets in a family  $\Psi_{mn}(x)$  can be generated from each other or from a common template  $\Psi(x)$  via the generative equations

$$\Psi_{mn}(x) = 2^{-m/2} \Psi(2^{-m}x - n). \quad (1)$$

(Conditions are now known for obtaining 1-D wavelet families which are orthogonal, infinitely differentiable, have strictly compact support, and constitute complete signal bases.) Generalizing the above generative equations to two dimensions and incorporating discrete rotations  $\theta$  into the generating function, together with shifts  $p, q$  and dilations  $m$ , the present (non-orthogonal) 2-D gabor "wavelet" set can be generated from any given member by:

$$\Psi_{mpq\theta}(x, y) = 2^{-m} \Psi(x', y') \quad (2)$$

$$x' = 2^{-m}[x \cos(\theta) + y \sin(\theta)] - p \quad (3)$$

$$y' = 2^{-m}[-x \sin(\theta) + y \cos(\theta)] - p \quad (4)$$

By using the relaxation network to find the coefficients for this self-similar multi-resolution wavelet scheme, in which 2-D Gabor elementary functions serve as the  $\Psi_{mpq\theta}(x, y)$ , we have been able to explore many new aspects of orientation-based, multi-scale, self-similar image codes. One particular application of this scheme for encoding and representing image structure is the design of an automatic system for high-confidence visual personal identification, based on encoding the 2-D Gabor wavelet coefficients from the real-time video image of a person's iris texture. This represents the first successful, very high confidence, practical system of automatic face recognition.

### (c) Neural Mechanisms for Interpreting the Dynamic Visual World: Figure Ground Segregation Based on Differential Motion Cues

Many of the problems we wish to solve in machine vision and robotics were solved millions of years ago in the natural evolution of animal species. The signal processing and computational strategies implemented in biological neural mechanisms should be of immense interest for engineering efforts to build systems which understand their environments well enough to navigate through them and recognize patterns. Even the visual capabilities of the common honeybee, whose entire nervous system contains only a few hundred thousand neurons—a tiny number compared with transistor counts in VLSI chips—far exceeds the abilities of man-made vision systems for adaptive real-time pattern recognition in an unpredictable, dangerous, 3-dimensional environment.

*new* ↓ We have assembled a neurophysiology laboratory for recording from isolated neurons in the lobula plate of the Sarcophagus Blowfly, stimulated visually with figure/ground moving texture fields. The rich visuo-motor behavioral repertoire of simple invertebrates can be elicited by motion discontinuities in the visual field. The demarcation of different velocity vector fields in the retinal stimulus array, even in the absence of any other cues (luminance, contrast, density, chrominance, disparity), constitutes a sufficient cue for visual figure/ground segregation. With physical motion through the environment over time, such discontinuities in the retinal velocity vector field can reveal the 3-D spatial configuration of the environment. We use conventional techniques for recording the electrophysiological activity of isolated neurons, and novel stimulation hardware of our own design. The electronic system is implemented in dedicated P.L.A. logic arrays clocked at 64 Megahertz, allowing the generation of multi-scale random texture fields instantiating arbitrary 2-D velocity vector fields specific to different image regions. The very fast 64 Megahertz clock permits a high frame rate of 200 HZ, which is a prerequisite for generating smooth motion displays. Both "figure" and "ground" texture fields, defined by specified boundaries, can be assigned any 2-D velocity vector field.

The identified neurons that process motion in the fly visual system supposedly do so in Cartesian vector form. Two orthogonal classes of motion-selective neurons, the V(ertical) and H(orizontal) cells, exist in the lobula plate. Each one integrates motion cues across the entire contralateral retina, in the form of vector projections onto these two basis vectors. By generating moving texture fields whose two Cartesian velocity vector coordinates can be independently manipulated, we have been able to study the interactions between the vertical and horizontal motion vector coordinates in the neural matrix. Specifically, we have measured the firing rates of individual lobula plate neurons stimulated by moving texture fields whose velocity vectors have the form  $S_1 = (v_x, v_y)$ ,  $S_2 = (v_x, 0)$ , and  $S_3 = (0, v_y)$ . All three velocity fields have different speeds and different directions of motion, but  $S_1$  and  $S_2$  have the same horizontal component of motion, while  $S_1$  and  $S_3$  have the same

vertical component of motion. Thus the classical framework for understanding the fly visual motion system predicts the same response from an **H** neuron for stimuli  $S_1$  and  $S_2$ , and the same response from a **V** neuron for stimuli  $S_1$  and  $S_3$ , despite all the differences among these stimuli in speeds and directions. We have shown that this is true over small angles, but we have been able to demonstrate definitive inhibitory interactions between the vertical and horizontal motion systems over larger angles. Thus Cartesian vector projection of the independent velocity vector components onto the **H** and **V** neural sub-systems is not an adequate model, and competitive interactions must be incorporated.

#### (d) Interpretation of Differential Visual Motion: Experimental and Theoretical Issues

An apparent paradox exists in motion-based visual figure/ground segregation. The measurement of motion in the stimulus array requires an interval of time and a region of space, with greater uncertainty in the motion estimate resulting from the narrowing of either of these. However, humans perceive motion discontinuity boundaries as phenomenally sharp, even with very small difference vectors between the two velocity fields. There should be an uncertainty principle limitation here. How can the visual system make a precise measurement of 2-D velocity vector fields, and simultaneously assign a crisp boundary in space (and in time) to the discontinuity between the two velocity signals?

We have investigated the parameters of motion-based figure/ground segregation in the human visual system. Using only velocity cues, with all other parameters of the moving texture fields identical, we have measured the magnitudes of the differences in speed and/or direction between the two motion vectors necessary to produce a figure/ground percept. These tend to be Weberian (differential signal proportional to vector norm). However, there are significant differences between the efficacy of a *speed* differential, and the efficacy of a *direction* differential, in producing a figure/ground segregation percept. We have also begun to map out the necessary and sufficient vector differentials for driving a *filling in process* based on the differential velocity vector field, versus only seeing the 1-D boundary contour.

As a challenge to existing motion models, we generated a counter-example to the popular model of motion processing based on the movement of Laplacian zero-crossings within the stimulus array. (See "Pattern and Motion Vision without Laplacian Zero-Crossings," *JOSA*, 1988.) We generated families of moving textures which, at all spatial scales of analysis, have only stationary Laplacian zero-crossings. Convolution of these spatio-temporal stimuli with  $\nabla^2 G_\sigma(x, y)$  operators, of all different blurring scales  $\sigma$ , produces an output all of whose zero-crossings are exactly stationary. Nonetheless, the motion of the stimulus is clearly perceived by human observers. We are trying to integrate these various observations about human visual capabilities into the frameworks employed in machine vision systems, specifically by developing anisotropic operators without these problems.

Finally, a parallel research project now underway concerns the perception of XOR'ed moving texture fields as two "superimposed" motion vector fields. The theoretical interest here lies in understanding the fact that the human visual system is capable of assigning not just one, but two, or more, motion vector fields to any given spatial position. How is this possible? XOR is a new form of motion transparency which we have developed, in which two 2-D moving binary texture fields are combined through the Exclusive-OR boolean operator (essentially multiplicative, rather than additive). The very fact that two independent motion fields can be seen at all is noteworthy, since such stimuli contain absolutely no Fourier motion energy in any direction. Nonetheless, the dual motion percepts are utterly salient. Thus this novel motion percept has considerable theoretical significance; no existing motion models appear to be capable of capturing it. We have begun to map out the differential vector parameters which are necessary and sufficient to drive this new class of motion transparency. It has theoretical significance for understanding how neural systems, as well as artificial systems, can interpret differential and conflicting velocity vector fields in the optic flow of dynamic visual environments.