The goal of this research program was to discover and develop real-time neural architectures capable of autonomously carrying out image processing and pattern recognition tasks in environments wherein noisy and unexpected events can occur. Such architectures are needed to cope with the fact that, in naturally occurring scenes, edges, texture, shading, size, stereo, and motion information are often overlaid and are viewed under variable illumination conditions. Special-purpose vision algorithms that can process only one of these types of information do not function well under naturally occurring conditions. The present work has analyzed a large body of data from visual psychophysics and neurobiology in order to discover and test neural principles and mechanisms whereby such a general-purpose competence is achieved by humans and animals. These designs are embodied in multi-level neural networks which are defined by novel types of nonlinear dynamical systems. The networks are computationally characterized for use both in explaining biological data about vision and pattern recognition, and in implementing novel image processing circuits for use in technological applications. Predictions of the theory have also been successfully tested in our psychophysics laboratory.
TABLE OF CONTENTS

Statement of the Problem

Research Summary
1. ART 3: Autonomous Learning and Distributed Memory Search 2
2. Global Motion Segmentation 2
3. Explanation of Short-Range and Long-Range Apparent Motion Data 2
4. A Symmetry Principle for Visual Form and Motion Perception 3
5. Visual Psychophysics 3

List of Publications 5
REAL-TIME IMAGE PROCESSING ARCHITECTURES
FOR PERCEPTUAL GROUPING, DEPTH SEGREGATION,
AND OBJECT RECOGNITION

FINAL REPORT

Professor Stephen Grossberg, Principal Investigator
June 1, 1988—August 31, 1991

U.S. Army Research Office
Contract ARO DAAL03-88-K-0088
Proposal ARO 25662-MA

Center for Adaptive Systems
Boston University
111 Cummington Street
Boston, Massachusetts 02215
(617) 353-7857

Approved for Public Release.
Distribution Unlimited.
STATEMENT OF THE PROBLEM

The goal of this research program was to discover and develop real-time neural architectures capable of autonomously carrying out image processing and pattern recognition tasks in environments wherein noisy and unexpected events can occur. Such architectures are needed to cope with the fact that, in naturally occurring scenes, edges, texture, shading, size, stereo, and motion information are often overlaid and are viewed under variable illumination conditions. Special-purpose vision algorithms that can process only one of these types of information do not function well under naturally occurring conditions. The present work has analysed a large body of data from visual psychophysics and neurobiology in order to discover and test neural principles and mechanisms whereby such a general-purpose competence is achieved by humans and animals. These designs are embodied in multi-level neural networks which are defined by novel types of nonlinear dynamical systems. The networks are computationally characterized for use both in explaining biological data about vision and pattern recognition, and in implementing novel image processing circuits for use in technological applications. Predictions of the theory have also been successfully tested in our psychophysics laboratory.
RESEARCH SUMMARY

Major progress was made on two types of problems: knowledge discovery and adaptive pattern recognition; and biological vision and image processing.

1. ART 3: Autonomous Learning and Distributed Memory Search

A new Adaptive Resonance Theory architecture, called ART 3 (Articles 5, 6, 7, 9) was designed by Carpenter and Grossberg. In common with ART 2, the ART 3 architecture is a self-organizing neural network capable of learning stable pattern recognition codes in real-time in response to arbitrary sequences of analog or binary input patterns. It does so by efficiently searching, testing, and learning hypotheses about the patterned data to which it is exposed.

In addition, ART 3 is capable of searching hierarchies of distributed pattern codes. This opens the possibility of simultaneously representing a pattern's most salient parts and wholes, and automatically reorganizing this representation based upon reinforcement feedback until it generates a maximally effective prediction. The architecture embodies new insights about how the dynamics of chemical transmitters and modulators act as a memory system capable of controlling a parallel search and hypothesis testing scheme.

2. Global Motion Segmentation

Grossberg and Mingolla (Articles 26, 27) have suggested a solution of the global aperture problem: namely, of how a coherent motion signal is imparted to all parts of a moving object, including parts that receive only locally ambiguous motion signals. This work clarifies, for example, how we can quickly see a sun-dappled leopard leaping under jungle trees. The sunshine and shadows upon the leopard's coat generate local motion signals in many directions that do not correspond to the overall direction of motion of the leopard's body. A similar competence is needed to rapidly detect camouflaged man-made objects moving in a cluttered background. Grossberg and Mingolla have suggested how these locally ambiguous signals are organized into a coherent global signal of the leopard's motion as a whole. and how challenging classical phenomena such as motion capture and induced motion may thereby be explained.

These analyses emphasize the fact that many, if not most, motion percepts are illusory percepts that can actively reorganize the data that reach our senses. Such an emphasis underscores the importance of explaining the large data base about apparent motion. We believe that such data should serve as a standard test of models purporting to provide computational insights about biological vision.

3. Explanation of Short-Range and Long-Range Apparent Motion Data

In this regard, Grossberg and Rudd have further developed this theory by explaining a wealth of data about apparent motion that other models have not been able to treat. In Articles 28 and 32, they describe further evidence for this 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 article uses the MOC Filter to explain a variety of classical and recent data about short-range and long-range apparent motion.
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; 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.

4. A Symmetry Principle for Visual Form and Motion Perception

In Articles 21 and 24, Grossberg has provided a computational foundation for unifying this emerging theory of motion form perception with our previously developed theory of static form perception. Prior to this work, these two important subjects were treated as wholly separate in the scientific literature. Our analysis suggests that the visual cortex develops to realize a symmetry principle, called FM Symmetry (F = form, M = motion). Both the Static BCS and Motion BCS can be generated by FM Symmetry as two parallel halves of a larger system which satisfies three properties: namely, it (1) computes all possible way of symmetrically gating sustained cells with transient cells and (2) organizes these sustained-transient cells into opponent pairs of on-cells and off-cells such that (3) their output signals are insensitive to direction-of-contrast. The properties of FM Symmetry suggest how the V1 → V2 → MT cortical stream helps to compute moving-form-in-depth, and how long-range apparent motion of illusory contours occurs.

An important consequence of FM Symmetry is its explanation of how the static form and motion form systems generate different geometries of perceptual space that are familiar from daily experience. In particular, cell pairs in the static form system define opponent orientations that differ by 90°, whereas cell pairs in the motion form system define opponent directions that differ by 180°. Changes in visual inputs can cause a rapid antagonistic rebound to occur between opponent members of each pair, with on-cell activation being displaced by off-cell activation. These antagonistic rebounds help to prevent smearing of percepts in response to rapidly changing scenes by resetting resonating boundary segmentations which could otherwise persist for a long time. In so doing, antagonistic rebounds can cause negative aftereffects, such as the MacKay illusion in static form perception, and the waterfall illusion and long-range motion aftereffects in motion form perception.

5. Visual Psychophysics

In Articles 41, 42, and 43, experimental tests of various aspects of the Motion Boundary Contour System theory of human motion perception were reported. Quantitative analyses of the dependence of stimulus contrast and scale on the probability of perceived direction of movement as a function of spatial separation and interstimulus interval were performed. The results verified the spatial and temporal form factors for apparent motion proposed in the theory. The data clarified the contrast and scale dependence of these form factors. Earlier data was insufficient for this purpose, because less sophisticated experimental techniques were employed in early studies, and because a very large number of trials had to be collected for individual subjects, in order to perform mathematical analyses of the data which could be related to parameters of the theoretical model.

Article 41 described experimental conditions in which the minimum interstimulus interval for perceived motion either is or is not dependent on the spatial separation between the stimuli matched by the motion detector. The fact that this minimum interstimulus interval is not dependent on the spatial parameters of the stimulus when the stimulus duration is
long was used to argue against a traditional spreading activation model of apparent motion threshold data, and in favor of a feed-forward motion receptive field model such as the Motion BCS. In Article 43, we showed that, when observer bias occurs in our forced-choice motion direction judgment task, it appears to act as a multiplicative gating of the signal-to-noise ratio of the mechanisms tuned to a particular direction of motion. This previously unnoted fact has important implications both for the construction of a correct model of the neural mechanisms involved in motion perception, and for understanding the interaction between preattentive and attentive processes involved in motion perception.

Further modelling studies in vision analysed aspects of visual detection that depend upon stochastic properties of neural processes. In Article 39, a theory of threshold vision was proposed that explains a greater amount of the visual sensitivity database as a result of light fluctuations than has previously been thought possible. The article proposes a statistical theory of contrast encoding by the visual system that reduces to a standard model in the deterministic limit. In the paper, recent mathematical theory and techniques developed in theoretical neurobiology are applied to the explanation of visual phenomena for the first time.
LIST OF PUBLICATIONS
Contract ARO DAAL-03-88-K-0088
June 1, 1988—August 31, 1991

BOOKS

ARTICLES


7


* Also supported in part by the Air Force Office of Scientific Research.

% Also supported in part by British Petroleum.

& Also supported in part by DARPA.

© Also supported in part by Hughes Research Laboratories.

+ Also supported in part by the National Science Foundation.

# Also supported in part by the Office of Naval Research.