Synchronization Opponent Networks: Dynamics, Computation, and Coding for Similarity and Object Recognition

by

David Lee DeMaris, B.S.E.E., M.S.E.

Dissertation
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin
May, 2001

Synchronization Opponent Networks: Dynamics, Computation, and Coding for Similarity and Object Recognition

Approved by
Dissertation Committee:

Baxter F. Womack, Supervisor

Dedication

To my family
for their encouragement,
and especially my mother
for nurturing my early
interest in science.

Acknowledgements

I thank my adviser, Baxter Womack, and committee members: Margarida Jacome, Risto Mikkullainen, Michael Marder, and Alan Bovik. Dr. Womack has of course been my greatest supporter from the start, when I had more enthusiasm than understanding of the complexity of what I was proposing. Gerhard Werner appeared at the last moment in my graduate career and has served as a great mentor, reader and foil for these ideas, enriching my graduate school experience considerably. Joydeep Ghosh gave me an initial online home in ECE; his suggestion to start with something simple led to the family of experiments on parametric curves. Bill Geisler suggested that I look into ambiguous perceptions, leading to an eventual publication in that area. Jake Abraham's vision class really gave me the opportunity to write code, some of which survived into the latest versions of the system. I thank Risto Mikkullainen for his attention to the work and hope we can find a way to cooperate further in the future, perhaps on improving the learning aspects of the system.

I've benefited from a good deal of interaction with researchers outside the U.T. community. Rosalind Picard stressed the need to write more critically, after generously taking time to act read my MS thesis. Fred Abraham, George Mpitsos, Daniel Levine, Vincent Brown, Mark Styvers, Bill Baird, William Calvin, Chris Langton, Ichiro Tsuda, Kuni Kaneko, Hava Siegelmann, Ben Goertzel, Melanie Mitchell, Dean Driebe, Chai Wah Wu, Todd Leen, Stephen Bressler, and Cris Moore have all provided helpful suggestions and opportunities to discuss this work in progress. Shimon Edelman and Bartlett Mel were very helpful as I tried to compare our work; it should be obvious that Dr. Edelman's comprehensive treatment of similarity and representation spaces came to be a major influence on this work. Many IBM managers have supported me in these studies, most recently and generously Paul Bassett. An IBM colleague in the late 80s, Jean-Claude Peréz introduced me to coupled maps in his own work and by giving me preprints from Kaneko; this made a large difference over time in my approach, in classic butterfly effect style. Melanie Gulick has been a wonderful help negotiating the paper mazes. Brenda Adrian helped out with copy editing and support of both the tech and emotional varieties in the late stages of preparing this thesis.

A few Matlab routines from Eero Simoncelli, and from Malcom Slaney care of Interval Research were used in this work. Michael Tarr of Brown University provided the image data used in the rotation in depth experiments.

Synchronization Opponent Networks: Dynamics, Computation, and Coding for Similarity and Object Recognition

Publication No._____________

David Lee DeMaris, Ph.D.
The University of Texas at Austin, 2001

Supervisor: Baxter F. Womack

Shape similarity, categorization and recognition of objects from outline shapes are addressed with a time varying coupled logistic map lattice trained by a genetic algorithm. Synchronization and spatial cooperation are key underlying principles, thus the network is designated a synchronization opponent cooperative activity (Soca) network. The computations are described in terms of well known frameworks – the dynamical recognizer and probabilistic finite state automata.

The network orders curve exemplars of parameterized curves by constructing a *partition cell metric space*, in which partitions of the network's dynamical phase space are considered as dimensions and the occupancy of all partitions locates an in the representation space. This statistical population representation contrasts with many connectionist representation spaces, in which dimensions capturing statistics of the modeled world are bound to individual output or internal nodes, even in a distributed representation. The latter style -*local* or *place coding* - remains the most widespread assumption in neuroscience and connectionism, but is increasing questioned by the neuroscience community.

Next I study a family of such networks acting as classifiers for paperclip objects rotated in depth. Recognition rates up to 85% are obtained for a set of 20 objects. The approach builds on the theory of view based recognition, but posits an alternative spatiotemporal computation and population code whose computation serves to combine local features (binding), capture local structural relationships, and handle view invariance.

The work raises principled issues in biological computation. It departs from previous recurrent networks and high dimensional chaotic networks by introducing two stages with sharp parameter changes, rather than stationary or smoothly changing dynamics. It is hypothesized that slow wave dynamics and oscillations observed in inferotemporal cortex are signatures of such a recognition process. Extensive neuroscience review, justification of the CML formulation via lower level models, and discussion of experiments to discriminate various hypotheses are included.

Table of Contents

List of Tables

List of Figures

List of Symbols

b	Bifurcation parameter
c	Coupling parameter
$b1$	Bifurcation parameter in first Soca stage (desynchronizing)
$b2$	Bifurcation parameter in second Soca stage (synchronizing)
$c1$	Coupling parameter in first Soca stage (desynchronizing)
$c2$	Coupling parameter in second Soca stage (synchronizing)
H_c	Cross entropy between partition cell distributions
H_s	Shannon entropy between partition cell distributions
k	Number of partition cells or bins
n	Number of coupled neighbors of a lattice site
N	Number of lattice sites, number of objects
S	State variable of logistic map
v_I	ith partion cell from vector of CML states
W	Weight multiplier in objective function
x	Spatial index in neighborhood computations
y	Spatial index in neighborhood computations
α	Word in some alphabet A*
ε	An incremental distance (representation space or synchronization)
π	Probability (transition probabilities)
ω	Last word in some alphabet A*

Chapter 1: Introduction

Classifying object shapes and identifying the category of a known or previously unseen shape most similar to a prototype are fundamental tasks in computer vision. Successful solution of these tasks is a prerequisite for higher levels of object and scene analysis in still images, and for object tracking in video sequences. Research into the problem is increasingly motivated by new applications in content-based retrieval for image, trademark, and video databases, in addition to traditional industrial inspection, biomedical, and target identification applications. In multimedia applications, it is often particularly important to conform to the human judgments of similarity, and to have a representation which can be matched with efficient computations.

While such practical interests are one driving force, the questions of similarity and object identification are of intrinsic interest from a psychological and neurophysiological standpoint as well; distinct perspectives on these problems have evolved within these disciplines, and will be given considerable space here. Computational and signal processing concepts have often informed these fields, framing the nature of experiments and interpretations of results.

In this thesis, the related problems of similarity and stimulus equivalence are addressed through a method involving the construction of a metric space by partial synchronization of states, under the twin influence of the intrinsic dynamics of the oscillators and of the shapes presented as initial conditions. The current form of the system is a hybrid of algorithmic processing for learning and recognition, with the construction of the space and the resulting recognition algorithm using a regular lattice of chaotic oscillating units with a local mean field coupling.

MARR'S PROGRAM IN VISUAL NEUROSCIENCE

Because the visual system has been studied most intensely, the mutual influence of theory and experiment in that realm has dominated thinking about neural mechanisms and sensory problem solving. Marr described three levels [1] of abstraction for the study of vision (Marr 1982) . The lowest level is the **mechanism** level, concerned with details of what the physical elements of the brain (or computer are doing). Historically, neurons have been taken as the elementary units of interest. The next level is **algorithmic**: a description of the process controlling the hardware. Finally, the highest level is the **computational** level, in which the problem should be understood in terms of information processing: what is being computed, why, and what are the appropriate models for these operations?

Marr produced such a computational theory for object recognition, in which information is reorganized by flowing through a series of parallel and serial modules or processing stages. The raw, two dimensional array of image intensity first goes through a

---

[1] The outline of Marr's work here follows closely the synopsis in (Yuille and Ullman 1990).

process of edge-detection to produce an intermediate representation known as the *primal sketch*. This representation undergoes further processing to produce a viewer centered representation (the $2\frac{1}{2}-\mathrm{D}$ sketch), Finally, an objected centered 3-D representation is computed, describing the objects in terms of volumetric primitives. Low level vision is concerned solely with the first two levels; interpreting and recognizing an object from a sketch is generally considered as high level vision. The notion of object centered representations has proven controversial; I will examine the controversy in some detail. In this thesis, I have adopted the *viewer-centered* approach.

Because the Marr program is especially influential and successful in replicating the types of function thought to be performed in *low level* vision, many researchers have tried to extend it to high level vision, but chiefly using the assumptions concerning *mechanisms* adopted from low level vision. These mechanisms are chiefly a “front end” of filters employing rate coding, with “back ends” consisting of specific arrangements of excitatory and inhibitory neurons, doing the work of the algorithms within a computational module. While the general framework remains valid, the role of detailed structural arrangements of individual neurons will be questioned in this thesis.

I will now present an overview of the problem - the computational analysis - and a brief treatment of the mechanism or implementation level.

SITUATING SIMILARITY AND OBJECT RECOGNITION

The over-arching problem chosen for study here is similarity, with a particular focus on similarity of curves and silhouettes. By silhouettes, I mean two dimensional projections of three dimensional objects, with all shading data removed. These binary images are essentially closed curves in the plane, with the enclosed region filled in. Considering curves, the similarity problem can be simply but generally formulated as:

Given planar, non-self-intersecting closed curves C1 and C2:

Define a distance function d(C1,C2) which returns a scalar value indicating their distance in some high dimensional space.

This formulation allows for geometric approaches relying on knowledge of explicit coordinates, as well as approaches which operate on the instantiation of the curves in an image. These approaches will vary according to various criteria for effectiveness, for sensitivity to noise and distortions, and with regard to plausibility as biological models.

In both machine vision and the psychological study of similarity, it has been common to construct a *metric space* and to embed objects as points in the space. The distance function defined above can then be computed as standard Euclidean distance, or by a variety of more sophisticated weighted distance functions. Non-metric methods of assessing similarity exist as well, and will be briefly mentioned; these emerged chiefly in psychology, in response to evidence that many aspects of similarity in cognitive and perceptual phenomena may not be metric.

Similarity is seen by cognitive science and psychology of vision as tightly linked to segregation and grouping processes in visual scene analysis, to identifying objects seen from different viewpoints, and to categorization. Thus, it underlies nearly all higher level cognitive phenomena regarding our processing of the visual world. In everyday life we spontaneously identify objects, and create categories from the diverse retinal images of an object seen from different viewpoints, or from the diverse individual members of a species. However, this fundamental problem of stimulus identity has only recently begun to be addressed satisfactorily in computer vision. While many geometric methods have been developed for handling translation and rotation in the plane ((Simoncelli, Freeman et al. 1992); ((Wolfson and Yehezkel 1992), rotation *in depth* of diverse objects has been addressed most recently and successfully through statistical approaches with a rich feature space (Mel 1997), and by an ensemble of radial basis function neural networks implementing a view normalization and interpolation strategy (Edelman 1999).

RESTRICTIONS ON THE PROBLEM SPACE

For exploring, the power of a class of dynamical networks on problems of similarity and stimulus invariance, I have chosen data limited to isolated outline objects (silhouettes), i.e. binary images in the plane. I do not address more general problems of segmentation and scene analysis; clearly extensions to this method and extensive preprocessing would be required to account for those aspects of visual processing.

The approaches just mentioned developed by Edelman and Mel for object recognition also use isolated objects, but include shape, shading and (in Mel's SEEMORE system) color. These additional cues may increase the performance of those algorithms, which have not been tested on silhouettes. Humans clearly can recognize common objects from outline information only (Hayward 1998), and the focus in this thesis has been on that domain.

Nevertheless, including shading data would allow discrimination of certain objects from views which occlude (hide) parts of their own structure. With silhouettes only, it would be impossible to discriminate a round ball from a clarinet seen looking directly down the bell, for example. Shading information would allow the convex and concave nature of the two objects to be discriminated. I do attempt to handle objects which present radically different views; the following figure illustrates an extreme view of one of the test objects in a family of recognition experiments.

Fig. 1. Two views of the same paperclip object, illustrating the extreme nature of the distortion due to rotation in depth. The left view is the 0°, the right view is +90°. Base images provided by Michael Tarr, Psychology Dept., Brown University.

In many object recognition studies, the performance of images degraded with noise and with distortions such as scrambling or occlusion is examined.
Examining performance in the presence of image noise is particularly important for a complete scene analysis system attempting noise sensitive processes such as segmentation.

At the time of writing, the approach here has not been tested with noise degradation, and I will assume segmentation is handled by pre-processing which must be noise tolerant and ideally has some noise suppression ability.

AN OVERVIEW OF THE METHODS

The use of a metric representation on a space of extracted features, or on higher-order features discovered by a learning process, is one of the oldest methods in machine vision, and in pattern recognition in general. The method developed here is novel in that the dimensions of the space are derived from phase space partition cells of a dynamical system[2]. Stated differently, the representation or encoding is based on the *statistics* of the *total network state* in the constructed space, without reference to the individual nodes in the network. This is in contrast to most connectionist representation spaces, where the dimensions of the space capturing the statistics of the modeled world are bound to individual nodes, whether these are internal or output nodes whose values constitute a code. The linkage of representation dimensions to nodes is referred to as *localist* or *place coding* and remains the most widespread assumption in neuroscience and neural network theory.

---

[2] This is translated roughly as non-overlapping intervals in the set of possible real valued states, with each node in a network having a state in only one such interval.

The present system and network model assumes an alternative spatio-temporal population code, and achieves some functional computing properties which help to overcome classical dilemmas in pattern recognition. Issues of combining local features (binding), capturing structural relationships at slightly larger scales, and handling view invariance can be addressed simultaneously by the spatio-temporal interactions that occur during learning and recognition in this system.

In addition, the work raises principled issues of computation in biological systems. While rooted in theory of recurrent networks, chaos, and complexity (high dimensional coupled chaotic systems), it makes some substantive departures from previous work in these fields. The most important is that it introduces two stages with a sharp change in the parameters, rather than stationary or smoothly changing dynamics. By some definitions, this precludes its consideration as a dynamical system altogether and certainly makes any analysis based on continuous mathematics more difficult.

The present work also breaks with typical practice in recurrent networks by concentrating on the *transient* regime of dynamics rather than on equilibrium states (i.e. attractors) of the network. In other words, the network is measured and finishes its work *prior* to reaching any stable asymptotic state. This staged processing, with desynchronization and partial synchronization of transient trajectories is designated as a *Synchronization Opponent Cooperative Activity* (**Soca**) network.

To date the study of high dimensional, spatially coupled nonlinear systems is chiefly experimental, with the experiments conducted by numerical simulation. Given this fact, there is little existing theory to build rigorous proofs of the system's capability, bounds on performance, or expectations on memory capacity and scaling. The most relevant recent theoretical developments are presented briefly, though none are directly applicable in their present form. The present study, like the bulk of spatially-extended network studies, is exclusively computational in nature.

OPERATIONS AND ARCHITECTURES FOR COOPERATIVE NEURAL COMPUTATION: AN OVERVIEW

> *The application of the Turing scheme to describe neural computation in a real brain is not completely obvious … Where is the program in the brain? And what is a memory? If a program exists, its mere definition will, in my view, be a revolutionary step toward the understanding of brain function. The mere demonstration of the existence of a program is beyond what seems imaginable. And all we have are noisy neurons and unreliable synapses.* (Amit 1995)

For the purposes of this introduction, let us note two observable properties of a *chaotic* system. It is a deterministic system charactized by some scalar or vector state *S* which

1. exhibits an aperiodic time evolution (called its orbit or trajectory)
2. exhibits exponentially rapid divergence over time for nearby initial conditions.

We can model such systems with differential equations or difference equations. Using difference equations, chaotic behavior can be generated with a one dimensional system – a system with a single state variable. Such a system, with a single time varying state variable, can be considered as a simple model of an *oscillator*.

If we want to compute similarity of inputs based on the state values of a system observed at some later time *t*, chaos seems to be the very opposite of what we want. Only early in the evolution of a chaotic system, in the *transient* stage, is there reliable correlation between the input and the state of the evolving system.

We can construct *higher dimensional* systems by connecting two or more such systems as nodes in a network. At every time step, an averaging function results in mutual influence of the state values, counteracting the divergent tendencies of chaotic dynamics. This coupled high dimensional system behaves quite differently; depending on the "strength" of chaos and coupling, the connected oscillators may *synchronize* in an aperiodic or periodic mode.

In this case, there are rather intricate dependencies in the approach to synchronization on the specific spatial form of the input to the system. The *basic intuition* underlying this thesis was the idea that by exploiting these two opposing tendencies, of divergence and synchronization, a high dimensional system may effectively compute a representation suitable for use as a distance function.

Since an important goal of this thesis is to extend and solidify the connections between high dimensional nonlinear dynamics, computational approaches to pattern recognition, and biological networks, it is necessary to understand developments in several fields which have motivated the approach taken. The relationship between arrays of chaotic network elements, programs and neurons would seem even less obvious than the relationship of programs and neurons.

I will argue that a more diverse, and ultimately clearer, picture of neural computation is emerging from theoretical areas such as synchronization dynamics of ensembles of coupled chaotic elements, and the field of *symbolic dynamics* which bridges classical computing concepts (symbols and formal languages) with dynamical systems theory. I will also argue that a wealth of experimental evidence in neurophysiology supports this view, even though (with a few exceptions) this has not been the interpretation framework used by experimentalists working on synchronization dynamics [3].

This style of computation differs from the classical connectionist models, including attractor neural networks, in at least one significant way. *The structure supporting the computation is implicit in the graph corresponding to the flow of states, when a particular partitioning (coarse graining) of the state space of network elements is chosen.* This state flow graph must adapt to both the statistics of the sampled world of patterns, and must also support a cooperative transformation of the input, such as the normalization function developed here in the context of solving the stimulus equivalence

---

[3] There is some evidence that this is changing. W. Singer, a well known investigator of synchronized oscillations, stresses that *aperiodic* oscillations may be synchronized in a recent review. The work of Hampel and Sompelinsky, reviewed later, is an outstanding exception.

version of the similarity problem. Some similarity of this computational structure to Markov chains and decision processes is apparent, but again the present emphasis on *changes* in the dynamical parameters precludes an obvious mapping of techniques from that field, which chiefly models random processes with stationary probability transitions between states.

It could be argued that the statement highlighted above is too strong; recurrent network architectures *in general* do their work through complex state flows controlled by weights. However, these still depend on fine structure of weights and specific connection topologies, while the networks studied here do their work with homogenous local connections between units in a single layer.

With this iterative computation style and corresponding implicit state flow graphs, networks with few (i.e. 6) parameters yield effective computations for a perceptual task. These parameters are spatially homogenous across a spatially regular, locally connected network, rather than requiring a specific topology of connections and weights fed by feature detectors. These parameters correspond to *average parameters* of micro-circuit component systems tolerating a great deal of randomness in structure and noise in their local, neuron level operation, because computation is carried out by hierarchies of oscillating sub-populations.

This is not to say that classical ideas of receptor fields, and of neural coding and computation mediated by single cell rates are unimportant, as they are relatively simple and predict the responses of superficial neurons to simple stimuli. However, in the emerging view, these classical concepts interface with more dynamically complex computational systems, functioning in part by rapid coordination between cortical regions and between cortical and sub-cortical areas. In sensory pathways, these simplest computational stages may act as *perturbations* to higher level computational systems supporting segmentation, object recognition, and attended search functions. In motor systems they serve as final actuators to the muscular systems, driven by complex dynamical pattern processes in pre-motor systems.

Still, even the classical concepts are being questioned and revised, based on a variety of evidence which I will review in some detail. Omitting citations for the moment, the evidence pointing to a revision of classical ideas on neural signaling and computation includes:

1. Modulations of response profiles from natural stimuli, including areas outside classical receptive field.
2. Multi-channel spike rate and field potential studies indicating a role for temporal synchrony and modulations in synchrony.
3. A changing role for dendritic action from passive conductance to an active role in processing, related to complex spike arrival time processing and synchrony.
4. Interactions between local fields (electrotonic coupling) and dendritic processing, again related to spike arrival time processing and synchrony.
5. Indications that the same neurons perform different functional roles in stages as the time course of perception unfolds.

One way to view the historical interactions between experimental neuroscience and neural network theory is to consider that ideas about the elementary operations computed by neurons condition ideas about network architectures, which in turn condition experimental approaches. For example, in the history of neural modeling, neurons were first considered as binary threshold units, combined into networks functioning as boolean logic systems. Experimentally observed excitatory and inhibitory synapses, along with the connectionist concepts of weights, advanced the neuron level computation model to support a variety of architectural approaches, shifting the architectural dialog towards signal processing, and to the mapping of learned environmental regularities to representations or behaviors. Competitive principles came to play a large role in network architectures.

The new operations conceived here derive from a research stream rooted in experimental and theoretical biology of neuron populations[4]. Small microcircuit models (Chapeau-Blondeau and Chauvet 1992), small living neural circuits maintained on a silicon electrode array (Kowalski, Albert et al. 1992) and living sub-networks in awake behaving animals (Freeman 2000) have all been demonstrated to show complex, time-varying population firing activity. Observation in living systems and detailed modeling of increasing scales of populations can result first in synchrony, then eventually chaos (by coupling synchronous excitatory and inhibitory *sub-populations* with incommensurable frequencies).

Perhaps finally (but awaiting explicit experimental confirmation) sequences of desynchronization and partial synchronization emerge as “populations of populations” are coupled, mediated by slower waves or impulses of activity *from spatially separate regions* which serve to modify dynamical parameters. The idea that slow wave rhythms in the brain act as a kind of clock was proposed long ago by Wiener (Wiener 1985), but a clock was conceived of as a kind of “gating” operator on signals as in digital logic. An alternative view expressed here is that such rhythms act as a clock controlling synchronization operators by modulating bifurcations in a nonlinear system[5] Synchronization and clustering operations can implement competitive interactions, but their hallmark is cooperative effects, resulting from the interaction of spatially organized input with the geometry implicit in the dynamics. Binding of disparate sensory features into a unified code which preserves compositionality – the ability to recognize sub-parts – is an important possibility for these new primitive operators. The binding problem has so far eluded a satisfying solution by localist, rate coded feature combination hierarchies (i.e. gnostic or grandmother cells), and by feed-forward and attractor networks, due to the compositionality issue. This issue will be taken up in a later chapter in some detail.

These new synchronization operators and the corresponding strategies for coding and computation operate in 16 iterations or less, an upper bound chosen to correspond to plausible biological recognition times (assuming certain spatial and temporal scales of neural computation). Simulations described in this thesis, and by other investigators, indicate that rapid synchronization can occur with coupled discrete oscillators. Emerging

---

[4] While there is a literature on chaos in single neuron responses, I will not discuss it here.

[5] Baird earlier argued that discrete time clocking dynamics in cortical assemblies establish (fast) entrainment and (slower) bifurcation frame rates, with Hebbian learning occurring during the latter.

theories on ensemble density evolution in chaotic systems and on parameter lower bounds guaranteeing synchronization may provide a deeper mathematical explanation for the time course of transients and synchronization, and will be introduced briefly.

In addition to the present work demonstrating that rapid synchronization of transient responses governed by chaotic dynamics has applications in shape recognition, many investigators have been exploring similar systems for the related vision tasks of segmentation and grouping. The generality of the computational strategy with chaotically evolving fields of coupled state variables, suggests that such dynamics and conceptual approaches, while originally suggested for large scale "mass action " models, may also apply to smaller networks. The state variables in large scale models may be *ensemble average firing probabilities or spike rates*, while in microcircuit models (where each node represents a single cell) the state variable can be mapped to *spike firing phase* relative to some reference slow cycle, for example, resulting in local phase distributions which can rapidly perform computations. In the rest of the thesis, I will adopt the convention of discussing firing rates, but the flexibility of these principles is a rather important point.

Historically, pioneering researchers in psychology and neuroscience have emphasized the *differences* between dynamical approaches and computational or symbolic approaches. I will be more conciliatory, following the lead of a research community sometimes termed "physics of information", which examines dynamical systems as information processing systems, often using the tools of symbolic dynamics to bridge these domains.

ORGANIZATION OF THE THESIS

The organization of the thesis reflects several goals beyond the description of a pattern recognition algorithm, chiefly the grounding in psychology, neurobiology and the justification of the coupled map approach as an appropriate technique in computational neuroscience.

In this introductory chapter and continuing in the next, I describe the selected problem of similarity and pattern recognition, emphasizing well-known dilemmas which are negotiated in the present study with a combination of novel approaches and existing methods.

Chapter 2 reviews psychological and computer vision approaches to the issues of similarity, as well as reviewing a selected set of algorithmic approaches to shape recognition. These algorithms were chosen for review from the vast literature on the basis of recency and some family resemblance to the present method. Finally, a section on pure computer science theory relating dynamical systems to pattern classification is raised in this context.

Chapter 3 reviews some neurophysiology and systems neuroscience which is especially relevant to the systems approach followed here. The emphasis is on recent findings and controversies over the role of single neurons versus larger assemblies in coding and computational tasks. Further, I explore the possible nature of computations in such assemblies and at various scales of organization.

Chapter 4 will more systematically introduce the required dynamical concepts. Since this dynamical framework at first appears remote from conventional neural network modeling, I discuss dynamics in the context of larger scale networks and the concepts of macrostate or ensemble variables. Some work of other investigators in dynamical networks is reviewed, emphasizing the visual task of segmentation where most previous work in oscillations has been focused.

Chapter 5 discusses representation and learning. I begin with some discussion of how I approach the problems of similarity and stimulus equivalence with this dynamical framework, then proceed to computation and framing the learning problem.

The dynamical complexity of recurrent, spatially extended systems poses challenges for analytical determination of network parameters, though some recent progress on related issues was described in Chapter 4. As a result, in order to construct network dynamics, the *evolutionary computing*[6] paradigm is used to construct networks. The emphasis on this study is more on dynamics than on learning per se, so the treatment of this aspect of the work is brief.

Chapter 6 describes the computational experiments undertaken. First I describe a variety of simple but novel studies of transients and convergence in single maps, pairs of coupled maps, and lattices with random or simple algebraic initial conditions. Most of these studies examine system responses over multiple parameter and stimulus dimensions, with 3-D animations of the system state versus a parameter plane produced as a result. These animations are provided as an internet based supplement.

Next, I describe a family of experiments involving the construction of a quasi-metric space which serves to order a family of parametric curves. The motivation for using this toy problem is that these visual forms can be unambiguously ordered in a way that maps naturally to their perceptual appearance. This is rarely the case for natural objects. The resulting family of images for a curve, considered as distortions of a canonical image, also resemble distortions due to scaling and rotation in depth of a single object.

Finally, an extensive family of experiments on recognizing the equivalence of different views of objects rotated in depth is described. This problem is formulated in two ways to follow experimental paradigms used by other researchers; one is a database search problem, with a second set following a psychological paradigm of determining whether two successive presentations are the same object or different. The *paperclip* images and variants used here have been used in both psychophysics and neural recordings, proving challenging to monkey and human subjects, with error rates ranging from 30% to chance depending on training circumstances, indicating that the problem is non-trivial. For the stimulus equivalence problem, the objective function used during evolutionary computation involves separate terms to achieve a balance between clustering different views of an object and avoiding the mapping of views with similar local statistics to the same region in the representation space. The concepts of

---

[6] While the term genetic algorithm is more widely used, some authors (e.g. J. Pollack) restrict that term to bitwise parameter encoding and mutation which is blind to "gene" boundaries, with evolutionary computing subsuming both blind mutations and gene-boundary-aware mutations.

regularization or normalization and cross-entropy are introduced. The results of the hybrid (network and algorithmic) system are compared with two other state of the art computational approaches.

Chapter 7 is devoted to discussion, future work and conclusions. I highlight some of the limitations of the study and system, suggesting possible improvements from the standpoint of both engineering and biological plausibility. I revisit competing theories of biological coding briefly, and provide a new interpretation and computational role for synchronous observations and modulations of synchrony (rate correlation) seen in multi-channel neuron experiments, and suggest neurophysiological and computational work. Finally, I summarize the main contributions and conclusions of the thesis.

Appendices are included introducing several technical topics: formal language theory, signal processing, and Markov chains.

Chapter 2: Similarity and Object Recognition

This section critically surveys historical and recent work related to the topic, with the aim of situating the present theory in various ongoing dialogs related to effective pattern recognition and the nature of biological perception. Given the heavy interactions between these disciplines, the assignments are perhaps somewhat arbitrary, but are simply intended to reflect the organization of the literature. Similarly, occasional references to other disciplines or technical topics covered in depth in later sections seem to be inevitable; it is hoped that the reader will exercise patience, and expect that such material will be approached again in a later section.

PERSPECTIVES FROM PSYCHOLOGY

The wealth of data from functional brain imaging has presented challenges to cognitive psychology and artificial intelligence, which for their first decade or so could presume that there might be universally effective algorithms in a given domain of perception or cognition. Given the distributed nature of processing revealed by imaging, a rough consensus exists across psychology and neuroscience that, for any given domain of cognitive and neural processing, diverse mechanisms come into play depending on the exact nature of the task faced by the organism. Cognitive processing is attributable to moment to moment shifting between large-scale brain states that interact with and control diverse linkages of distributed and local processing networks, not simply a sequential flow of information through static modules. Thus there is no *universal* processing flow for perception which is independent of the time course of the information presented, nor from the task context in which a perceptual act is embedded. This is stated only to reinforce the appreciation that such changes in the nature of the task may induce a different cognitive or neural network architecture, subtle changes to the task may change the picture considerably.

To be concrete, much of the literature on visual recognition discussed below is focused on rather contrived conditions of matching objects presented briefly in sequence. This task differs from a natural ecological embedding condition, involving access of short or long-term memory representations of objects with some meaning to the organism, and searching for them in a natural scene. Thus observations at the levels of psychophysics and neural activity may give only limited insight into other modes of object recognition. For match-no match tasks on rapidly presented objects, *preattentive* neural dynamics, which can rapidly make discriminations with minimal involvement from more complex representations, are likely to dominate. The psychological meaning of similarity in such a task setting and in the context of object recognition, will differ from other classic work, such as Gestalt era studies of similarity (Goldmeier 1972). In the latter, subjects make choices at their leisure between various drawings; their judgements shed light on perceptual issues of the interaction and dominance of aspects of form such as orientation, size, and spacing on grouping processes. Whether the same processes come into play in very rapid recognition process is an open question, but most of the computational procedures to be examined here do not focus on issues of grouping, or other areas of figural processing which includes phenomena such as length distortions, illusory contours and size distortion illusions.

Similarity and Metric Spaces

Similarity plays a fundamental role in theories of knowledge and behavior. It serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations. Indeed, the concept of similarity is ubiquitous in psychological theory. It underlies the accounts of stimulus and response generalization in learning, it is employed to explain errors in memory and pattern recognition, and it is central to the analysis of connotative meaning

Similarity has a long history in mathematical psychology, with two major branches: set-theoretic and geometric. The emphasis in the present work is on *geometric similarity*. In this formulation, objects are identified with points in a space, with categories corresponding to volumes in the space. The *dimensions* of the space are identified with primitive features in the input space. Similarity is conceived of as proximity, and a space which supports a *distance function* or metric is a metric space.

Three properties serve to define a distance function as a metric:

1. The *identity* property asserts that an object should be most similar to itself.
2. The *symmetry* property asserts that the order of presentation should not affect the measure.
3. Finally, *triangle inequality* should be satisfied; two dissimilar shapes should not both be similar to a third.

While these properties must be satisfied to conform to the mathematical definition of a metric, it is less clear that they are relevant to human perceptual processes. Symmetry, in particular, does not hold; but examples where it fails are most readily drawn from the realm of semantic or conceptual constructs (Tversky 1977). To better model semantic information and the symmetry violations, Tversky proposed a set-theoretic framework which counts shared features and independent features to produce a quantitative similarity measure, but dispenses with the notion that the set properties need to be ”dimensionalized” .

Proponents of *prototype based categorization* for shapes have argued that this asymmetry is due to the fact that prominent features of a shape in memory establish it as a prototype, and the absence of this feature can quickly be detected. Edelman (Edelman 1999) argues further that the critique of metric similarity at verbal and conceptual levels is of limited relevance in the assessment of geometric objects, where the objects can be decomposed into objective primitives. The same assessment was made by Tversky and Hutchinson (Tversky and Hutchinson 1986), stressing that this is particularly true when physical stimuli involve a small number of dimensions. When a larger number of feature dimensions is involved, metric models become problematic and the set-theoretic alternative performs better. Uttal, in a review of similarity and categorization, stresses that the boundary between perceptual and semantic representations may hinge on the number of dimensions (Uttal 1988).

The metric space concept was enhanced by the addition of a local density factor by Krumansl (Krumhansl 1978), in a way that compensates for some of the weakness noted by Tversky. In this formulation, the distance between objects is modified by the number of nearby neighbors in the space so as to increase distances in densely populated areas.

Measuring Similarity Experimentally

How can perceptual similarity be measured in a psychophysical setting? A variety of methods have been developed.

When subjects are asked to perform tasks, three major methods of experimentally assessing the perceived similarity or similarity of representation are seen in the literature. *Reaction time* (also called judgement time) are measured, but in some cases short reaction times are interpreted as a sign of similar internal representation (Bower and Clapper 1989), while others interpret long reaction time as confusability (Mumford 1989). *Error rates*, particularly false positives or confusions, are often interpreted as indicators of stimuli sharing similar representations. A final method is to measure *interference*, which is essentially reaction times or error rate effects of attending to, ignoring, or holding in working working memory multiple stimuli.

Levels Of Categorization and Object Recognition

Similarity, in the broad context of category formation and object recognition, is normally framed in a discussion of the properties or features constituting a category. Seen broadly, several problems must be solved for effective and flexible object recognition[7] (Tarr 2000). One is the multilevel nature of categories. Very broad distinctions suitable for concise representation, naming and rapid recognition define the *basic* level of categories (Rosch 1975). *Subordinate* level categories are more fine grained and require more time for naming. The categories ‘birds’ and ‘humans’ are entry level, while ‘blackbirds’ and ‘song sparrows’ are subordinate categories of bird. A further level of categorization is the individual or exemplar level, where an individual sparrow or human could be identified.

The term *entry* level refers to the level that is accessed first and typically named when a subject encounters a familiar object. This is normally the basic level, but for some categories (notably faces) the subordinate or individual level is the entry level; for anomalous objects in a basic category, such as penguin among birds, the subordinate level may be named.

Another related issue in recognition is the variability in viewing conditions for objects. Objects may be obscured by other objects (occlusion), or parts of an object may

---

[7] The categorization section follows Tarr’s review closely; the review provides a concise statement of issues and recent work including some relevant imaging and single neuron studies, but focuses on the structural vs. view based controversy to the exclusion of other important issues such as visual search, and overlooks additional single and multi-neuron studies which will be addressed in a subsequent section here.

shield other parts from view (self-occlusion). For non-rigid objects, recognition must account for variation in the configuration of parts. Viewpoint changes affect the retinal image of an object; changes in size, position in the plane, and rotation in depth.

Categorization level interacts with object recognition, implying the likely prospect of multiple subsystems. There is evidence that the ability to compensate for viewpoint changes depends on the categorization level (Edelman 1995). Subordinate level discriminations – those between very similar objects – increase the costs of recognizing unfamiliar views.

Another form of interaction is that certain stimulus classes, notably faces, seem to be interpreted chiefly at the subordinate level. There is a long history of claims of specific face detector neurons (Rolls, Baylis et al. 1989). Sub-regions of inferotemporal (IT) cortex, the putative cortical high-level pattern recognition area, were shown to be more active during face recognition tasks (Sergent, Ohta et al. 1992). However, by synthesizing a novel class of stimuli and intensely training subjects to make fine discriminations between members of this class, Gauthier and Tarr (Gauthier and Tarr 1997) make a strong case for an alternative interpretation: that it is chiefly *stimulus expertise* which results in a specific syndrome of configuration sensitivity and automatic assumption of the subordinate level in recognition, rather than the specific stimulus category. This may still be associated with specialized regions of IT cortex; the imaging research of Gauthier and colleagues suggests that such localization does occur (Gauthier, Anderson et al. 1997).

A final area of interaction between categorization levels and recognition involves the type of representation or features used in recognition. Structural description theories assume that *view-independent* or i*nvariant* features underlie the representation of objects, and there is evidence that features such as the major axis of a 3-D shape are used to make entry level categorizations. View based theories build representations from various local features extracted from separate learned views, and have historically been associated with the subordinate level. Advocates of *view-based representations* have recently claimed that basic level categorization can emerge in a natural fashion from the clustering involved in making subordinate level distinctions (Duvdevani-Bar and Edelman 1999). These two strategies are considered in some detail in the next section.

View-Based and Structural Description Theories: Strategies for View Independent Recognition

A growing body of experimental evidence now suggests that performance on recognition tasks is proportional to the distance from the nearest familiar view. This includes both error rate measures (Bulthoff and Edelman 1992) and recognition time (Tarr and Pinker 1989); (Tarr, Bulthoff et al. 1997).

The structural approach derives essentially from the early proposal of Marr and Nishihara (Marr and Nishihara 1978) that the ultimate task of object recognition is the recovery of 3-D structural relationships from the 2-D retinal projection. If such a structure could be derived, then recognition of the object would be largely viewpoint independent. Faced with evidence that recognition is not invariant but varied linearly with distance from learned views for "paperclip" objects (Bulthoff and Edelman 1992), the recognition by components (RBC) approach (Biederman and Gerhardstein 1993) was refined to suggest that the maximal viewpoint invariance occurs when three conditions hold:

1. Objects consist of geon-like parts. Geons are a set of simple volumetric components which can be configured to produce a wide range of everyday objects.
2. These parts form qualitatively distinct configurations in different objects.
3. These parts are visible over the range of viewpoints for which invariant performance should occur.

These predictions were tested by substituting a unique geon for one of the cylinders in the chain, which resulted in nearly invariant performance. A subsequent study by Tarr and colleagues (Tarr, Bulthoff et al. 1997) revealed that the single geon case is exceptional, and generally performance falls off with distance from a learned view with three or five geons. This 'paperclip with added geons' image set from the Tarr group (henceforth denoted here as the paperclips+ set) was selected for study in the simulations described later in this thesis.

In contrast to the experiments just described, in my simulations the original gray scale images are reduced to silhouettes, reducing the 3-D information available by shading and occlusion in the raw synthetic images. This is done chiefly to simulate putative edge extraction mechanisms in early visual layers in an attempt to rely on form alone. Another recent recognition study by Hayward (Hayward 1998) found no significant difference in viewpoint-dependent performance between silhouettes and shaded, part-boundary-visible versions of objects, indicating that features in the boundary contour are largely responsible for recognition and changes in performance.

Network Implementations of View Interpolation

The view based approach has been developed extensively with feed-forward neural networks, stemming from a general strategy first described by Poggio and Edelman (Poggio and Edelman 1990). This strategy is essentially view interpolation by regularization, or by normalization[8] in Tarr's terminology (Hayward and Tarr 1997). Normalization refers to the concept that different, perhaps novel views are mapped by some computational process to a representation derived from the trained views. This mapping occurs by transformation of several views of an object in a high dimensional measurement space to a lower dimensional shape representation space. A learning process operating over the presented views (e.g. the adjustment of network parameters) ensures that the transformation approaches the same point in the representation space for all trained views.

The dimensions of the measurement space correspond to an assembly of tuned filters. This transformation occurs by approximating the statistics of activation and their changes with basis function units (Poggio and Girosi 1990). The statistics are captured

---

[8] I will use the term normalization; regularization implies a certain underlying mathematical approach is used (Poggio and Girosi 1990); a major result in the present work is to demonstrate an alternative mathematical approach and network realization to accomplish the normalization.

by the selection of centers and widths of Gaussian radial basis function (RBF) units and weight values from each unit in the basis function unit to an output summing layer.

A more elaborate version of this approach is found in the Chorus system (Duvdevani-Bar and Edelman 1999); (Edelman 1999). In Chorus, a set of prototype objects is chosen as representative of a larger world of objects. The *n* prototypes are drawn from a smaller number of categories. RBF classifiers for each of these prototype objects are designed (trained). With advance knowledge of all the prototype objects, optimal views for each can be chosen which achieve the best normalization (constant activation of the classifiers for all views) as well as maximizing inter-cluster distances in the space of all prototypes. Any known or novel objects applied to the measurement units (200 tuned filters) map to a point in the *representation space* whose dimensions correspond to each prototype. This point is signified by activation values on each prototype unit.

Given this representation space, categorization can be performed by various strategies. Nearest neighbor match chooses the category of the object with minimum distance; another, k-nearest neighbors, examines the category of the k nearest neighbors and selects the category based on majority vote.

Critique of the feed-forward view interpolation theory

Supporters of the view based stragegy generally attribute recognition time effects to a normalization (i.e. orientation correction) process, similar to that assumed for mental rotation processes. However, Chorus, a well knnown computational view-based model with claims for biological relevance, does not actually predict any differences in reaction time for the normalization process. The one-shot feed-forward flow through the network is the same for any view presented to the network, whether novel or previously learned. A previously proposed network, with a spreading activation architecture, had a more natural interpretation for reaction time (Edelman and Weinshall 1991) . In general, some form of iterative computation and competitive interactions progressing toward a decision state have been invoked to explain reaction times in connectionist models, while more abstract theories such as the diffusion model (Ratcliff, Van Zandt et al. 1999) claim to explain response time distributions and differences in response distributions for error and correct responses.

Nonlinearities in response time vs. distance from familiar views have been noted by several investigators under certain testing conditions. Hayward and Tarr used a set of qualitatively distinct single part geons previously used by Biederman and Gerhardstein (Biederman and Gerhardstein 1993), but changed the experimental conditions to eliminate possible opportunities to learn multiple views and exploit local diagnostic features (Biederman and Gerhardstein 1993). They designed objects and training viewpoints such that for ± 45° rotations from the trained view, one direction resulted in no qualitative changes, while the other produced qualitative changes, such as the disappearance of areas of curvature. For these conditions, they found that response time varied between quantitative and qualitative conditions (610 vs. 650 ms) and error rates also varied similarly (13.5 vs. 4.5 %). Hayward and Tarr conclude, then, that these results contradict a normalization process based simply on magnitude of rotation.

The source of the well-established reaction time effect is arguably an artifact of some matching process of stored codes and codes formed in the early visual pathways, possibly involving synchronization processes. Based on these nonlinearities, Edelman has argued that the mental rotation hypothesis has weak support based on the evidence mentioned above, and that the disappearance or reduction of delay with practice represents a faster path to recognition, not an increase in the rotation rate.

The subject of visual search and attention has to date been given relatively little attention in the literature on view-based object recognition, but problems with the feed-forward recognition model are also apparent in this context. It is easy for humans to search a visual scene for a familiar object and to know that it is not there, but feed-forward models do not readily address this case. Extensive experiments by Miyashita indicate a repeatable, stimulus-specific response in anterior ventral IT cortex during a 16 second delay interval in a delayed match to sample task (Miyashita and Chang 1988). The stimulus is not present during this interval, and the response is statistically distinguished from the period when the stimulus *is* present. This is interpreted by Miyashita as a neural correlate of short-term memory for the particular shape. A feed-forward model considers the network weights to be the essence of memory, and predicts no stimulus specific response during a delay period.

Single unit studies addressing attention and search aspects of object recognition in IT cortex have led to considerably different interpretations of the functioning of IT than those cited by Edelman and Tarr. These are described in more detail in a later section, but for now I note the findings of Eskandar et al. (Eskandar, Optican et al. 1992) that the best prediction of the stimulus from spike trains results from interpreting the trains as a multiplication of a target code and the incoming stimulus code during a search process. This could be interpreted as an intermediate computation (a weighting process) in an RBF-style computation leading to activation in a certain area. Alternatively, it might be interpreted as a cooperative synchronization process, also ultimately resulting in activation in a few areas which are structurally and dynamically suited to synchronize with the stored memory representation. The latter type of computation is the focus of the theory and experiments here.

A final issue I raise regarding the neural correlates of psychological phenomena was pointed out by Tsuda (Tsuda 1992), that of the difficulty of breaking the life of an organism into clean epochs of learning and recognition. It seems likely that normal exploratory behavior involves both of these activities proceeding in parallel, or at least that a system is poised to be able to rapidly switch from one to the other as the dominant mode. Dynamical models involving continual *bifurcation* (dynamical parameter changes), but more explicitly recast in terms of synchronization dynamics, may be a more appropriate architecture for combining learning and recognition in a natural way (Skarda and Freeman 1987). The approach in this thesis, while consistent with oscillatory representations and synchronization-based computational strategies, does not yet step up to the challenge of dynamic shifting between learning and recognition modes.

In summary, while view-based recognition has deservedly emerged as a leading computational theory of object recognition and representation, several issues have not been addressed: Primed search (search for an object held in short term memory in a visual scene), variance in reaction times, and the longer time scale contextual shifts between learning and recognition. While these issues are not addressed or resolved theoretically or experimentally by the present work, we will return to these subjects in discussing the relative merits of three recognition approaches with claims to biological relevance.

PERSPECTIVES FROM COMPUTER VISION

Most contemporary work in the psychology of vision and perception uses computational and signal processing concepts. The converse is not true, in that many algorithms proposed for recognition have no ready interpretation in neural network terms. In reviewing developments in computer vision most relevant to the present work, I will first focus on a few classical dilemmas related to object recognition. I then review some recent work claiming to be biologically motivated, and finally mention some recent algorithmic approaches which share aspects of the computational style. In spite of the emphasis here on dynamics and neuroscience, the synchronization opponent lattice network also has something in common with recent trends in computer vision including nonlinear diffusion, deformation, and feature histogram methods; thus it may be improved by drawing on continued progress in those areas.

One algorithm (geometric hashing) is presented which may seem a bit out of context with the rest of the discussion. I include it because it handles two problems – invariance for discontinuous point sets and embedding of objects in a scene – which I do not believe can be handled by any methods discussed here, including my own.

Classical Pattern Recognition in the Image Domain

Some attempt must be made to situate the present work in relation to the long and diverse history of image recognition methods. To concisely present the history and recent trends of such a vast field is challenging; I will emphasize the areas of transformation and multiple scales that characterize recent geometric methods, and will stress the way in which transformative methods can blur traditional distinctions between structural and syntactic approaches and scale issues.

Several surveys on image processing methods identify the major classical methods as either statistical or structural (Freeman 1985); (Leedham 1991); (Del Bimbo 1999). Del Bimbo describes more recent approaches as “shape through transformation”. Thus classical recognition methods – both statistical and structural (or *syntactic*) – are relatively passive, in that they do not modify the base image. They merely subject it to some interpretive framework, such as a particular feature set. In contrast, my method (and others I will survey) *modifies* the image in some way prior to measurement on a modified image, or possibly measurements over a sequence of modifications.

Classical Methods and Dilemmas

Statistical Methods

In statistical approaches, pattern data is represented by a feature vector which is used as input to some classifier or decision process. Features may characterize global form (area, elongatedness, major axis orientation) or local elements (corners, characteristic points). Shapes are viewed as points in shape feature space. For effective recognition, the requirement is to choose features such that patterns of the same class are tightly clustered in N dimensional space corresponding to N features, and patterns of different classes are in other tightly clustered regions well separated from each other (Duda and Hart 1973).

A key problem in statistical methods is the reduction of the dimensionality of the feature vector. This may be accomplished by a feature selection process, in which low significance features are deleted, or by a feature space transformation method, or both. Classically, a particular class was represented by a template with matching against templates; this matching was considered to be intractable for large numbers of objects due to the need to compare with inputs which have been rotated, scaled, partly occluded, non-rigidly transformed, or presented under varying lighting conditions. Recent schemes employing normalization (the RBF networks underlying Chorus) and interactions among multiple well chosen prototypes, or the sophisticated weighting of a large feature set (Mel 1997) have overcome this to some extent.

Another approach to the use of features is to create a transformed representation space on the basis of correlations among the dimensions to enhance cluster tightness and inter-class separation. Feed-forward supervised networks, or competitive networks such as self organizing maps can use feature vectors as input, and via training transform the features into activation levels in a set of network elements corresponding to classes.

Decision methods may generally be classed as non-parametric or parametric (Leedham 1991). Non-parametric methods include linear discriminant functions, minimum distance classifiers, and nearest neighbor classifiers.

The most widely used parametric decision rule is the Bayes classifier. The main distinction from non-parametric methods is that the decision rule involves class conditional densities and *a priori* probabilities of occurrence of classes. Bayesian classifiers are particularly important with large object databases, where setting classifier decision boundaries properly and defining the optimal feature set are crucial for good recognition performances.

The description of statistical pattern recognition methods presented here thus far has been in general terms, applicable to any data set. Recognition of object shapes in a statistical framework poses additional problems unique to this class of data. Non-rigid objects are composed of parts which can assume different poses – human and animal figures are good examples.

The changing projections of three dimensional objects seen from different viewpoints constitute the *stimulus identity problem*. Different features and feature conjunctions will be present in each view. This problem has been addressed by geometric methods seeking invariants (treated in the next section), or by neural networks exploiting regularities in the changing distributions of raw features (the Chorus RBF ensemble approach). Recently, however, progress on stimulus equivalence within a "raw feature" paradigm has been demonstrated, by careful design.

Mel, describing the design goals for a recent high performance feature based system (Mel 1997), notes the following expectations on feature sets to overcome these problems:

1. Features should be large in number; sparsely occupied high dimensional representations are most robust to noise.
2. Features should be useful; they may be sensitive to object quality (occlusion, poor lighting) but should be robust in the face of pose or configuration changes.
3. They should be dominated by spatially local features; this is particularly important for non-rigid objects , which preserve local but not global structure in any particular view.
4. They should be driven by multiple visual cues to maximize discrimination, represent diverse objects, and buffer representation against degradation which affects different cues (feature channels) more or less severely.

The use of these principles led to the creation of his SEEMORE system, which achieves recognition rates above 90% in a 100 object world, even for scrambled images. The high performance achieved with these first order[9] feature channels is interpreted by Mel to support the idea that a simple feature space is all that is needed and attempts to extract structural information or otherwise "bind" collections of features may be unnecessary for biological systems. It is easy, however, to construct images with identical first order statistics which will fool such a system but are readily distinguished by humans. It seems likely that some of SEEMORE's recognition success depends on diversity in first order statistics of the object world, along with limited use of second order statistics for some feature channels.

Structural or Syntactic Methods

The other major family of classic pattern recognition approaches, chiefly developed for image or shape processing applications are structural or syntactic methods (Pavlidis 1977). Here, the input image must first be segmented into primitives; the primitives must be recognized, and spatial or topological relationships between these primitives extracted. Finally, with this information, a syntactic analysis and classification on that basis can proceed. None of these problems are trivial.

Within computer vision, structural methods based on raw image data have been largely superseded by related methods which capture structural information implicitly by *multi-scale* representations or by deformations. In the psychological examination of human vision, structural approaches still command a good deal of support. In part, this is

---

[9] First order features implies that no information on the spatial proximity of other features is present. Second order features would capture adjacencies of feature pairs at one or more scales, with increasing high order features preserving this trend.

due to the fact that task specific or language mediated descriptions of objects offer evidence that *compositional* representations are used. Statistical approaches, and feed-forward neural networks have been problematic in regard to this issue.

Compositionality is essentially the separability of the components of a composite representation, i.e. the ability to use or talk about them independently after the formation of that representation (Van Gelder 1990). Recurrent networks have been demonstrated to exhibit a so called functional compositionality, in which tree structures can be represented and their constituent parts derived (Pollack 1990).

Some Geometric Methods for Shape Description

Finding a representation of shapes which is invariant to viewpoint has been approached from a variety of methods which are difficult to justify biologically. However there is evidence that classes of stimuli, such as point sets in a regular geometrical arrangement, are recognized even in a noisy background (Uttal 1988). It is unlikely that other methods discussed here involving local receptive-field computations (e.g. SEEMORE) would handle this situation well. The first geometric method to be examined deals explicitly with point sets and is designed to work in scene analysis.

Geometric Hashing

Geometric hashing (Wolfson and Yehezkel 1992) was proposed as a means of performing model based recognition in scenes, with robustness to partial occlusion and to transformations in the plane. This is accomplished by considering an object as a point set, and by remapping coordinates of every point in terms of all possible triplets of non-collinear points. For four points $A,B,C,D \in \boldsymbol{R}^2$, affine invariant coordinates of D are coordinates with respect to axes defined by $\overline{AB}$ and $\overline{AC}$. Affine transformations (translation, scaling or rotation in the plane) will produce a new set of points A',B',C',D'; the coordinates of D' in the A',B',C' coordinate system are unchanged. First, signatures are generated from interest points on one or several views of an object. Interest points are endpoints or intersections of segments extracted by some edge extraction procedure. During scene analysis, interest points are selected and processed by a similar computation. The signature generation procedure is outlined here:

```
procedure signature_generation
  for each model object
        extract m interest points for the object
        for each ordered non-collinear triplet (affine basis) do
              a) compute coordinates of all m-3 model points in the affine coordinate
                 frame for the current basis;
              b) use the coordinate as an address to a hash table;
              c) record in the table entry a pair {model, basis} for which the
                 coordinate was obtained
        end for
  end for
```

```
end procedure
```

The complexity of signature generation is of order m^4 per model. The creation of the hash table is viewed as a learning process, in which a memory is formed relative to different foci of attention.

The corresponding matching procedure is then:

```
procedure match_model_in_scene
        a) extract n interest points from scene
        b) choose arbitrary ordered triplet of non-collinear points, compute scene
        points referenced to this triplet as affine basis.
        c) for each such coordinate
                check the appropriate entry in hash table;
                for every {model, basis} pair , tally a vote for the model and affine
                basis.
          endfor
        d) If a certain {model, basis} pair scores many votes, decide this is the
        object.
        e) Consider all {triplet, image point} pairs which voted for winning
        {model, basis} pair
        f)  find the affine transformation giving the best least-squares match
        between corresponding point pair views.
        g) transform the whole low level representation of model according to affine
        transform and verify it vs. the scene.
end procedure
```

Multiresolution Methods

One major problem with feature vector classifiers is that the relevant features of an object tend to vary with scale in a way which is unknown *a priori*. Overcoming this defect is a major goal of *scale space* approaches, such as geometry-driven diffusion. While such methods can adaptively tune feature representations for shapes with detail at many scales, the mapping of the resultant curve family to a representative feature vector can be computationally expensive, and some of the features advocated are difficult to discover (i.e. the detection of singularities in evolved curves). Wavelet decompositions also perform well in terms of capturing details at multiple spatial scales, but early formulations had problems with translational and rotational invariance; newer methods, such as *steerable pyramids* (Simoncelli, Freeman et al. 1992) claim to overcome these limitations.

Transformational or Deformation Methods

Scale spaces are a generic term for families of derived images or shape which attempt to capture aspects of shapes at various spatial frequency bands. In *smoothing* approaches to scale space shape characterization, a gray-scale luminance image is subject to a smoothing evolution by a family of Gaussian kernels with increasing neighborhood size. Alternatively, in a more geometrical, abstract formulation, a curve may be evolved by displacement at each point by moving in the direction of the normal vector by an amount proportional to the curvature at that point. Each evolved shape in this iterative process can be characterized by some feature. Zero crossings of derivatives, inflection points, curvature extrema, and symmetry axes have been used as features (Kimia and Siddiqi 1994). The set of features extracted after the evolution process captures the shape characteristics at a variety of scales. Extrema that survive larger smoothing extents may be considered more significant, and might be weighted more heavily during feature based distance computations.

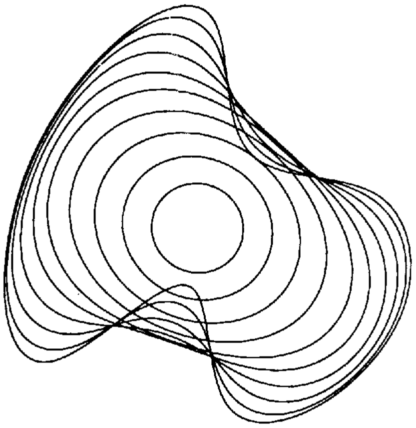


Fig. 2.Example of curve evolution by geometry -driven diffusion. Smoothing occurs by displacing each point from the original curve proportional to the local curvature. The series of curves generated serve as the basis for characterization of the original shape. From Kimia, B. B. and K. Siddiqi (1994). Geometric heat equation and nonlinear diffusion of shapes and images. Computer Vision and Pattern Recognition, Seattle, IEEE Computer Society. Used with permission, IEEE.

The literature on curve evolution is primarily concerned with theoretical problems and extensions and short on applied comparisons to other approaches. Curve evolution scale space methods are argued to give good qualitative descriptions of shape, but are rather expensive to compute and do not allow easy reconstruction in contrast to decomposition methods (e.g. wavelet transforms) which also capture information at various spatial scales. The listing of inflection points at each scale or iteration involves difficulties in the data representation and efficient computation of a similarity function, since the number of inflection points for each shape is not constant. In addition, if the shapes to be characterized are not simple solids bounded by a closed curve, missing interior components would need to have separate evolutions and feature vectors, possibly leading to ambiguities, singularity problems or problems handling an extended set of vectors for each component when applying a similarity function.

Many of these defects have been reduced by a closely related, more recent approach utilizing *shocks* (Tek and Kimia 1999). Shocks are the sites where diffusion wavefronts from an object outline collide internally (on the medial axis of a shape), and externally as well. Resulting shock graphs and their grammars have been used to describe shapes, with similarity functions on the resulting graphs defined. The representation has been proven sufficient to reconstruct the local shape from the medial axis, tangents, velocity and acceleration of shocks (Giblin and Kimia 1999) . Medial axis representations are normally sensitive to deformations in the outline, but methods have been developed to distinguish stable from unstable shocks to ameliorate this problem (Giblin and Kimia 1999).

While the neural mechanisms which might implement such transformations are rather opaque, there is evidence that axis representations or skeletons are computed in some fashion and influence the response of cells in primary visual cortex (Kovacs and Julesz 1994); (Lee, D. et al. 1998).

Morphological Scale Space

A related approach relying on nonlinear transformation of the image at multiple scales is designated *morphological scale space* (Korn, Sidiropouls et al. 1996). In this approach, morphological operators of increasing scale are applied to the original image or curve, again resulting in a family of transformed images. A *pattern spectrum* has been proposed to characterize such an image family (Maragos 1988). The spectrum consists of the accumulation of successive differences in area between a pattern and its successor as opening and closing morphological operators of increasing scale are applied. Similarity functions can be applied to the resulting histogram.

Comparison of Computational Methods and Psychological Responses

The application of geometric algorithms for shape similarity to the problem of image retrieval in multimedia databases has motivated studies of how well a particular algorithm corresponds to human judgements on the same task. One such study with a large and diverse set of images (Scasseleti, Alexopoulos et al. 1994) found that each of several algorithms performed very well for certain target images, but poorly on others. Turning angle, the most robust of the algorithms across the images, was the best match to human preferences on only 8 of the 20 target images used. Turning angle methods require a search or fitting procedure to insure the best alignment between the feataure vectors prior to computing the distance function; also, such raw curvature descriptions are sensitive to scale. Descriptions based on sets of inflection points, like the sign of curvature approach, reduce scale sensitivity in comparison with raw curvature.

This finding of an apparent psychological primacy of local contour based measures corroborates the finding of Hayward already mentioned (Hayward 1998) that contours of binary images (silhouettes) are recognized with comparable performance to gray scale images, and that contour (or some transformation of contour) is the major information source for human object recognition and similarity.

Summary: Situating the Soca Approach in Computer Vision

The framework and implementation I develop here draws from many classical computer vision concepts, as well as the image transformation, and representation space concept introduced earlier in the context of the Chorus system. Statistics are used to match histogram templates; currently a nearest neighbor decision function is used. These statistics are over an abstract representation space which is designed to achieve the goals of within-class tightness and inter-class separation, where each class is a depth rotation-invariant description of a three dimensional object. The rotation invariance is formed by a transformation method involving diffusion and blurring as part of its mechanism, like the heat equation deformation methods. These scale space methods typically avoid *creating new spatial structure*; in contrast, the procedure described here is completely dependent on creating fine structure, and on cooperative interactions derived from those structures.

A fundamental aspect of the Soca network implementation is that there is local, receptive field like processing, with a diffusive “spreading of activity” character. However, this activation is not to be understood as a monotonic variable associated with detection of some feature. In image processing terms, the process can be considered as a nonlinear filter with feedback, or iterative nonlinear convolution. This combination of diffusion and highly nonlinear (non-monotonic) transfer function forms a representation determined by both local features (e.g. curvature and corner elements) and medium-scale structural relationships. The scale of interactions is determined by a window proportional to the number of network iterations used to generate a representation meeting some criteria. A particular juxtaposition of local curvature changes may, with appropriate network parameters, result in a unique distribution or histogram in the representation space. This type of process is, to my knowledge, a unique approach to combining local feature and structural information; thus it represents one of the main contributions of the thesis.

Forming such a representation - one that captures the co-occurrence of local features - is a hotly debated subject in neuroscience, referred to as the *binding problem*. While the problem is typically presented in terms of separate channels (such as color and shape), the situation of decomposing an image or outline into a set of orientation frequency detectors presents the same difficulty. Opinion on the subject ranges from claims on the neural correlates of binding to assertions that there *is* no problem. This will be discussed in some detail in a subsequent section on neuroscience.

PERSPECTIVES FROM THEORETICAL COMPUTER SCIENCE

Computer science has certain perspectives and emphases that result in a characteristic way of framing the problems of similarity and recognition. In this section, I focus on two such perspectives. First, any spatial or temporal pattern can ultimately be represented as a string in some alphabet. One traditional theoretical approach is to consider families of such strings as a formal language, and to frame recognition problems in terms of recognizing a language. This way of formulating recognition problems is also relevant to the present thesis because it allows the tools and language of symbolic dynamics to be applied. Symbolic dynamics will be taken up in more detail in the section on dynamics and representation, and in theoretical discussions on proving the representation-forming capability of Soca style transformations.

Dynamical Recognizers and Computational Mechanics

The problem of learning to accept positive exemplars of a language while rejecting negative exemplars is known as language induction. Classical machine learning approaches to this problem construct a finite state automaton to affect recognition. Formally, a finite state recognizer is a quadruple $\{Q,\Sigma,\delta,F\}$, where Q is a set of states (with q_0 denoting the initial state), Σ is some finite alphabet, δ is a transition function mapping $Q \times \sum \Rightarrow Q$, and $F \subset Q$ is a set of final or accepting states. A string of tokens from alphabet Σ is *accepted* by the recognizers if, starting from initial state q_0 the sequence of state transitions indicated by the tokens in the string ends up in one of the final states in subset *F*.

A pioneering attempt to formulate the language induction problem as a dynamical system, in the form of a recurrent neural network, was the study of Pollack (Pollack 1991). The dynamical recognizer is a quadruple $\{Z,\Sigma,\Omega,G\}$, where $Z \subset R^k$ is a state space; $z_k(0)$ is the initial condition. Σ is the input "alphabet", where a particular closed interval in Z corresponds to each element in this alphabet. (This correspondence between *intervals of state-space* and *symbols* is a cornerstone of *symbolic dynamics*, which will be mentioned again later). Ω is the dynamic, a sequence of transformations $\omega_l:Z \rightarrow Z$ (one for each token) with an associated set of dynamical parameters; these parameters are fixed for a particular recognizer during the induction (training) process. $G(Z) \rightarrow \{0,1\}$ is the decision function which maps one or more states in the sequence produced by the dynamic to an accept/ reject decision. In Pollack's work, only the final state and token are used in the decision function. Within this general framework, the dynamics and decision function are normally much weaker in computational power than a Turning machine. Pollack notes that G may be generalized to a graded function indicating "fuzzy" acceptance, or could return a more complex categorization or representation.

The Soca network and recognition method I describe later is quite consistent with this extended dynamical recognizer framework. A key difference is that the Soca net operates on an image "string" in parallel (thus the state space has higher dimensionality R^N, where N is the number of pixels or sampled image elements), and the tokens are used only once as the initial state. Such a parallel recognizer framework for *picture languages* with constrained states and transformations was studied in a series of papers by Rosenfeld (Rosenfeld 1979). In Rosenfeld's formulation, the transition function at each pixel is now a function of several tokens in some spatial neighborhood; this is the cellular automata formalism, which is described in the dynamics section later. The decision function is necessarily modified by this larger state space. Rosenfeld proposed several possibilities:

1. every spatial element reaches an accepting state
2. any element reaches an accepting state
3. one particular spatial element reaches an accepting state.

In the Soca network and recognition strategy, my approach is to form a metric representation space, but that space consists of *statistics measured instantaneously during a high dimensional, parallel dynamics,* rather than a direct map of the input features or measurement space. These statistics naturally support an acceptance function; simply define some threshold distance for each classifier, and accept an object as an instance of language L if it satisfies this distance test. The distances might vary by class, depending on the cluster density of that class in the representation space. Another contribution of the thesis, then, is adding another type of decision function to the repertoire defined by Rosenfeld. While such a distance threshold decision function is common in statistical pattern recognition, it is novel for processes operating with local dynamics.

Other researchers have recently been concerned with decision functions over spatial patterns processed by cellular automata, a form of spatially-extended dynamical systems closely related to those used in the present work (Mitchell, Hraber et al. 1993); (Mitchell, Crutchfield et al. 1996); (Hordijk, Crutchfield et al. 1998). Genetic algorithms were used to generate and test particular one dimensional cellular automata (CA) which decide, for example, whether a random initial condition has majority ones or zeros. The group then examines space-time plots (i.e. plots of successive iterations of a 1-D spatial array) of the resulting successful computations and develops an explanatory framework based on physical metaphors; this framework is designated by this group as *Computational Mechanics*.

Computational Mechanics seeks to reconstruct the computations embedded in space-time behavior in terms of regular domains, particles, and particle interactions. *Regular domains* are regions visible in space-time plots consisting of words (spatial configurations) in the same regular language, i.e. regions that are computationally homogeneous. *Particles* are localized boundaries between such domains; they serve as information carriers. *Collisions* between particles are the loci of information processing. This processing can be conceived in terms of operators such as decay of one particle to many, reactions (state transitions between language domains at collision sites), and annihilations (the disappearance of an interface as one language domain dominates future spatial evolution at a collision site). The computational strategy can then be expressed in the more concise language of particles and their interactions, substituting for a more verbose description in the language of CA rule lookup tables and raw spatial configurations.

While the computational mechanics group does not explicitly state this, the decision function in the majority task can be considered a type of *synchronization* – until all cells reach the same language domain (which, in this simple case, is all 0 or 1) the system is undecided. Consider a k-block as an adjacent set of k cells, with the entire CA consisting of overlapping sets of such k-blocks. Each k-block of cells in this automaton is defined by the state transition graph of a finite state automaton (FSA) with k states. As the automata evolves, at each time step we can label each state with the fraction of k-blocks which currently hold that value (we say they occupy the state). Synchronization in this context implies that over time, the occupancy statistics of the graph converge to sharp peaks, or an unchanging sequence of sharp peaks; particular sub-graphs of the state-transition graph are active, while others become blocked, as their predecessor states become unreachable within increasing spatial "territories".

Note that particles have a characteristic velocity, and for certain kinds of terminating conditions (such as a particular site or region reaching a value in a set of accepting states F) one possibility for variance in the temporal processing is dependence on the emergent particle velocities on initial configurations in a family of inputs, when a "synchronization" decision function is reached.

Summary: Situating the Soca Approach in Computer Science

In summary, the Soca system extends the tradition of dynamical language recognizers over spatial configurations, and attempts to unify this approach with traditional metric representation space of statistical pattern recognition. Similar to the work of Mitchell, Crutchfield, and their colleagues, the approach taken here is to discover successful computations within a particular family of spatially distributed computations, then analyze the result. I have generally proceeded with more constraints on the search process, guided by general principles of pattern recognition.

The decision functions used here also involve synchronization in the sense defined above, but the synchronization is partial and not defined to contiguous regions as in the regular domains. Another key distinction of the Soca work from the research in the computational mechanics group is that the present search strategy focuses on solving the decision problem within a fixed number of iterations, rather than an open ended synchronization process. This led to the hypothesis that dynamical changes (non-stationary parameters or rules) might lead to superior performance relative to constant dynamics, by forcing more rapid synchronization.

Chapter 3: Topics in Neuroscience

In this chapter I will review the current issues and controversies in neuroscience which have motivated the form of the network dynamics I investigate. The review will encompass both single neuron and *integrative* neuroscience, with the latter referring to larger scales and interactions across scales.

I begin with an overview of those objects of neuroscience research that bear on the research topic of this thesis. Since the focus of the dissertation is on similarity and object recognition, I survey recent experimental work on inferotemporal (IT) cortex, and the theoretical commitments of the various research groups. This region is clearly implicated in recognition and memory of objects, so it is important to review the findings in that area before evaluating computer vision systems with claims to biological motivation and plausibility.

Finally, I describe different approaches to neural systems modeling and how these relate to the structures and methods of experimental neuroscience. A high level treatment of the assumptions of connectionist neural modeling is contrasted with what might be called *dynamical pattern* network modeling. I situate the present work in this modeling context. The story will remain incomplete until the next chapter on nonlinear dynamics, where many concepts required to discuss networks with more complex dynamics are presented. I revisit biological and neural modeling concepts after introducing such dynamics, and again in the final discussion.

BASICS OF NEURAL ORGANIZATION

The nervous system of humans can be regarded as consisting of *peripheral* sensory and motor nerves, which connect transducers to the *central* nervous system. The brain itself consists of various concentric layers progressing from the evolutionarily oldest brainstem, through the midbrain, to the neocortex. Separate divisions within the dense brainstem and midbrain regions are typically called nuclei or loci. The cortex, in contrast, is generally divided into regions (areas, modules) distinguished long ago either anatomically (by staining), functionally (by observing the effects of injury on that region) or both. Modern imaging techniques reveal which of these modules are co-active in the performance of a perceptual or cognitive task, and to a limited extent the sequencing of their activity.

While cortical regions can be distinguished functionally, the amount of structural similarity is striking. The basic cellular unit is the neuron, with hundreds of specialized types organized into micro-circuits and larger systems. Cortical regions have a columnar organization. There is a fine structure of minicolumns (30 μm in humans) and a more coarse structure of macrocolumns (.4 – 1 mm). For a recent review, see (Calvin 1995). The number of neurons across thickness of cortex (30u cylinder) are remarkably constant around - 110 in motor, somatic sensory, frontal, parietal, temporal, in mouse, cat, rat, macaque monkey, and man (Mountcastle 1978). Primary visual cortex columns are more dense, with perhaps 160 cells with complex intra-modular connectivity. By intra-modular connectivity, Mountcastle means that the connections between regions are not simple one to one projections of every minicolumn to a corresponding one, but consist of "subsets, each with a particular pattern of connections to similarly segregated subsets in other regional entities".

Within each column, neurons are organized into horizontal layers, again with a relatively uniform scheme of projections into and out of layers. Layers are numbered 1 (outermost) to 6 (deepest). In the visual system, layer 4 receives input from thalamus, basal ganglia, and other cortical areas. It projects feed-forward to layers 2 and 3; these are chiefly connected in their own layers and laterally to other columns. These in turn project to 5 (motor output) or 6 (return to thalamus and other sub-cortical structures. Note that loops or *recurrent* structures are pervasive in at least three groupings: intra-layer, inter-column, and cortico-thalamic.

While investigators such as Mountcastle and Calvin emphasize the pervasiveness of columnar structures in the cortex, counter arguments questioning the role of columnar structures are found in (Purves, Riddle et al. 1992); (Swindale 1990).

Many regions are organized as *maps*[10], meaning that the spatial relationship between some sensory field is preserved through one or more registered regions. The visual system is well known in this regard; the sense of touch is also organized to produce a topographic representation on a contiguous region of cortex. There are often characteristic topology transformations in this mapping structure. The retina introduces non-uniformities in sampling projected into visual cortex (area V1). For example the macaque monkey areas V2, V3, VP are elongated in the horizontal (central to peripheral). Some areas (inferotemporal) emphasize the central, densely sampled region while others (parietal) emphasize the sparsely sampled periphery.

Anticipating the topics of the later chapters, it should be pointed out that much of the history of neural modeling beyond low level vision fails to take into account these two spatially regularities: mapping and laterally connected columnar structure. This remains a gap between biologists and the connectionist community. The family of models in the present thesis, known as coupled map lattices or discrete time cellular neural networks, resemble the mapped column structures, with each "cell" representing collective behavior of hundreds or thousands of neurons. Thus such models are considered *medium scale* models, in contrast to *microcircuit* or *small-circuit* models. Such regular physical architectures with chaotic units and recurrence have *implicit* connection structures between communicating states, which may or may not be bound to specific column-like structures [11]. The units of representation and computation are these states, rather than activation values of specific output neurons of a microcircuit.

---

[10] The term map is used in a different sense elsewhere in this thesis, in the sense of a discrete time function mapping values in a phase space.

[11] For more detail on this topic of implicit structure in oscillating networks, see(Ito and Kaneko 2000) and (Kaneko 1990)

THE NEURON

With a few exceptions, most theories of nervous system functioning focus on the neuron as the main actor, playing roles in both signal processing and more abstract computational processes. The assumed complexity of the operation of single neurons has tended to increase with time. In their introduction to a recent compilation of neuron and small-circuit level modeling, Abbot and Sejnowski (Abbot and Sejnowski 1999) offer the sobering conclusion that at the time of writing (2000), there is actually little consensus about the mechanisms leading to details of spiking activity in a single neuron, particularly regarding fast, stimulus linked spike rate *modulation* that would be required for *temporal codes*.

The classical model of neuron operation sees it as a threshold device, integrating positive (excitatory) and negative (inhibitory) inputs on a graded potential input system. This input system is the dendrite; a typical neuron receives input from as many as 10,000 other neurons. When a threshold is exceeded at the cell body or soma, the neuron fires a spike down the output (axon), which contacts other neurons. The actual contact between axons and dendrites is via complex electrochemical activity at synapses. Neurotransmitters are chemicals released during firing of the axon, and received at the dendrites; local production and absorption (*re-uptake*) of these chemicals are modulated by complex electrochemical processes. The modulation of thresholds by transmitters at the synapse and the propagation of spikes both rely on the complex dynamics of ion channels. The dynamics of transmitter production, release, and channel chemistry are subject to *modulation* at a variety of time scales, allowing networks with constant anatomical connectivity to perform quite differently.

Any particular synapse is either excitatory (increasing voltage) or inhibitory (decreasing voltage). A particular neurons axon terminals are all either excitatory or inhibitory; while most neurons receive a mix of excitatory and inhibitory input. Families of neurotransmitters play predominantly excitatory or inhibitory roles.

Early in the history of neuroscience it was thought that cross-scale interactions between field activity of the neural mass and individual neurons (electrotonic coupling) might play an important role in neural computation (Lashley 1942). Recently, there has been some revival of this concept via the idea that gap junctions[12] sensitive to activity in the dendritic mass (neuropil) may affect behaviorally relevant synchronization properties in brainstem neurons, even though the gap junction mediated currents may be as little as 2% of the total dendritic currents (Usher, Cohen et al. 1999). No equivalent demonstrations of such effects in cortex are known to the author, but these results are interesting in light of the fact mentioned earlier about relative sparseness (e.g. 10%) of inter-column connections. It is possible that rapid communication between columns could be mediated through this mechanism in less time that would be required for synaptic transmission. The reader should bear this in mind when reviewing arguments against any role for recurrent computations in visual processing, which are often made on

---

[12] Gap junctions are a cell membrane structure similar to ion channels; they are universal in intercellular communication, not limited to neurons. Specific protein regulation mechanisms in gap junction complexes have been implicated in disease, e.g. high frequency deafness.

the basis of the hypothetical performance characteristics based on synaptic transmission delays.

EXPERIMENTAL NEUROSCIENCE METHODS

What we know of neural functioning comes from a variety of experimental methods developed over the last 100 years; the organization of the next few sections reflects the knowledge gained from, and theoretical biases associated with, several techniques. The emphasis is on signal flow and dynamics; methods for ascertaining structure at various scales, down to detailed receptor types, are beyond the scope of this review.

Neural measurements on organisms are typically done in one of three modes. Measurements may be performed on *slice preparations (in vitro)*, with a section of tissue cultured for some time. With such cultures, transmitter dynamics can be observed with the fast cyclic voltammetry technique (Stamford 1990). Neuron cultures have also been preserved directly on silicon electrode arrays, allowing extensive measurements (Kowalski, Albert et al. 1992).

Temporary or permanent (*chronic*) electrode implants are on used on live animals (*in vivo* processing). The animals may be anesthetized, which of course makes them easier to handle, but may give a distorted picture of neural functioning. More recent research tends to use awake, behaving animals, giving a better picture of normal neural functioning; still, the animals often view impoverished scenes and are far from natural ecological contexts.

The workhorse of experimental neuroscience is the single electrode measurement. There has been a strong mutual reinforcement of this technique with the *neuron doctrine*, or *localist* processing: the idea that most representation and computation in the nervous system is performed by single neurons and via small, specific functional networks.

This technique has been supplemented more recently (beginning in the seventies) by simultaneous multiple-electrode techniques. Most of the experiments motivating oscillatory models are based on observations with two or more electrodes simultaneously recording with the same stimulus present. There is also the possibility of an intermediate technique of moving a single electrode to nearby areas while repeating the stimulus, but due to habituation or learning this cannot really substitute for simultaneous measurements. Observations with multiple electrode techniques have given insights promoting alternatives to the neuron doctrine, emphasizing so called “dynamic assemblies”, cooperative processing, synchrony, and variable coupling. Signal processing techniques have been developed to assess the correlations between nearby neurons.

In both single and multiple electrode techniques, the actual signal being measured is voltage induced by conduction currents. These include spikes and possibly graded potentials – many neurons do not produce spikes.

Moving up to medium or *mesoscopic* spatial scales, *local field potential* electrodes and electrode arrays have been employed to measure summed dendritic currents over hundreds or thousands of neurons. These currents are effectively a measure *of ensemble average spike density* (Freeman 2000). Some groups, such as Eckhorn and colleagues in Marburg, measure both spikes and slow wave potentials in the same system, allowing correlations between these levels of processing to be observed.

These local field potential arrays are a medium scale version of the older electroencephelogram (EEG) technique. Most EEG studies record external to the scalp, in single or multi-electrode (multi-channel) configurations. These signals are typically bandpass filtered and analyzed for temporal changes and for inter-channel interaction in various frequency bands. EEG signals are rather diffuse (many current sources sum at an electrode) and historically considered as noisy; to investigate stimulus correlated signals, a common technique is to repeat a trial many (e.g. 100) times, and sum the resultant waveforms. Noise components are assumed to cancel, with the resulting *evoked potential* signals showing correlations to perceptual and cognitive events.

The EEG was the first "whole brain" measurement technique; investigations at this scale are usually designated as imaging or mapping. More recent imaging techniques include MEG, which is costly but has higher spatial resolution relative to EEG. A variety of slower temporal, medium spatial resolution techniques (PET and fMRI) measure cortical blood flow, which is correlated to the activity in some area relative to a baseline condition. Finally, a variety of optical methods are in use, in some cases on awake behaving animals with implants.

One drawback of all imaging techniques is that networks consist of inhibitory and excitatory neurons, and activity *per se* does not distinguish between them. As we will see, many theories of large scale network operation depend on interconnected excitatory and inhibitory pools, with the resulting activity not always easily characterized in simple active inactive terms. A better understanding of mutual influences of long range (i.e. between areas) activity and its local effects on multi-channel measurements is likely to emerge in the near future, through efforts aimed at discovering the flow of information and causal influences between co-active areas (Kaminski and Blinowska 1991);(McIntosh and Gonzalez-Lima 1994; Taylor, Krause et al. 2000).

CLASSICAL, NON-CLASSICAL, AND DYNAMIC RECEPTIVE FIELDS

With all these techniques, experimenters must choose to define the stimulus presented directly or indirectly prior to recording. Especially for single and multi-neuron techniques, theoretical assumptions on the nature of the processing operations constrained the range of stimuli used for many years in what, in retrospect, appears to have been misleading fashion. Typically such studies were done with sine wave gratings of various orientation, frequency, and contrast levels. Under these circumstances, the concept of local, anatomically determined micro-circuits performing feature detection became well established.

The usual formulation of a feature detector neuron involves a local configuration of excitatory and inhibitory neurons called *on-center off- surround*. A neuron receiving such input will respond most strongly (i.e. produces spikes at a maximum rate) to specific frequencies and orientations, forming a *receptive field*. The set of neurons sensitive to particular orientations and spatial frequencies is considered as a channel, with the function of primary visual cortex essentially acting as a filter bank with some adaptive dynamic range correction capability.

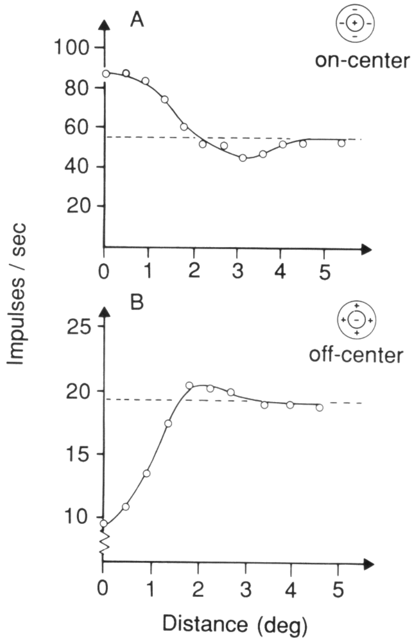


Fig. 3. The response of (A) on-center off-surround receptive fields and (B) off-center on surround in the retina as a function of the distance of a bar stimulus from the field center. From Maffei, L. (1968). "Inhibitory and facilitatory spatial interactions in retinal receptive fields." Vision Research **8**: 1187-1194; reproduced with permission of Elsevier Science.

The modern view of receptive fields, while still formulated as static structural elements in the vision system, is to consider simple cells as optimal two-dimensional Gabor filters. This representation has been demonstrated to achieve theoretical maximum in both frequency and spatial resolution (Wilson and Knutsson 1988), and is now commonly used as the "front end" for many successful computer vision applications. Texture analysis, in particular, is seen to be largely a function computable by the receptive field structures of primary visual cortex without additional back end processing (Bovik, Clark et al. 1989).

The receptive field concept has of, course, been subject to criticism. Fundamentally, practical limitations in biological recording preclude exhaustive sampling of the possible stimulus space. Harth showed that determining a receptive field via his Allopex biofeedback device, which changed a video stimulus until a local maxima was reached in the neural response, gave a completely different picture of the field shape than conventional grating methods (Harth and Tzanakou 1974). The modern version of independent, fixed (i.e. structural) Gabor channels has been questioned based on psychophysical findings that pre-trained vs. naive subjects exhibit different confusion patterns for a mirror image compound Gabor stimulus (Rentschler, Hubner et al. 1988); this is interpreted as evidence for cooperative interactions and stimulus dependent adaptation. More recent objections to the receptive field channel model of primary visual cortex function based on experimental observations of multi-neuronal correlations and temporal modulation in response profiles are treated below.

From a theoretical standpoint, the Gabor filter bank and subsequent wavelet representations represent an improvement in distinguishing textures with the same Fourier spectrum; this comes at the expense, however, of translation and rotation invariance. The concept of shiftable transforms and steerable pyramids represent one approach in computer vision to overcome these limitations (Simoncelli, Freeman et al. 1992).

Further studies with less constrained visual stimuli revealed a more complex picture. The presence of orthogonal components modulated the response of feature detectors; eventually it was determined that rather distant features could modulate the response of a classical feature detector (Allman, Miezin et al. 1985), and that a single neuron's output carried information on the global character of the stimulus at later epochs (Lee, D. et al. 1998). The term *non-classical receptive field* has been introduced to acknowledge the changing nature of this concept.

Systematic examinations of the temporal response in primary visual cortex in cat (Area 17) undertaken by Dinse and coworkers (Dinse, Kruger et al. 1991) revealed that receptive fields have a *dynamic (time-varying) orientation sensitivity* and size, inconsistent with a static structurally determined inhibitory surround. Further, four different families of neuronal subsystems were found within the area. Type I neurons (24%) showed an initial period of non-selectivity, with selectivity emerging after about 40 ms. Type II (34%) similarly showed emerging selectivity, but the selected orientation changed over time. Type III neurons (25 %) showed more conventional orientation tuning with no time dependence, but with broader tuning than that exhibited by the other types. In the time varying types, the response epoch (70-90 ms) with the sharpest response did not coincide with the highest rates of response.

The overall pattern of the response was characterized as a *damped aperiodic oscillation* of low frequency (6-20 Hz) superimposed by higher frequency oscillations.

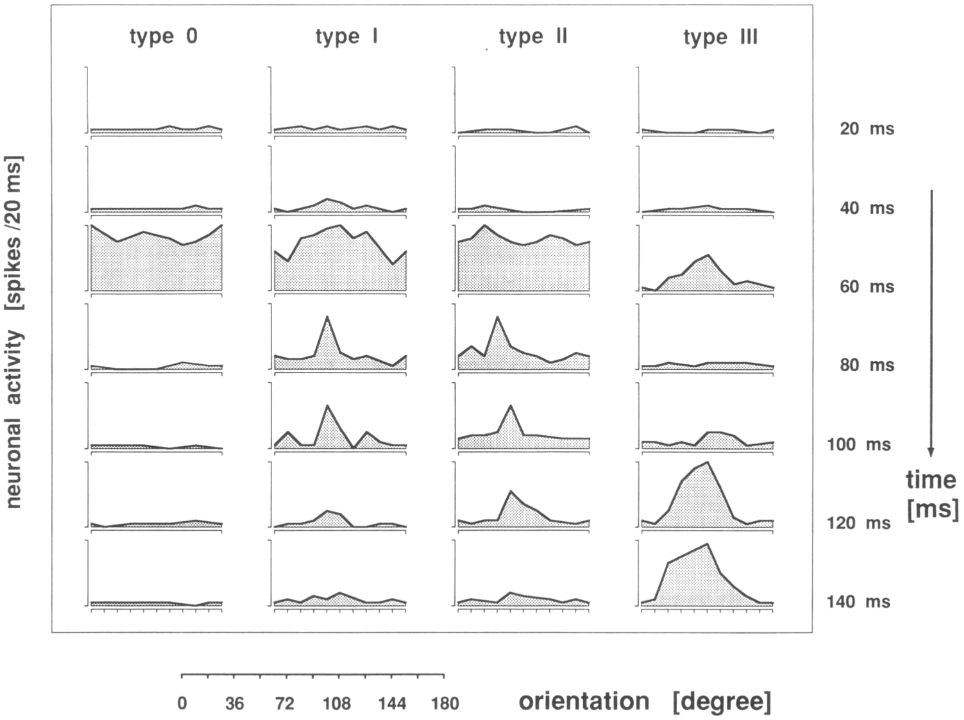


Fig. 4.Time varying receptive field structures. See text above for description From Dinse, H. R., K. Kruger, et al. (1991). Temporal Structure of Cortical Information Processing: Cortical Architecture, Oscillations, and Non-Separability of Spatio-Temporal Receptive Field Organization, in Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag: 68-104. Reproduced with permission of Springer Verlag.

These two key aspects of neural response identified by Dinse and coworkers – diverse types and temporal variation in response profiles - should be kept in mind when evaluating work on object-level recognition in inferotemporal cortex., which will be surveyed in a subsequent section. [13].

One study states that no difference in discrimination capability of single neurons is found by comparing fine-grained temporal epochs in IT cortex (Tovee and Rolls 1995) compared to the rate over an entire 400 ms response window. However, the possibility of different types of neurons, or a population response is not considered.

The models I will describe later correspond to large neuronal populations, but also exhibit this temporal response variability. There is a brief *optimal readout window*

---

[13] A recent cortical column *model* (Hansel and Sompolinsky 1996) also exhibits temporal variation in response profiles.

in which objects can be maximally discriminated. In my model, this readout window applies to the entire population of high level units.

THE PROBLEM OF NEURAL CODING

The receptive field idea was one of the first biological demonstrations of a plausible neural computation and code, though theoretical studies of neural coding and computation predated the demonstration of receptive fields by Hubel and Wiesel (Hubel and Wiesel 1962). I will take up the idea of coding a bit more formally in a subsequent chapter; for now it is worth stating that coding in communications theory normally implies a sender and a receiver. In experimental neuroscience, the complexity of the total system dictates that researchers focus on a small piece of a large network of processing and transmissions leading to behavioral distinctions. Particularly when the object of study is the spiking activity of a single neuron, the ultimate receiver is often unknown and assumed to be capable of using the information; in fact the *researcher interpreting the code of the neuron* is the only receiver known with certainty. In general, this *readout problem* is a major issue that must be addressed by neural modeling with claims of biological plausibility.

The receptive field or feature detector, grounded on the assumption of spike rate coding in single neurons, has been challenged based on this readout principle. Neurons show a high degree of variability even when presented with the same stimulus. Spike generation is precise and reliable, so this variability arises from fluctuations in input that drives neuronal firing. But this view is problematic due to the large (e.g. 10^4 for cortical pyramidal cell) numbers of inputs integrated; one would expect small variability due to the central limit theorem. Koch & Softky thus argue that neurons do not integrate excitatory synaptic inputs over a reasonable period of time. (Softky and Koch 1994).

Another readout-based challenge to rate codes is that the most reasonable window of integration for a rate code is the longest duration over which the stimulus can be approximated as taking a constant value. Some neurons only fire about one spike in such an interval, and thus arguably cannot encode and convey information through a rate code. If the neuron is part of a large population, this may be overcome by encoding (and reading out) a *population* firing rate (Abbot and Sejnowski 1999).

If the use of rates is viewed as problematic, what are the alternatives? Two major alternative paradigms are under intense investigation. One is that individual spike arrival times serve as a code, and in general spike arrival time coincidences are significant for algorithm level neural computation. Another is that significant computation is done by *cell assemblies* of one form or another, with population coding of intermediate results, memories and motor outputs.

The variety in potential coding schemes has been recognized for a long time, but practical difficulties in experimental methods have inhibited the investigation of many possibilities other than local rate codes. Bullock, for example, drew up a list of possibilities summarized in the outline below, and suggests that many or all of these coding strategies are employed somewhere in the nervous systems of different organisms (Bullock 1993).

Candidate Neural Codes

I. Subthreshold Graded Events
II. Impulses in Unit Neurons
  A. Representation by Identity of Active Fiber
  B. Codes based on Temporal Properties of Impulses
    Time of Occurrences
    Interval Statistics
      Frequency: Weighted Average
      Frequency: Instantaneous
      Frequency: Increment Above Background
      Frequency: Rate of Change
      Frequency of Firing / Missing at Fixed Intervals
      Coefficient of Variation
      Higher Moments: Interval Histogram Shape
    Temporal Patterns of Impulses
    Number of Impulses or Duration of Burst
    Velocity Changes in Axon
  C. Codes on Other Properties

Candidate Neural Codes, continued

III. Ensemble Activity
  A. Representation by Spatial Array
    1. Topographic Distribution of Active Fibers
  B. Codes Based on Temporal Relations
    1. Latency Distribution
    2. Phase Distribution
    3. Probability of Firing After Stimulus: PST Histogram Shape
  C Representation by Form of Composite or Multi-Unit Activity
    1. Evoked Potential shape
    2. Slow waves in ongoing EEG

The coding strategies treated in the present work are a small subset of these potential forms of coding and computation compiled by Bullock over 30 years ago from a workshop on the nervous system. The following diagram presents a restricted set of possible choices and locates some of the functional network types discussed in a *space* of choices for coding.

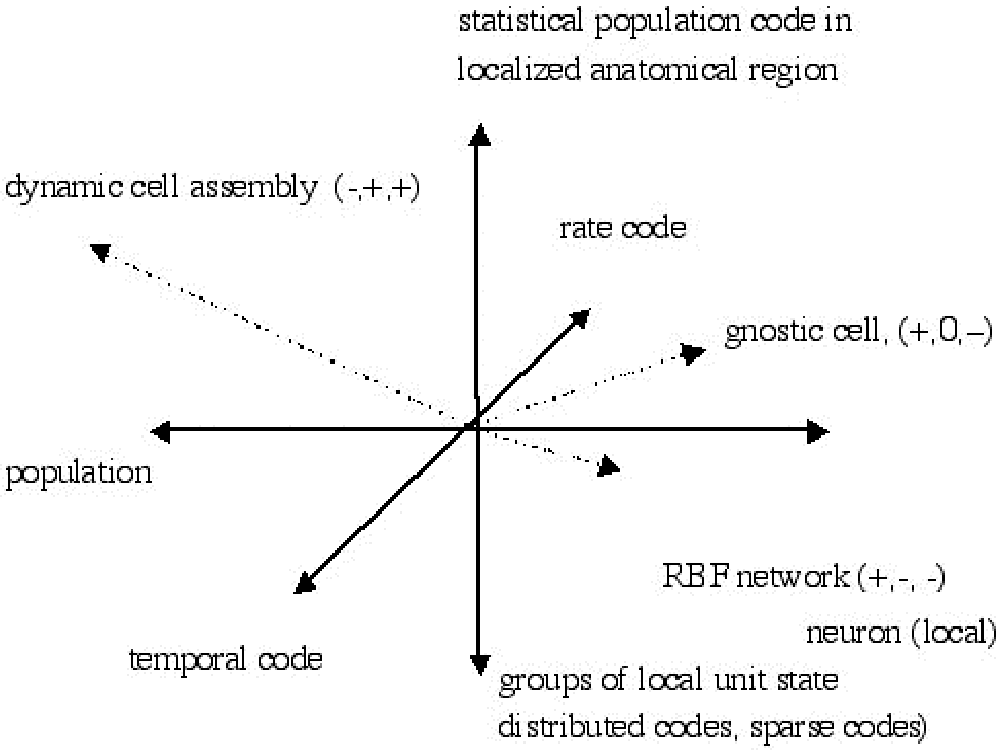


Fig. 5. A space of neural codes, with the dotted lines pointing to prototypical functional units. The x axis indicates whether coding occurs at single neurons or via groups. The y axis indicates whether, for group or population codes, whether the code is local or statistical. For the former, the code may be distributed but still depend on activation values at particular units; for the latter, statistics over the entire population carry the code. The z axis indicates whether the relevant units and the code are monotonic activation values, or involve temporal coding. To produce temporal codes, neurons act as spike timing detectors and / or constituents of distributed bifurcating subsystems in medium scale oscillatory networks, resulting in temporal patterns in phase dispersion and ensemble average frequency.

NEURONS IN CELL ASSEMBLIES

In theoretical and computational neuroscience, the alternative to coding and computation with neurons and small circuits has been the concept of a cell assembly.

Early research in cell assemblies concerned conditions for activation and stability in memory. Over time, generalized notions of assemblies have come to be replaced by a variety of specific functional circuits of connectionism, though a few researchers have always focused on larger scale models. Even in these models, activation was the dominant paradigm for analysis and simulations.

Much recent research implicated temporal correlations and synchronization for certain classes of coding and computation in early visual areas, with substantial evidence for this view reviewed below. To distinguish the classical assembly based on activation from new forms emphasizing spatio-temporal patterns, correlations, and synchronization phenomena, the term *dynamical cell assembly* was recently introduced (Fuji, Ito et al. 1996); however it was presented in the context of small spiking networks. This may be a worthwhile distinction, but need not be limited to spike correlations and small circuit models.

I suggest that in higher areas, such correlations are functional correlates or observables of larger scale computational processes, for which the simplest *algorithmic* explanatory level lies in non-stationary dynamics of coupled oscillator systems. This algorithmic understanding may contribute to an extension or revision of extant cell assembly concepts.

One contribution of this thesis is the demonstration that recurrent dynamics, spatial coupling, and temporal modulation of synchrony can support computations with relatively *simple and homogenous structure*. Synchronization is affected through modulations of network control parameters; these may reflect intrinsic rhythms, be generated in response to stimulus, or a combination of both. In this formulation, **collective variables** (i.e. population codes) measured on a set of oscillators, supplement classical coding concepts of **activation variables** on **localized units** or **localized cell assemblies**. Each of these oscillating assemblies may correspond to neurons rather widely distributed, such as a column of cortex with connections to interacting subcortical areas. Such assemblies may interface with rate coded activation assemblies for readout or memory coding.

In contrast to small spiking microcircuits, larger scale dynamical states are more easily correlated with medium and large scale electrophysiology (Freeman and Barrie 1994). The validation of connectionist models from single neuron or localized multi-channel recordings is a difficult task, because measuring all the neurons in a network is simply impractical.

For some researchers, the large amount of inter-region connectivity and the fact that most neurons project *both* locally and between regions, argues against local modular networks as envisioned by classical "small circuit" connectionism and classical modular cognitive science (Elbert, Ray et al. 1994,); (Mumford 1994).

THE CONTROVERSIAL ROLE OF RECURRENCE IN CODING AND COMPUTATION

We have seen that the anatomical connections may be described as reciprocal or recurrent at a variety of scales - within a single layer of minicolumns, between the laterally connected columns constituting a macrocolumn, between cortical areas and the subcortical areas which project to them, and between different cortical modules. In spite of this connection pattern, much theory in sensory and especially visual neuroscience is based on feed-forward models of computation, with lateral connections limited to special roles such as the on-center off surround receptive fields. In the context of feed-forward theories, if any role for local recurrence (i.e. within a column) is envisioned, it is to form population codes. These may be simple averages (to overcome the response variability cited earlier by Softky and Koch) or to exploit a Gaussian distribution of rates to perform function approximation (Poggio and Girosi 1990).

At larger scales, recurrent connections projecting back toward earlier sensory pathway regions have been ascribed roles in attentional processes, shifting and rescaling control (Van Essen, Anderson et al. 1994), or contextual modulation (Allman, Miezen et al. 1985) via linear signal processing mechanisms such as inhibitory gating of pathways or responses of receiving cells. Grossberg and colleagues conceive of some recurrent back projections as expectation signals (Grossberg 1980).

Several investigators working on problems of object recognition stress that based on the rates of recognition observed and the number of modular stages thought to be involved in processing, only feed-forward processes are possible. The putative site of invariant object recognition is the inferotemporal cortex, and stimulus invoked responses in this area are sufficient for experimenters to identify the stimulus form presented (from a restricted range of possibilities) in about 100 ms.

If the coding assumption governing both computation and representation is localized rate coding, the arguments are compelling. However, as we begin to review evidence and theory derived from multi-channel recordings in IT cortex, other coding and computation possibilities are raised based on spike coincidence, and deterministic changes in rates in larger networks of aperiodic oscillators. It is unclear that round trips between cortical modules, between cortico-thalamic areas, or widely separated lateral connections are required to achieve correlations between areas.

Local recurrence in columns, between neighboring columns (possibly exploiting gap junctions), and with *ongoing background input functioning as bifurcation control signals* are assumed to be the biological correlates of the models I develop in the next chapter. The ongoing background input may indeed be rhythmic signals from recurrent cortico-thalamic loops, but no signals need make the round trip. Freeman estimates the average time for signal transfer between pyramidal cells at distance of 1-2 mm at about 8 ms (Freeman 1992). This puts an upper bound of 10-20 iteration cycles for perceptual computations which produce a response in 100-200 ms.

TEMPORAL VS. RATE CODING STIMULUS PREDICTION FROM POPULATIONS IN PRIMARY VISUAL CORTEX

Kruger and Becker conducted studies with a regular spatial array of microelectrodes in Area 17 (primary visual cortex) of the cat, in which they assess the ability to predict which of 16 moving bar stimuli was shown (Kruger and Becker 1991); (Kruger 1991). An average response vector over several trials was computed for each stimulus, with the prediction based on a single non-averaged trial. The total response time was broken up into temporal bins, so that the vectors to be compared are of size:

$$channels * \frac{responseTime}{binWidth}$$

where channels is the number of electrodes (30), responseTime is the post-stimulus time interval recorded (300 ms), and binWidth is the time in ms. of a temporal bin on each channel. Two measures of vector comparison were used; the cosine of the generalized angle method gave better predictions than a sum of difference measure. This implies that scaling all bins by a constant rate increase does not change the stimulus prediction. The authors interpret this to mean that the temporal pattern codes stimulus identity but the rate codes the *importance* of the response. Further, they suggest that this helps resolve the apparent coding paradox related to how attention could modulate the early vision response if rate codes alone were used. The major finding is that the best predictions are obtained with more fine-grained bins, indicating that temporal codes rather than rate codes are the best population measure. A broad maxima from 20 to 80 ms is seen (figure below on following page).

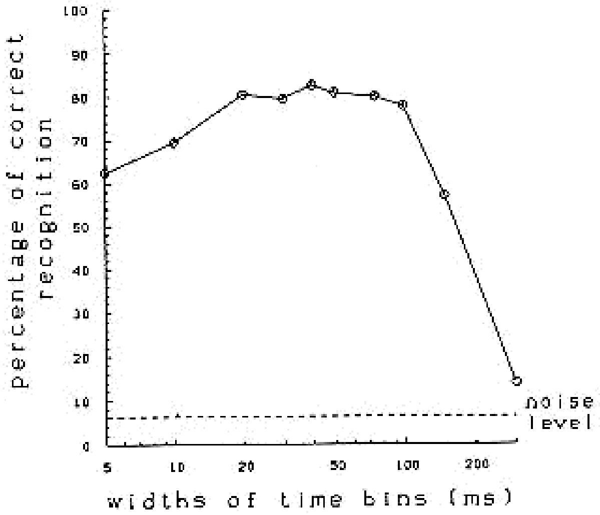


Fig. 6. The percentage of corrected responses as a function of the temporal resolution for averaging spike rates of a 30 channel spatio-temporal response vector for various time bins. Using a rate code (i.e. averaging over the entire 300 ms response interval) approaches a chance prediction level. From Kruger, J. and J. D. Becker (1991). "Recognizing the visual stimulus from neuronal discharges.", Trends in Neuroscience 14: 282-285. Reproduced with permission of Elsevier Science.

SYNCHRONIZATION AND CORRELATIONS: PHENOMENA AND ANALYSIS METHODS

I now turn from the discoveries and controversies surrounding single electrode measurements to a new set of ideas on coding and computation which have emerged from simultaneous multiple electrode measurements of temporal variation in neural microcircuits. The work of a few different groups will be briefly mentioned; an ambitious survey by Fuji et al. treats much of this work in greater detail, along with some modeling work on spatio-temporal coding (Fuji, Ito et al. 1996). Fuji et. al model these phenomena with spike detector neuron models with delay. In the next chapter on dynamics and methods, I will present models which embody similar spatio-temporal computation and coding principles but correspond to larger scale network dynamics.

To begin, I will note the definitions given in that survey for some common terms and attempt to augment them slightly. **Coincidence** refers to local events in the synapses of a *single neuron*, essentially the probability of receiving spikes in a given temporal interval. **Synchrony** refers to simultaneous (phase locked) firing of a *group* of cells. The two are closely related. Models of neural functioning involving precise spike timing posit that neurons are maintained near threshold by balanced excitatory and inhibitory input; in this state, they tend to fire based on sufficiency of coincident input with less regard for the average rate of firing on excitatory inputs. Coincidence is essentially a pathway to synchrony. Synchrony can be considered at the level of spikes or rates; for the latter, some integration window may be chosen and synchrony measured at the level of corresponding rates, even if individual input spikes on neuron events showed coincidence only at chance levels.

Correlation comes into play at a system level; Fuji et al. define a *dynamical cell assembly* as a group which, in response to the context of stimulus or another group is temporarily "bound by coincident timing of spikes". This is a more complex phenomena than synchrony, as it may imply *temporal variations* in correlation which lead to the formation of spatio-temporal patterns and to the formation of clusters of synchronization. Peaks in cross-correlation between members of such an assembly may occur with delay.

Historically considered as stationary processes, neuronal inter-spike intervals show characteristically Poisson distributions in cortex, but Gaussian distributions in motor systems. Given this fundamental irregularity, fluctuations in correlations are expected, and expectations and variances for correlation between two neurons can be defined. These are dependent on the firing rate, computed over some integration window. What is of interest to experimenters, then, are repeatable stimulus invoked correlations which *are significantly above or below the expected values*.

These measurements clearly depend on the time windows chosen to compute rates; classical cross-correlation methods assume that the neural signal is stationary. Methods have been developed to handle time varying rates by Aertsen, Gerstein, Vaadia, and coworkers (Gerstein 1988; Aertsen, Gerstein et al. 1989; Aertsen and Gerstein 1991), (Vaadia, Ahissar et al. 1991). The resulting data for pairs of neurons show peaks and troughs in correlation over time, which are interpreted as changes in *functional coupling*. This functional coupling or effective connectivity is in contrast to structural (anatomical) connectivity. It emerges rapidly, and is observed to be context dependent and dynamic on several different time scales.

An example of a two channel non-stationary correlation measurement for a neuron pair is shown in the following information rich figure. Each x, y point of the 100x100 matrix corresponds to the correlation strength for a different time lag between signals. Along the x and y axes are conventional post stimulus time-locked histogram (PSTH) spike counts for each neuron. This matrix has been normalized by subtracting the individual neuron PSTH cross product and dividing by the cross product of standard deviations of the individual PSTH. It is this normalization procedure which attempts to extract modulations in coupling from stimulus related modulations of firing rates.

The diagonal base of the "T" on the right half of the figure is the PST coincidence histogram, obtained by integrating over a 4 bin radius orthogonally from the diagonal of the left normalized matrix. The diagonal crossbar is a standard cross-correlogram. The two orientation and direction sensitive neurons from cat area 17 were exposed to moving bar stimuli, switching from non-preferred to preferred direction.

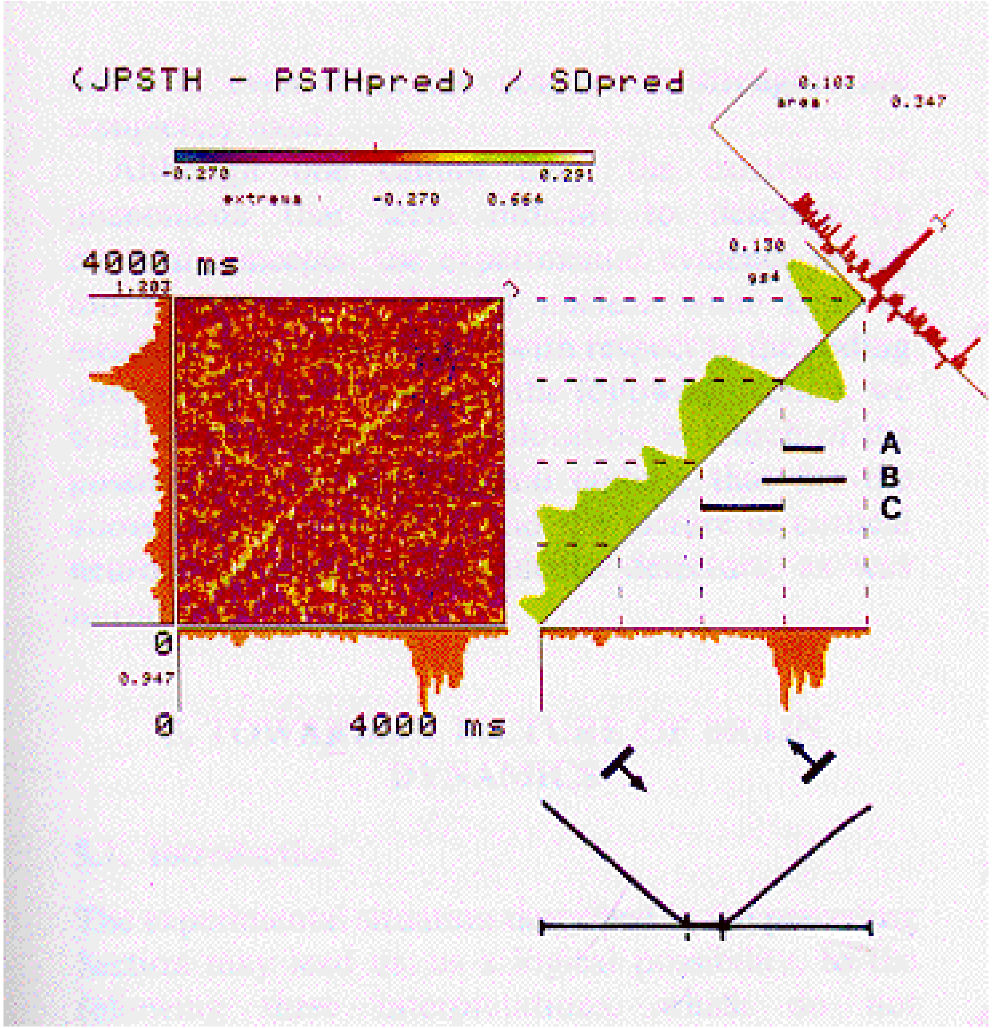


Fig. 7. Temporal changes in correlation structure for non-stationary signals. See text above for explanation. From Aertsen, A. M. H. J. and G. Gerstein (1991). Dynamic aspects of neuronal cooperativity: fast stimulus-locked modulations of effective connectivity. Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag. Reproduced with permission of Springer-Verlag.

Recent study of synchronization patterns in motor cortex indicates that synchronization and firing rate modulations are both important and may play different roles (Riehle, Grun et al. 1997). For one-third of 359 neuron pairs recorded, significant synchronization was observed, loosely time-locked to behaviorally relevant events. Synchronization in spikes was accompanied by firing rate modulation for external stimuli; for internal events (i.e. stimulus expectancy) synchronization occurred but significant firing rate modulations were absent.

PERIODIC AND APERIODIC OSCILLATIONS AT SMALL AND MEDIUM SCALES

The early history of cell recording was performed chiefly within the framework of feature detectors described above. From about 1985 on there has been a significant trend toward using multiple electrodes, documenting a variety of oscillatory phenomena. Seminal experiments of this type were rabbit olfactory system recordings (Skarda and Freeman 1987); (Freeman 2000), and observations of brief stimulus linked periodic oscillations in cat Area 17 (Eckhorn, Bauer et al. 1988; Gray and Singer 1989). Further observations of periodic stimulus linked oscillations in monkey primary visual cortex followed. The phenomenon has been controversial, perhaps in part by an early emphasis on periodic oscillations which are less common than aperiodic oscillations. With aperiodic oscillations, the apparently deterministic nature of the process and the role in cognition only becomes evident by coherence studies carried out in a behavioral context (Bressler and Nakamura 1993).

The interpretation of these results has varied, but usually there has been a focus on feature linking or binding for the periodic oscillations, while aperiodic oscillations are interpreted as possible temporal codes, or as products of deterministic chaos. As computational studies of coupled high dimensional chaotic systems have proceeded in parallel with the experimental work, the interrelated nature of all of these viewpoints has become apparent. Strongly coupled chaotic systems can become phase locked in periodic or aperiodic modes, and exhibit transient episodes of periodic oscillations; such systems will be described in the next chapter, and are the foundation for the *algorithmic level* of the similarity and stimulus invariance problems.

In the next several sections, some specific results are cited in support of this view. For more in depth reviews of oscillation and synchronization experiments and theory consult Bressler (Bressler 1995), Singer (Singer 1996), Eckhorn (Eckhorn 2000) for diverse perspectives; Elbert et al. emphasize chaotic oscillatory dynamics (Elbert, Ray et al. 1994), while Fuji et al (Fuji, Ito et al. 1996) emphasize the formation of cell assemblies.

The chief sense in which the present work offers a perspective distinct from these investigators, is an emphasis *on temporally structured (e.g. non-stationary) modulation of synchronization* in coupled chaotic systems[14]. I assert that this perspective holds some potential to explain many experimental datapoints, or at minimum to widen the

---

[14] It seems on reflection that the work of Bressler and colleagues cited earlier, involving modulations in synchrony over many bands (with alpha showing a slightly different profile) is closest in spirit to this emphasis.

discussion. This point of view also offers a mechanism for how distributed inter-regional networks might cooperate to perform computations in task specific brain states.

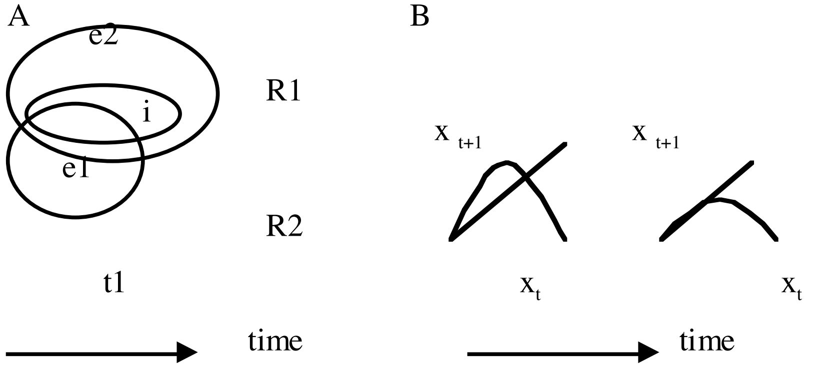


Fig. 8. Illustration of assemblies in two regions R1 and R2 cooperating to change dynamical parameters. A) The ovals represent coupled pools of randomly connected excitatory and inhibitory neurons. At time t1 one of the pools decreases its activity. B) Each curve on the right indicate spike firing probabilities in successive time windows; The height of the curve can be modulated by changing thresholds or excitatory inhibitory ratios. (See discussion of (Anninos, Beek et al. 1970) in the following chapter). The two curves correspond to dynamics before and after removal of the R2 excitatory pool at time t1.

PERIODIC STIMULUS LINKED OSCILLATIONS IN VISUAL CORTEX

Gray and coworkers found that cells 17 mm apart in cat Area 17 (primary visual cortex) with similar orientation preference showed both oscillations and significant correlation for a long moving bar which passed over both cells simultaneously (Gray, P. et al. 1989). Weaker correlation was seen for two separate moving bars, while no significant correlation was seen for bars moving in opposite directions. The spike counts produced by the two cells were similar under all three conditions .

Later studies showed that such effects were replicable in awake cats and monkeys, and synchronization could be observed even across hemispheres. These have been called fast cortical oscillations, occurring in the gamma range (30-80 Hz).

Eckhorn (Eckhorn 2000) suggests that the classical receptive field concept should be extended or supplemented by *association fields* which capture feature context, and therefore larger scale spatial regularities in the input. Using synchronization dynamics and coding, these association fields represent small visual objects or parts of larger ones, extending over a few millimeters of cortical surface.

MODULATION OF COHERENCE BETWEEN REGIONS

Seeking to explain large scale integration of networks, Bressler proposed three steps required to map functional connections among cortical regions. Sites from each area should be sampled simultaneously, and all possible site pairs should be examined for synchronization. Studies should be carried out within the context of a well defined task; and the measures of synchronization should allow for the possibility of aperiodic synchronization.

Several interesting findings resulted from this approach. Synchronization of distant (frontal sites) could occur at roughly the same time it appeared in visual sites, indicating that serial cascades from visual areas to frontal played little role. Synchronization appeared in episodes lasting from 50-299 ms. These episodes were broadband, not limited to γ synchronization seen in visual cortex. Differences were seen in the GO (motor response) and NO-GO (response withheld) conditions, particularly for non-visual sites. This is interpreted as indicating a functional role for synchronization.

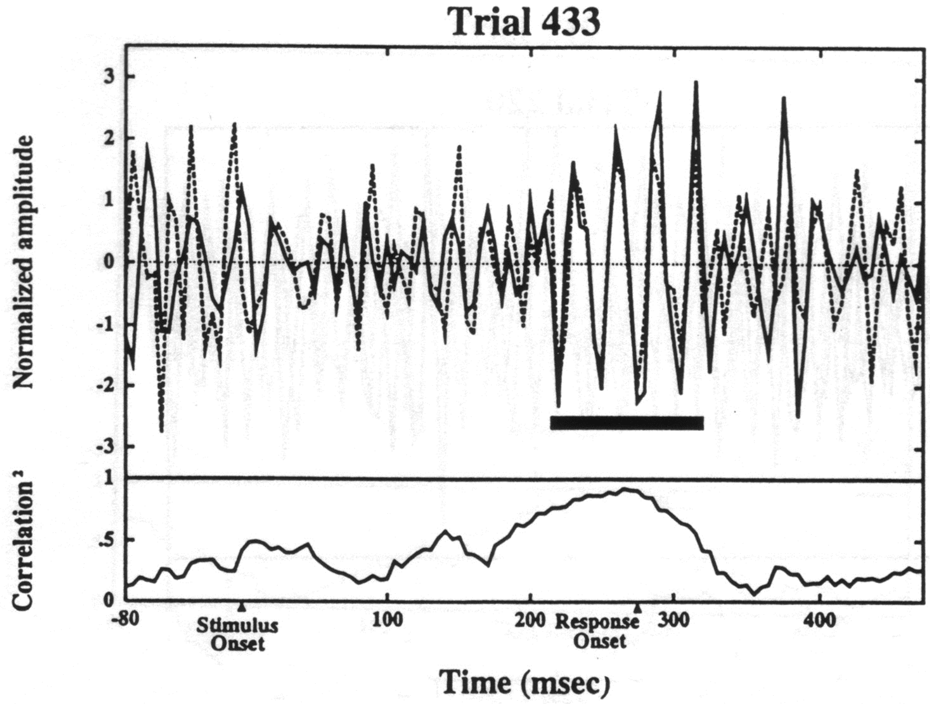


Fig. 9. Local field potential, filtered from 30-80 Hz, for a single trial, showing modulations in coherence of IT and striate cortex. In the framework of the Soca model described here, these modulations might correspond to transfer of phase synchronized contour border shapes from V1/V2 to IT (stimulus onset to 50 ms) in a restricted subspace of the dynamics, followed by subspace desynchronization (reduced local coupling), and subspace synchronization epochs. From Bressler, S. L. (1995). "Large-scale cortical networks and cognition." Brain Research Reviews 20: 288-304. Reproduced with permission of Elsevier Science.

SPATIOTEMPORAL OSCILLATION PATTERNS IN POPULATIONS

Walter Freeman and coworkers have refined experiments and a theory of medium scale (mesoscopic) neural function in the olfactory system over the course of nearly 40 years, in one of the most comprehensive, multi-scale research programs in neuroscience. Single neurons play little role in this program, with the emphasis on group activity, and with oscillations apparent at higher levels and observable in EEG. Individual activity is important chiefly in forming a transfer function of pulse to wave mode in dendritic mass (neuropil) groups. These groups are exclusively either excitatory or inhibitory, with primitive oscillating groups emerging from connecting the two types. Beginning in the mid 1980s, their interpretation of olfactory neural dynamics emphasizes spatial patterns of amplitude modulation, possibly as a manifestation of underlying chaotic dynamics. Further research has complicated what initially promised to be a novel and possibly comprehensive theory of neural function.

Skarda and Freeman documented physiological evidence for complex dynamical behavior in the rabbit olfactory bulb and provided an analytical model and numeric simulations with good fit to the EEG signals measured *in vivo* (Skarda and Freeman 1987). One key result was that the encoding of learned perceptions are "wings of chaotic attractors" of the global network, in contrast to the point attractors of Hopfield and feed forward networks. The term wings, to my knowledge, rarely appears elsewhere in dynamics literature; perhaps a more mainstream and contemporary reading would be dynamical motion with reduced phase space volume, exhibiting the underlying unstable periodic orbits.

Two key roles for the intrinsic chaos of the olfactory bulb are envisioned. The network cycles between highly chaotic and convergent (quasi-periodic) *phase regimes*, corresponding to exhalation and inhalation. The background chaotic state is deterministic, but is modified by each new perception-learning cycle of the network. The chief claim made for the function of these chaotic-period cycles are that they provide a *novelty filter* to stimulate motivation; a novel perceptual field will *not* fall into an existing cyclic attractor, but instead falls into a so-called *chaotic well*. Entry to this dynamics triggers a change in the network parameters to allow formation of a unique attractor and category, which is then learned by weight modification. It is also conjectured that the chaotic state provides rapid and unbiased access to existing attractors, and further that the chaotic half of the cycle avoids any accidental entrainment of attractors due to co-activation of neural assemblies from intrinsic structural resonances.

Regarding the theory of binding by oscillations, Freeman states that within his spatio-temporal pattern framework, the narrow band oscillations (the observations of Gray et al., Eckhorn et al., and subsequent workers described above) should be considered as *transient coherences* or intermediate products of integration; he asserts that they are not given as output (of a modular region) unless their phase and frequency are consistent with the whole (larger scale emerging pattern).

In recent writings, Freeman de-emphasizes his previous commitment to pure chaotic attractors, in favor of a more general non-stationary processing scenario (Freeman 2000). He has stated:

> The issue is not whether cortical dynamics is chaotic, that is not only unresolvable but unimportant. The issue is whether masses of neurons forming an area of neuropil are capable of establishing spatial patterns of cooperative neural activity with characteristic broad spectrum carrier over areas far greater

> than the mean length of dendritic and axonal arbors. These patterns must be formed in time periods much shorter than their 0.1 sec duration. [My work] shows these are reliable properties of laminated neuropil.

He has explicitly argued against *strongly* coupled chaotic oscillators as a model of brain dynamics recently (Freeman 1999), stating that "the synaptic coupling of multiple chaotic domains in the neuropil of cerebral cortex is not revealed in phase locking or synchronization at zero lag of local mean fields in the time domain".

The work described in this thesis, while using coupled chaotic oscillators as a starting point, does not assume phased locked synchronization or stationarity. In fact, nonstationary dynamics and avoidance of what might be called subspace synchronization are built into the learning system. While *phase* synchronization in the sense discussed in this review plays only an implicit role in the networks dynamics, this second meaning of synchronization derived from symbolic dynamics and graph theory, does play a crucial role. The concept will be further developed in the next chapter.

LARGE SCALE DYNAMICS: EEG AND MEG MEASURES AND THEORY

For nearly 70 years, large scale electric fields have been observed non-invasively in animals and humans, for both research and clinical applications. The electroencephalogram (EEG) is sensitive to both cortical and sub-cortical components, has poor spatial resolution, and incurs distortion in spatial patterns due to impedance differences in tissue. More recently, the magnetoencephalography (MEG) technique has permitted observation of magnetic fields, which offers several advantages (Basar 1998). MEG is selective to *cortical* activity, in contrast to EEG which mixes sub-cortical and cortical sources. The skull and extra-cerebral tissue are practically transparent to MEG. EEG requires selection of an arbitrary ground reference state, while the MEG field does not. The dipole moments measured by MEG are higher resolution; 10nA of current is estimated to correspond to the emission from 200-500 mm^2 of cortex (Hari 1997). However, the MEG technique is more expensive and less widely available.

Both EEG and MEG are generally broad spectrum, with changes in band power and inter-regional correlations studied for functional relationships to perceptual and cognitive processes. The following table summarizes the classical bands of interest; certain boundaries, especially gamma, seem to escape consensus.

Table 1. Major Rhythms in the EEG

band	frequency range (Hz)	behavioral and cognitive correlates
delta	0.5-3.5	deep sleep
theta	4-8	early stage sleep
alpha	8-13	mental activity, memory, attention, association
gamma	30-80	sensory processing

There are two main approaches to studying the interaction of large scale electromagnetic oscillations with cognitive, perceptual, or motor activity.

1. Searching for stimulus (or internal event) related trends in band power. This is the field of *evoked potential* studies. Correlates of behaviorally related *internal events*, such as decision processes or motor preparation events, may be studied in addition to responses to an external stimulus. Typically, many responses to a repeated experimental trial are summed, under the assumption that any signal is embedded in a large noise component which approaches zero mean over many trials. The resulting curves show changes in spectral content at characteristic latencies. More recently, there has been an emphasis on more subtle changes in the shape of spectral peaks.
2. Examining correlations in both phase and amplitude between regions thought to be active. Techniques showing changes in cortical blood flow may be used to guide this research, establishing spatial regions of interest for analysis.

Another approach pursued by several investigators over the last generation is to apply analysis methods of nonlinear dynamics, attempting to characterize the signals as low-dimensional chaos; see Elbert (Elbert, Ray et al. 1994) for a survey. This approach has fallen out of favor, since the methods require long periods of stationarity. Freeman, one of the early pioneers in this approach, has recently called that program a failure, given that dynamics appear to be non-stationary, irreducibly high-dimensional, and blending elements of determinism and stochasticity to serve specific functional roles (Freeman 1999).

Signal stationarity (constant parameters in the underlying dynamics) and the relationship of noise are constant source of difficulty and controversy in the field. Though many of the classical techniques (Fourier decomposition of signals, correlation analysis) assume stationarity, there is widespread agreement on the non-stationarity of EEG, leading to the recent emphasis on spectral shapes and correlation changes as inter-regional control systems. A variety of methods for analyzing non-stationary signals, or segmenting them into stationary windows have been developed (Gersch 1987; Pardey, Roberts et al. 1996).

Another problem in EEG analysis is that the background state of an organism is highly variable and partly determines the evoked response. To compensate for variability in evoked response dependent on the background conditions at the time of stimulus variation, Basar and coworkers more recent methodology consists of recording pre-stimulus EEG and post-stimulus evoked potential. A so called *enhancement factor* (the ratio of evoked to background power for frequency bands) is considered as a measure of resonance phenomena (Basar 1998).

A number of methods have been proposed which attempt to go beyond correlations, by determining the flow of causal influence between co-active and correlated regions (Kaminski and Blinowska 1991); (McIntosh and Gonzalez-Lima 1994).

INFEROTEMPORAL CORTEX AND OBJECT RECOGNITION

In this section, I will review some experimental data and interpretive approaches related to inferotemporal cortex (IT). This is the region - more accurately a complex of several regions - most implicated in invariant object recognition and memory formation, based on lesion studies and a long history of experimental work. The experimental data reviewed in this section represents both classical and newer concepts introduced above, with little consensus on the neural coding and representations strategies evident at the time of writing.

Architecture of the Ventral Visual Pathway

Neuropsychology has long recognized two visual systems operating in parallel, the so-called ventral and dorsal streams. Historically these have been considered the *what* (ventral) and *where* (dorsal), with both assumed to serially proceed to higher level cognition and motor input, but these distinctions have been revised somewhat (Milner 1999). The ventral stream provides visual contents of perceptual experience, and codes information in form suitable for processes like imagining, recognizing, and planning. The dorsal stream serves the immediate function of guiding actions from moment to moment, and needs to code information in a quick, ephemeral and view-specific form. The posterior parietal cortex is a major locus of the dorsal stream, with growing evidence for several modality specific visuospatial coding systems; for example, separate systems for the eye and the hand to reach to the same visual location.

Both streams are driven by the retina, proceeding through lateral geniculate nucleus (thalamus) to primary visual cortex (V1, also called *striate* cortex; area 17 in cat). Beyond this point, the streams separate, with the ventral stream terminating[15] in inferotemporal (IT) cortex. The following diagram illustrates this flow, along with the connection to thalamic and brainstem regions; note that reciprocal connections between adjacent pathways and subcortical regions are present:

---

[15] It is somewhat misleading to speak of the stream as "terminating", given the multitude of recurrent pathways; however, it is conventional to consider ascending (sensory) pathways and descending (motor) pathways, and this is arguably the last stop in the ascending pathway.

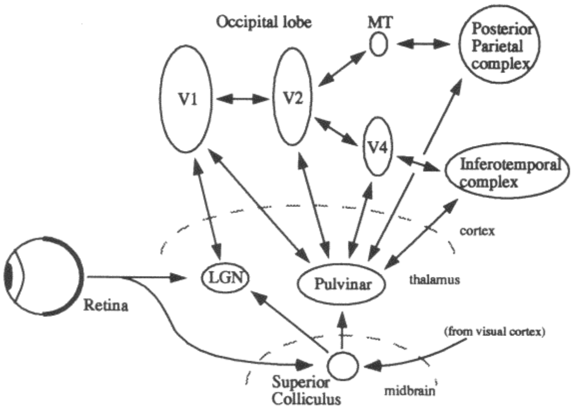


Fig. 10. Connections in the ventral pathway. From Van Essen, D. C., C. H. Anderson, et al. (1994). Dynamic Routing Strategies in Sensory, Motor, and Cognitive Processing. Large-Scale Neuronal Theories of the Brain. C. Koch and J. Davis. Cambridge, MA, MIT Press: 271-299. Reproduced with permission of MIT Press.

From V1, areas V2 and V4 perform further processing of form. V2 is known to be involved in contour completion; neurons have been shown to respond to illusory contours. V4 has recently been implicated in processing of contours; many cells showed preferential responses to certain classes of spiral, concentric, or radial forms (Gallant, Braun et al. 1993).

IT cortex encompasses posterior inferior temporal (TEO, pIT) and anterior inferior temporal (TE, aIT), which has a number of subdivisions including the superior temporal sulcus (STS). Of course, these have connections to non-visual areas. IT cortex has strong reciprocal connections to the amygdala (associated with reward systems and with social and emotional cues), connections to hippocampus (via the entorhinal cortex) and to prefrontal cortex. These connections are shown in the following figure; again reciprocal connections between cortical areas are the rule.

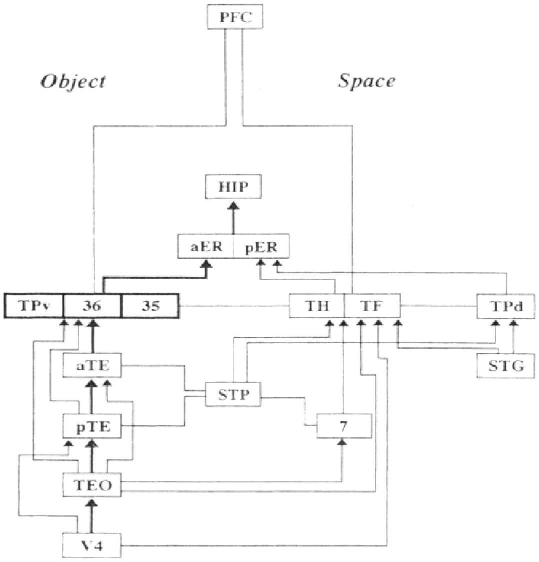


Fig. 11. Flow of information in subregions of IT cortex and connected cortical areas. pTE and TEO are implicated in processing of features, while aTE (anterior) is implicated in visual memory (ER = entorhinal, HIP = hippocampus, PFC. = prefrontal cortex). Connections to subcortical areas are not shown. Arrows represent reciprocal (feedback as well as feedforward connections). From Nakamura, K. and K. Kubota (1996). "The primate temporal pole: its putative role in object recognition and memory." Behavioral Brain Research 77: 53-77. Reproduced with permission of Elsevier Science.

One important aspect of the ventral stream is increasing receptive field sizes, as determined by classical single unit measurements. The following table adapts information from (Wallis 1994) and Rolls (Rolls 1992) on receptive field size and functional characteristics.

Table 2. Receptive field sizes and characteristics of ventral stream areas.

area	field width	functional characteristics
aIT (aTE)	50°	view independence
pIT(pTE)	20°	view dependent, configuration sensitive feature combinations
V4	8°	orientation and form processing
V2	3.2°	form, color, depth strips illusory contour completion (32% of cells sensitive)
V1	1.3°	orientation, frequency, location; movement direction in complex cells

Many neurons in IT fire at a slow spontaneous rate (3-4 spikes sec). Some of the research reported below emphasizes modulations in this firing rate *above and below* the background. A large literature exists on tuning properties evidenced by rate *increases* of IT cells to preferred oriented gratings, Fourier descriptors, etc. and to specific objects; many studies on preferential responses to faces exist (Rolls and Baylis 1986). For more in depth reviews of classical single neuron IT studies see (Rolls and Treves 1998) and (Logothetis and Sheinberg 1996). A review by (Nakamura and Kubota 1996) encompasses single neuron data and a variety of clinical studies.

The following discussion presents a sampling of recent work, emphasizing different approaches. Not surprisingly, I focus on work supporting temporal patterns and synchronization which underlie the computational approach developed in the next chapter.

Temporal Codes, Multiplex Filter Hypothesis, and Cross-Correlation

In a series of papers, a view of temporal coding via frequency modulation (the multiplex filter hypothesis) has been advanced by a group at U.S. National Institute of Mental Health. The earliest work (Richmond, Optican et al. 1987) examined single neuron responses to Walsh functions, finding that information was conveyed by *temporal modulation of spike rates*, based on an information theoretic comparison of principle components with simple rate coding. They suggested that multiple dimensions of a stimulus could be decoded from principle components of the modulated spike train. The figure below illustrates the stimulus pattern and typical averaged response, with a set of individual trials shown to give a feeling for the variability.

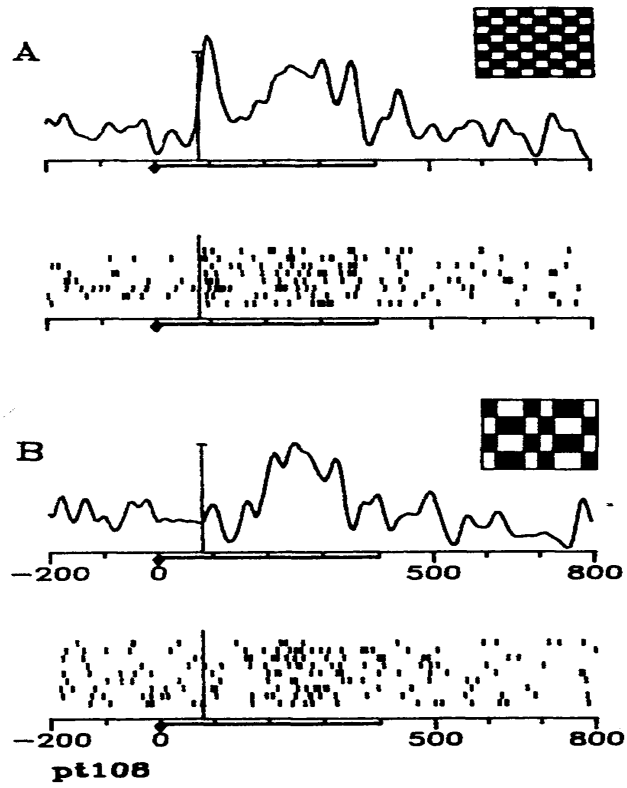


Fig. 12. The mean response of one neuron to two different Walsh patterns A and B, with rasters from individual trials plotted. Both excitation and inhibition are seen, in agreement with data of Gochin and coworkers described below. The bar height represents 50 spikes/ sec. The bar beneath shows the stimulus duration. From Richmond, B., Optican, L., Podell, M. and Spitzer, H. (1987). "Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex." Journal of Neurophysiology **57**: 132-146. Used with permission of American Physiology Society.

In more recent work recording from monkey TE (anterior) in a behavioral matching task (Eskandar, Optican et al. 1992), the same group found evidence that neuronal responses encode information about *both* the stimulus and the memorized target. In addition, local functional differences were found, with more information on the current stimuli in IT gyrus, relative to the superior temporal sulcus which appears to be biased toward target (memory) information. In concurrent modeling activity (Eskandar, Optican et al. 1992), the output of a pointwise multiplicative model of IT neurons was claimed to be a good fit to the recorded responses, suggesting that these neurons multiply temporally modulated waveforms arising from separate visual and memory systems in the *comparison* step of a visual memory task.

Arguments Against Temporal Encoding

Other researchers remain unconvinced by evidence for temporal coding in inferotemporal cortex. Tovee and colleagues (Tovee, Rolls et al. 1993); (Tovee and Rolls 1995) analyzed information available in varied short temporal epochs (100, 50, and 20 ms) of a 400 ms response series, and at different time offsets from stimulus onset. Again using information theoretic methods, they claim that 20 ms gives 30% of the information present in 400ms; the specific 20 ms interval chosen has little impact. More information is available at the start of the spike train than at the end, based on their analysis.

I argue that we should not be completely convinced by the latter demonstrations, for the following reason. If a cortical region is involved in the formation of a distributed representation, information about the stimulus must be present at the beginning of the computation in at least a subset of neurons contributing to the distributed representation; in fact, if the neurons participate in a meaningful way in the computation or are recruited into the distributed representation, the peak information about the stimulus *readable from a single neuron* should be available early, even though the code *used by the organism* is distributed and might require some recurrent cycles to form. Thus, the stimulus prediction *from an ensemble* at a later time might still exceed the prediction that can be obtained from a single neuron, for a wide variety of complex stimuli. Methodologically, the neural response needs to be measured on a behavioral task and correlated with the response to know whether the observed coding is actually the one used by the organism.

This is similar to the dilemma raised by Tsuda (Tsuda 1992), mentioned already in the psychology review section. It is unclear if different operational epochs for learning and recognition exist, or whether both are occurring concurrently; activity supporting both learning and recognition may coexist, but rapid task context shifts will activate the *effective connections* that allow one or the other to dominate. If the system is capable of rapid learning and forming representations we might *always* observe computational artifacts of the representation formation process, even though the representation and its decoding (readout) are distributed within the region or by interactions between regions.

Combination Codes in IT Columns

In a series of papers based on single neuron recordings (Tanaka, Saito et al. 1991; Tanaka 1993) and more recently optical imaging (Wang, Tanaka et al. 1996), Tanaka and coworkers examined the response of single neurons and trends in localized regions to stimuli of intermediate complexity. The methodology involves presenting progressive simplification of images to obtain the exact combination of primitive features (within a restricted object universe) which gives the maximum response over background rate. The data supports a hypothesis of *combination coding*, that specific combinations of features elicit responses in a small set of columns about .5 mm in area TE. Some of the optical experiments showed that center positions of the active areas move systematically with rotation of a face stimulus.

Tanaka acknowledges that critical feature columns cannot be the whole story, as the feature combinations could only represent a portion of complex objects. The spectre of the binding problem is raised again; it appears in slightly different form for TE, because the wide receptive fields will include multiple small objects, which must be discriminated from each other while their component features are bound. He raises possible solutions compatible with observations, including *aperiodic* synchronization, attentional enhancement, or the formation of loops of activity back to earlier stages in the ventral pathway (Tanaka 1996). He also notes that the optimal stimulus paradigm is based on a rate code assumption, but he has no evidence against the hypothesis of temporal coding.

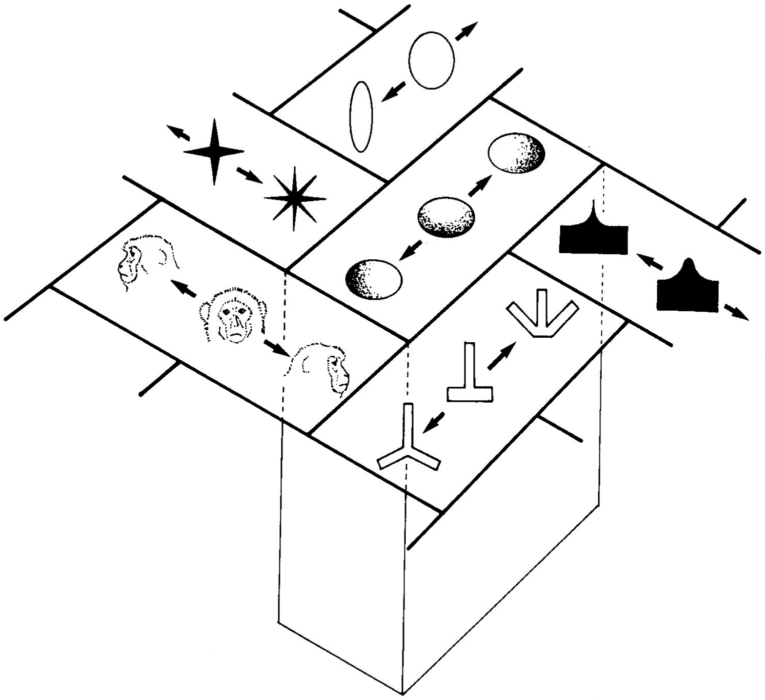


Fig. 13. A schematic view of overlapping critical feature columns hypothesized by Tanaka based on single neuron and optical recordings. From Tanaka, K. (1996). "Representation of visual features of objects in the inferotemporal cortex." Neural Networks 9(8): 1459-1475. Reproduced with permission of Elsevier Science.

Invariance vs. Broad Tuning to Specific Object Views

Logothetis and colleagues have examined single neuron responses to a variety of objects rotated in depth, including paperclip stimuli similar to those used as recognition targets in the present work (Logothetis, Pauls et al. 1994; Logothetis, Pauls et al. 1995;

Logothetis and Sheinberg 1996). They found that many neurons respond preferentially to a single or limited range of views of objects. Some responded both to a view and a near mirror image, while others did not. Tuning of the neurons to preferred views was found to be fairly broad. This data has been interpreted as evidence for RBF models as the Chorus network described previously (Logothetis, Pauls et al. 1994); (Edelman 1999). This is, of course, founded on the assumptions of rate coding and representation in single neurons that guide the study; the substantial evidence contradicting these assumptions, favoring temporal coding (or perhaps computation involving temporal modulations which *construct* a code) and population representations, is not discussed in that review.

Activity During Delay Period in Matching Tasks

While the other work in IT cortex described so far involves examining responses during the presentation of a stimulus, Miyashita and coworkers have studied the activity in anterior ventral IT cortex during the delay period of a visual short-term memory task. One of 100 possible fractal patterns is shown for 0.2 sec., followed by a 16 second delay, with a second pattern shown for 0.2 sec. and the monkey required to choose “match” or “no match”. Sustained increases or decreases from background rates were found in 95 of 144 cells. Of those, 77 showed variable frequency depending on the pattern, and many showed strong activity to only a few patterns (Miyashita and Chang 1988). Further experiments showed learning was crucial to generating the delay response; also, that the optimal response was often to rather dissimilar patterns, but substantial correlation in responses between successively presented patterns was seen (Miyashita 1988). Thus, it seemed that learning produced a kind of *temporal binding* of observed patterns. This is not surprising, since objects are observed from a series of viewpoints (as an animal moves around a stimulus or manipulates it) during the formation of a view-invariant representation.

These observations lend support to some models of object representation in the literature. They have been interpreted by Griniasty, Amit and coworkers (Griniasty, Tsodys et al. 1993); (Amit 1995) as evidence for the representation of objects as attractors; they show similar temporal correlation in patterns presented to a symmetric network with fixed point attractors. These are presumed to correspond to different sets of active neurons, but with correlations between successively learned patterns. The Visnet model of Wallis and Rolls (Wallis 1994); (Rolls and Treves 1998), is predicated in part on a short term temporal association of different views, similar to the observations of Miyashita.

It is less clear what role such delay activity should play in feed-forward recognition models like Chorus, which to date have not addressed primed search or matching tasks[16].

---

[16] The Soca model advanced in this dissertation has also not yet taken on primed search in a biologically realistic way, but I will attempt to sketch a strategy in the discussion section.

Stimulus Inference From Ensemble Responses

Gochin et al. (Gochin, Colombo et al. 1994) found that 5 stimuli could be inferred from a histogram of excited and inhibited responses of 40-50 cells in awake monkey IT cortex. The response was a rate integrated over the interval 100-500 ms post stimulus, rather than instantaneous statistics as in the present model. These results suggest an ensemble coding interpretation in terms of histograms of numbers of excited and suppressed units without regard to location. They found that reducing the integration window to 100 ms reduced correct inferences from the ensemble; this is in direct contrast to the results noted earlier by Tovee and Rolls on stimulus inference from single cells, which peaked in the first 100 ms.

Slow Oscillations Correlated With Stimulus

Nakamura and coworkers (Nakamura, Mikami et al. 1991) have observed oscillations in the anterior tip or pole of the temporal complex (TPv) , including areas 36 and 38, (areas not mentioned by the other researchers surveyed here). Neurons in this region responded to complex stimuli (photographs) but not to oriented bars. In the context of a visual memory task, the response of many neurons consisted of relatively slow oscillations, in the range 3-28 Hz with most occurring in the range 4-7 Hz. For some oscillating neurons, the oscillation frequency varied with the stimulus presented, i.e. 3.3 Hz for stimulus A, 5.6 Hz for stimulus B.

This finding is particularly intriguing in the context of the theory and algorithms I present in subsequent chapters. The computational model developed here identifies fast oscillations with computation toward a certain goal (metricity over partition cells or stimulus identity), and slow oscillations with modulations of synchrony; it predicts that slow rates are one of the controlling variables which serve to define a dynamical recognizer for a particular object. In memory tasks, these oscillations could play a role in modifying the response of local groups during comparison operations against incoming stimuli, perhaps allowing the activation of those TE regions observed by Tanaka.

MECHANISMS AND ALGORITHMS: A SURVEY OF THE NEURAL MODELING TERRAIN

The previous discussion focused on experimental methods and relatively data driven theory. I now turn to a brief discussion of *modeling* methods, many of which proceed from a similar abstraction for the neuron and an assumption that it is the locus of computation.

Modeling of neuronal processing is somewhat fragmented between various disciplines and scales of modeling. Some research is very application and mathematics driven, emphasizing proofs on the power and optimality of methods, with less regard for correspondence with the data of experimental neuroscience. Other approaches seek correspondence with higher level cognitive data and psychophysics, but not necessarily data from electrophysiology studies.

Most contemporary neural modeling is carried out in one of four major styles, with a certain amount of overlap between the camps.

1. Detailed electrochemical level models for neurons, including compartment models of dendritic processes, scaling up to small networks. Large networks (e.g. 10^4 –10^6 neurons) may be addressed through heroic parallel processing efforts. Generally the level of computation addressed is detection of features, or response to specific categories of spike input. These researchers would self identify as computational neuroscience.

2. "Neuron" computing units with idealized transfer functions, with weights on connections between neurons and possibly layers. The neurons may be connected in recurrent patterns, i.e. with direct or indirect feedback. Researchers would in the main identify themselves as connectionists or neural network researchers. In most recurrent networks, the end state of the computation is a fixed point attractor, hence the overarching term attractor neural networks. A new class of *spiking network* models has emerged recently, with behavior and computational complexity intermediate between the low-level models of computational neuroscience and connectionist models.

3. Models based on similar assumptions about single neuron dynamics as the first two categories, but typically with time delays, mixtures of excitatory and inhibitory nodes, or other parameters which result in *oscillatory* behavior, possibly including chaotic or other complex forms of oscillation. The literature of such models can be found under the topics neural networks, computational neuroscience, mathematical biology, biological cybernetics, and biophysics. Since the computation and coding usually results in dynamic or oscillatory states, these are sometimes called *dynamical networks* or *dynamical pattern networks*.

4. Oscillatory models with non-monotic or chaotic transfer functions defined for neurons or large neuron groups, with the oscillating units connected in regular spatial patterns. Connections are usually modeled as coupling strengths rather than weights, though in some cases these can function essentially as weights. The transfer functions for the groups may be more complex than the widely used sigmoidal activation. The concepts of excitatory and inhibitory connections often still appear, as in types 1. and 2. This is something of an emerging research field; much of the publication is by researchers in physics, centers for nonlinear dynamics and complex systems, and a few centers of dynamical neuroscience. In electrical engineering, the term cellular neural networks (Chua and Yang 1988) is used for very similar networks, with a greater emphasis on analog computation.

The style employed here is closest in spirit to the third and fourth types. Both types 3 and 4 have been described as *statistical mechanics* approaches. Physicists developed the statistical mechanics strategy for modeling large ensembles of similar elements. In this strategy, each state variable corresponds to the aggregate response of many such elements, and each parameter in a model may also capture in a single number a quantity arising from a distributed set of objects. For example, the bifurcation parameters in the present models are interpreted as excitatory-inhibitory ratios or delay time ratios of more detailed networks within each oscillating map site. Connections between sites or units at this level are not intrinsically excitatory or inhibitory, in contrast to connections between single neurons. The responses of *populations* should not be thought of in terms of monotic *activation*, but are typically time varying; the distribution of time varying response values over large population may be important. Silent sites or sub-populations may play an important role in perceptual or cognitive functions. This concept of population coding has become well established in motor cortex, but has advocates in the sensory and higher or associative areas.

These models can typically be related more easily to the experimental literature of medium and large scale[17] electrophysiology than those of the previous categories. Population responses may not be detectable in the response of single neurons which may not fire in every "cycle", and cycles will clearly be less evident in aperiodic population responses. Yet if nonstationarities in the response – evidenced by changes in correlations and synchrony at multiple scales – play a computational role, the modeler can use such changes as hints toward the type of nonstationarities likely to be effective in a task. This is the approach taken in the next section, where the linked observations of local aperiodic oscillations and large scale changes in synchrony directly motivate the dynamics and constrain the search space in a network.

NEURAL NETWORKS: AN OVERVIEW

I have now surveyed, albeit at a surface level, several research areas in system level neuroscience. As noted, there is often a reciprocal influence between theory and the experimental observations; but so far much of the discussion here has been data driven. Now I will turn to some discussion of neural modeling techniques and concepts. The goal here is to understand how these relate (or fail to relate) to the biology surveyed above, and to situate the recurrent, time-varying chaotic networks that are the central focus of the present work.

Activation Functions and Topologies

A network consists of processing units (nodes, cells) connected in some physical organization or graph. Certain typical directed graphs or topologies have been studied, such as feed-forward networks with hidden layers and recurrent networks (i.e. acyclic graphs), with feedback connections from a unit to itself or to a unit or units upstream.

Each unit performs an operation on one or more input signals, sending the resulting activation function value to its output function. While a wide variety of functions have been studied, until recently the combined activation output dynamics have been *monotonic*, i.e. the output function is strictly increasing with increasing (excitatory) input. This is viewed as a natural model of rate coding, with most computational power deriving from weights on input connections, specific topologies, and learning or self organization processes which update weights. While units are disclaimed from corresponding directly to neurons, it seems hard to escape the association of a unit with a parallel pool which attempts to account for neuronal variability through averaging.

Historically, the monotic activation functions are not time varying functions on the inputs. Recently some models incorporate delay between units, with resulting

---

[17] By medium scale, I mean studies involving multi-channel spike studies, arrays of local field potential macroelectrodes directly on cortex or optical methods. By large scale, I mean studies involving scalp recorded multi-channel EEG or multi-channel MEG.

interesting effects on dynamics and spontaneous formation of assemblies. Chaotic behavior at the network level is one possible result.

Non-monotonic activations have also appeared in the literature. When used as an associative memory, non-monotonic networks have been shown to have increased storage capacity compared to a comparably sized network with monotic activation (Yoshizawa, Morita et al. 1993).

For recurrent networks, one recent trend in research has been to investigate the computational power of particular mapping functions (Pearlmutter 1990); (Moore 1998), in terms of the formal language complexity class they are able to recognize. Holden and coworkers proved that homogeneous diffusively coupled maps are less powerful than Turing machines, and suggest that anisotropy (in connections, coupling, or evolution rule) is needed to increase the computational power (Holden, Tucker et al. 1991). These spatially distributed maps are the architecture I extend to nonstationary or staged operation in a subsequent chapter; I am unaware of any similar investigations on the recognition power of homogeneous but nonstationary networks.

Weights and Couplings

The numeric values which propagate between nodes are scaled by weight values in typical networks; this naturally corresponds to ideas that learning takes place in rate coded networks through slow modifications of synaptic efficacy (effectiveness). Many models have also incorporated a concept of fast synapses, with rapid adaptation of some or all weights in a network based on input or correlations in the network. Such fast synapses are important in the formation of dynamic assemblies and correlation coding.

In networks with non-monotonic functions, similar scaling takes place on the inputs to units. However, since small positive or negative changes on inputs may lead to changes in the output of the opposite sign, the term weight is less appropriate. The term coupling often appears instead, but conventions for mathematical notation for weights and couplings are interchangeable.

Learning Strategies

Learning strategies are chiefly divided into *supervised* and *unsupervised* types. In supervised learning, the desired output state of the network is known, and a teacher signal must adaptively modify weights or other network parameters to match the output state by an error minimization process. In unsupervised learning, the output state or encoding of an input pattern is not dictated by the network designer, but is arrived at through some means such as satisfaction of competitive constraints, or correlations between previously activated paths and those activated by the current input.

Representation and Coding

A *sparse* code implies that activation of only one or a few output units is significant. *Distributed* codes imply that the values of all output nodes are important. A code is *fully distributed* if the values of all output units must be known, i.e. the distribution of activities is balanced for all possible encoded values. A code may be both sparse and distributed, when the percentage of units active for each separate coded element is much smaller than the total.

The term *population code* appears less frequently in modeling, but has been the subject of some interest in experimental work and associated modeling. A population code is distributed code in which the *statistics of the population* response code the information about stimulus or the result of some computation, but not the activation patterns of particular output neurons.

Functional Classifications and Putative Biological Roles

Apart from the basic distinctions on the basis of activation functions, topologies, and of learning and coding styles, artificial neural networks have been categorized on a *functional* basis. Rolls (Rolls and Treves 1998) describes several functional families and considers their plausibility as biological models; there is considerable overlap with the learning strategies. He also gives a set of constraints on plausible models for cortical computation, based on anatomical criteria and psychophysics.

Pattern association networks use supervised learning. A particular input pattern presented to the network gives a response through a feed-forward topology. A supervised learning techniques (e.g. back-propagation of error) usually updates weights in layers somewhat removed from the output stage where readout occurs. Rolls sees little biological role for such networks, as they violate the local learning principle of his cortical constitution: all factors determining alterations in synaptic weights to implement learning are present in the pre-synaptic and post-synaptic firing rates of each neuron.

Autoassociative networks perform *pattern completion.* After training, presentation of a partial or noisy pattern will recover the original trained pattern. These networks are also called *attractor neural networks*, with the Hopfield network the earliest and best known example. A large role for autoassociative networks, particularly for episodic memory, is envisioned. Specific brain regions which have been proposed to function in this manner are the hippocampus CA3 region and anterior inferotemporal cortex (Miyashita and Chang 1988). An important issue for the biological validity of attractor networks is the time required to reach an attractor (convergence time); according to the analysis presented by Treves and Rolls, biological networks might reach attractors as rapidly as 20 ms.

The third major functional type of network are *competitive* networks. The Self-Organizing Map and Radial Basis Function (RBF) networks are included under this type. Differential connections or initial weights distinguish input-output pathways; the competitive aspect of the network results from strongly activated output neurons inhibiting others. In the Self-Organizing Map architecture , associative modification of active inputs to active outputs occurs during presentation of a pattern, increasing the chance of future activation of an active set with similar patterns.

The RBF network is a hybrid type with a competitive hidden layer (but not winner take all) feeding an output layer using supervised learning. Each hidden layer unit has a Gaussian activation function, giving a maximum response when an input vector is centered at its weight vector. In his review, Rolls suggests that it is not clear how Gaussian activation would be implemented biologically; however, recent investigations of spiking networks indicate that spike arrival times can implement such Gaussian units, and local learning based on pre- and post-synaptic firing times is possible (Natschlager and Ruf 1998).

This concludes the list of network functional types surveyed by Rolls. He alludes to the observations described above on oscillations in visual cortex and theories of feature linking, but generally dismisses the binding hypothesis on the grounds that synchronization processes would be too slow. His analysis assumes that direct communication is required for synchronization. However, in the next chapter I will show that, given a regular spatial arrangements of coupled chaotic oscillator units with uniform parameters, synchronization may occur *without direct connection between units* simply by deterministic dynamics tending toward synchronization, operating on similar local configurations (i.e. oriented lines or contours) in the stimulus space.

While the three artificial network types presented map well to distributed rate code theories, I have emphasized other temporal and population coding strategies, with evidence for their existence in cortex presented. A correspondingly rich variety of additional oscillatory or dynamic neural network models have been proposed by others, which align more with ideas of temporal, population codes. I now briefly survey some early oscillatory models and will focus on a few in more detail in the next chapter.

DYNAMICAL NETWORKS: OSCILLATIONS, CHAOS AND EXOTIC ATTRACTORS

Particularly since the recent experimentally driven interest in synchronization and modulated synchronization, computational studies of many types of *oscillatory networks* have been performed; recent workshop volumes include (Taylor and Mannion 1992) and (Brown, Levine et al. 2000). Different authors vary in their emphasis; some focus on mathematical properties, others on biological realism,, still others on modeling of perceptual and cognitive phenomena. I will emphasize modeling of perceptual phenomena with oscillatory networks in the following chapter on dynamics and methods; here I will note a few significant early efforts and surveys.

In an early modeling paper motivated by the experimental observations of synchronized oscillations mentioned above, Mannion and Taylor discuss both binding and *separation* of bound objects with oscillations, outlining parallel and serial strategies for separation. Separation refers to the need to handle multiple bound objects in parallel. The *parallel* strategy involves separating objects by frequency, while the *serial* strategy involves time slicing the activity of objects (i.e. phase separation). They indicate that little biological support for the frequency strategy (without specifying what prior studies support or fail to support any segmentation strategy). Developing a firing rate model with a regular spatial array of units producing sinusoidal oscillations, they demonstrate a serial strategy of alternating phases (Mannion and Taylor 1992).

Grossberg, a pioneer of many network architectures for specific psychological phenomena, published an early article on feature-linking with synchronous oscillations [Grossberg, 1991 #6]. The next chapter will discuss several more recent variations on the theme of synchronization and segmentation of scenes. For now, I turn to other oscillatory models emphasizing *memory formation and perception*.

Freeman and colleagues have implemented and refined over time a detailed, multiple-scale network model called the KIII model; they have validate the model with experimental results obtained in rabbit olfactory cortex and rats. The base organizational units are the K0 models for neural mass, having excitatory or inhibitory output. A K1 set consists of mutually inhibitory or mutually excitatory pairs. These are incorporated into KII sets with both excitatory and inhibitory elements. A KII set corresponds to the activity of a recognizable anatomical unit in the mammalian olfactory system; the olfactory bulb, anterior olfactory nucleus, and prepyriform cortex are separate KII sets. Finally, a set of KII sets are connected in a regular spatial arrangement with feedback pathways incorporating delay to make the KII model (Yao and Freeman 1990); (Kay, Shimoide et al. 1995). The architecture results in spatiotemporal patterns qualitatively similar to those observed in the olfactory system.

The Freeman group has historically employed numerical integration of ordinary differential equations as a modeling technique; a recent reformulation of the model to the discrete space and time, coupled map style was undertaken (Kozma 2000).

The olfactory system differs considerably from visual cortex of course, notably in the lack of spatial structure in the input transducers. Other investigators have explored the interaction of chaos, synchronization, and orientation sensitive structure at a similar modeling granularity (i.e. the detailed neuronal models of computational neuroscience). In a model with neurons of various orientation preferences in local pools with long range excitatory coupling to other pools, Hansel and Sompolinsky established first the important characteristic that synchronization between neurons of similar orientation preference could occur in a few cycles (Hansel and Sompolinsky 1992). In later work they have introduced inhibitory connections and modeled orientation tuning effects in a visual cortex hypercolumn (Hansel and Sompolinsky 1996). That model exhibits time varying orientation sensitivity and spatio-temporal excitation of intermediate neurons when instantaneous changes to the input stimulus orientation are made, matching effects seen in cortex .

Tsuda developed a network displaying complex, non-equilibrium temporal behavior during the recognition or memory recall dynamics, and postulated many possible cognitive functions for chaotic dynamics (Tsuda 1992). He has more recently used the term *exotic attractors* to distinguish these from fixed point or periodic oscillatory attractors. In his model, a Hebbian learning stage establishes attractors in the network based on intrinsic phase correlations arising from input stimulus vectors. During recognition, presenting one such learned pattern to the network *through an alternate pathway* causes it to cycle through the previously learned attractors in a pseudo-random order, with spurious transitional states between visits. This dynamical behavior serves to semantically link previously memorized attractors into more complex combinations, providing an alternative means of binding component perceptions into a composite memory. Compared to the binding by synchronization proposals, Tsuda's network seems to have attracted less attention and criticism; this may be in part because the concepts and dynamical behavior are unfamiliar to many experimentalists.

SUMMARY

The various threads of research introduced above form a network of interrelated issues, but were selected to underscore the following points

1. Neurons and local fields exhibit complex oscillatory behavior and synchronization phenomena. These may be involved in computation and coding, in contrast to computation and coding with rate coded and modulated “activity” networks. Using such dynamics in support of algorithmic processing is a relatively new field, relative to more established connectionist models employing the rate code assumption.
2. At neural and larger scales, “stages” of processing within the same regions and local circuits are seen, in contrast to serial presentation from one computational stage to the next. These are evidenced in several ways, from changing amounts of information seen in different time windows, to differences in sensitivity based on context, and changes in correlation structure. Interpretation of functional roles for such stages is not very advanced.
3. Stages may be related to changes in the correlation structure observed in signals at different scales and in the synchronization measures of larger scale (inter-regional) networks. This provides a new way of interpreting the interactive role of multiple regions observed in imaging in contrast to a serial presentation.
4. Oscillatory behavior may be aperiodic, which is characteristic also of coupled, spatially distributed chaotic dynamical systems. The study of networks of chaotic *units* is relatively new, but shows promise for modeling aperiodic oscillatory phenomena. Synchronization increases with increased coupling of such systems.
5. Stages may be related to major operating rhythms; in the view of Basar (Basar 1998), major operating rhythms (theta in prefrontal, alpha in occipital visual areas) control the evolution of the faster bands. In agreement, I further suggest that this control may be interpreted in terms of changing control parameters of nonlinear oscillator arrays to perform specific computations. The computations often involve changes in dimensionality corresponding to desynchronization and synchronization. This is in contrast to classical gating, excitatory, and inhibitory control architectures.
6. Retinotopic maps of organized columns are the regular organization seen widely in cortex. Network models of object recognition such as Chorus do not map to this organization, instead using retinotopic arrays of feature detectors as a front end to a specialized structural network. The model here uses a regular array of units to perform a task traditionally considered part of high-level vision.

The network style described in detail in the next chapter, coupled map lattices, is a relatively new (Kaneko 1986) approach to modeling *spatially extended* physical systems, of the type classicaly treated with partial differential equations. In this emphasis on spatially adjacent nodes it differs from some connectionist models, which do not have a regular, localized spatial structure similar to the columns and hypercolumns of cortex . By using discrete time iterations rather than differential equations, the couple map style is closer to recurrent attractor networks in connectionism than to the detailed “small circuit” dynamics of computational neuroscience.

Because of the computational cost, detailed small circuit models are often limited to reproducing spike level input output, rather than approaching perceptual and cognitive phenomena. The efficiencies of coupled map modeling allow a direct attack on *perceptual* problems, while retaining many characteristic aspects of biological neural systems and signals. The gap between Marr's algorithmic and implementation levels is reduced, though the neural modeler still bears the burden to justify a particular map, like the chaotic function used here.

Two additional differences with most previous work in similarity and object recognition are in the relevant site of encoding in the network, and the nature of the dynamics involved in the encoding. The feature detector class of neural models use what has been called *place coding*, locating meaning in the firing rate of particular neurons. This leads to the well-known binding problem and combinatorial difficulties representing the large feature space.

Connectionist models use a distributed representation, typically sparse distributed coding across a vector of output units. Still, the individual units are significant (local or place coding) and the activation or encoding is usually sigmoidal. Further, the network dynamics at the end of a recognition process are typically a one shot response, or a stable attractor controlled by the network parameters (connection weights).

In this thesis, I use a population encoding across an array of chaotic units. This encoding is in accord with the data and interpretations of Gochin et. al., with the chaotic dynamics of the units resulting in *aperiodic (frequency modulated) time series,* resembling their observations, along with the single neuron "temporal codes" observations of Richmond et. al. Thus, there is some correspondence between the model presented here and observed micro-circuit dynamics in IT cortex.

However, the correspondence between the Soca network and function of IT cortex cannot be taken to be a literal one. The current encoding involves sampling of the transient orbits of a nonlinear evolution process across the entire network *at a particular time* in a structured, non-stationary dynamics. While such an instantaneous population rate code is envisioned in Bullock's list of possible codes, the readout and comparison methodologies used here are very un-biological. In my recognition system, the statistics of this sampled state are *numerically* compared with other such samples statically stored in memory by conventional digital, algorithmic procedures. Such an instantaneous code must be seen as an input to some other readout or memory formation process in a biological system.

Chapter 4: Dynamics of Spatially Extended Nonlinear Systems

As background for the experimental work in the next chapter, the fundamental concepts of nonlinear maps and chaotic systems are now introduced, followed by extensions to high dimensional systems, and particularly to uniform spatially extended systems[18]. The aim is not to rigorously develop the material, but to survey the field in enough breadth and depth so that the meaning of the networks and algorithms treated in this study becomes transparent. Many texts providing more depth in low dimensional dynamical systems exist; Ingraham (Ingraham 1991) is a concise survey. To date, no comparable surveys of high dimensional or spatially extended systems exist; furthermore, terminology varies considerably across different disciplines. A special issue of the journal Chaos with an introduction by Kaneko is one possible point of entry (Kaneko 1993), along with a collection of papers on various applications (Kaneko 1993).

There are two approaches to problem solving using complex dynamics. Given the form of a desired outcome, search methods in the control parameter space can be used without requiring deep understanding of the underlying dynamics. The treatment of dynamics, then, is to provide the rational for the search methods and constraints on search developed here.

The other approach involves analytical treatment of the problem so that exact solutions, or at least bounds on the state space are obtained. At the time this project was initiated, the prospects for analytical solutions to problems where the desired outcome involved state distributions on sets of coupled oscillators seemed remote. Accordingly, I hypothesized that search methods might prove effective, even in the absence of a strong theory on bounds of the technique or direct solution methods; the results presented in the next chapter support this strategy. However, techniques have emerged which might lead to more direct solutions than the search methods used here. I will mention relevant mathematical approaches briefly in this review so that it is a useful overview of the evolving state of the art.

In the following discussion, many important terms from the literature of dynamical systems are introduced. For the benefit of readers encountering this material for the first time, these are highlighted in bold type.

ONE DIMENSIONAL NONLINEAR MAPS: DEFINITIONS AND TERMINOLOGY

A map is an iterated difference equation

$$S_{t+1} = f(S_t)$$

where S is a real valued state , f is some function mapping S within a subset of the real number domain $\boldsymbol{R}$, and t is a discrete time step. Iteration implies that the result of applying the function at time t is fed back into the computation at to produce the result at

---

[18] The term complex systems, to the extent that it is standardized, refers to spatially extended systems.

time $t+1$. The sequence of states $S^1, S^2, \ldots S^T$ by iteration for t=1,2,...T , is the image of the map. The sequence of states preceding any state S^I are the pre-image of the state. The terms trajectories and orbits also appear in the literature for the sequence of states, with trajectories used for continuous systems and orbits normally used for discrete iterations.

Nonlinear maps use some nonlinear function[19] f, resulting in diverse types of asymptotic behaviors; these asymptotic behaviors are reached after a *transient regime* of variable number of iterates. The duration of this transient depends on the exact initial conditions as well as the exact function and parameters. This variety in transient length and complex structured of the trajectories approaching a stable state, supports the algorithms for forming representation spaces and performing pattern recognition tasks. I will explore this structure by some simple parametric studies in the next chapter.

An **attractor** of a map is the asymptotic state sequence after many iterations, if such an asymptotic state exists. The term attracting set or limit set is also used. The **basin** of an attractor is the set of all pre-image states which converge to the attractor, after some number of iterations.

One crucial distinction for a system is whether for a particular fixed control parameter, different inputs converge to a single attractor or to one of multiple attractors. For the logistic map used here, a single attractor exists for all input states, but the basin structure and transient sequences leading to the attractor are highly variable depending on the particular instantiation of the map (i.e. the exact value of the chosen control parameters).

A dynamical system with multiple coexisting attractors can be used as a model for perceptual and memory processes. Training a supervised neural network consists of shaping the dynamics evolution of a network through its parameters such that the attractor basins map input states into the categories (attractors) desired. This basin structure can be considered as an intrinsic categorization by partitioning the input states into categories corresponding to the attractors.

A well studied map used as a network node (cell, neuron unit, site) in the models described later in this chapter is the asymmetric logistic map:

The equation for the asymmetric logistic map is

$$S_{t+1} = 1 - bS_t^2 \ , \ \begin{cases} -1.0 < S < 1.0 \\ 0.0 < b < 2.0 \end{cases}$$

where b is a bifurcation parameter; changing this parameter forces a structured transition between phases following the sequence of attractor types, which are introduced below:

fixed point → limit cycle cascade of increasing period and instability → intermittency → chaos → {limit cycle cascade →chaos} ...

The changes in attractor type occur abruptly, even with smooth changes in the parameter. The brackets and ellipses indicate that beyond the transition to chaos, there

---

[19] A nonlinear function is one for which the solutions are not subject to the principle of superposition, i.e., the solutions do not add linearly to generate a new solution .

are windows of periodic behavior in the bifurcation parameter values surrounded by regions of chaos, and this repeats infinitely.

A **fixed point** attractor is one for which every initial state leads to the same value. A **limit cycle** attractor is a repeating state sequence of period *P*; all initial states lead to the sequence, though the **phase** of the sequence (relative to time t modulo *P*) may vary. A limit cycle attractor may also be referred to as a **periodic attractor**.

This state sequence is the simplest form of **oscillation**. In continuous systems theory and circuit analysis, the conventional meaning of oscillation is a limit cycle or periodic oscillation. In nonlinear dynamics, more complex aperiodic motion is also referred to as oscillatory, which may be a source of confusion in discussions between neuroscientists and nonlinear dynamics investigators.

A **chaotic** attractor is an aperiodic orbit which exhibits sensitive dependence on initial conditions. A system can be more or less chaotic, essentially a measure of how rapidly nearby initial conditions diverge. **Lyapunov exponents** can be computed for a system as a measure of nonlinearity. Since the rate of divergence varies over the set of initial conditions, a system is commonly characterized by the **largest** Lyapunov exponent over the full range of possible initial conditions.

Typically the transition points between these phase regimes are visualized by bifurcation trees for systems with one bifurcation parameter, or phase space plots for **coupled systems** with multiple parameters governing transitions between regimes.

Time series (orbit) plots under various initial conditions, the bifurcation tree for the map, and an example phase space plot are shown in the following figures.

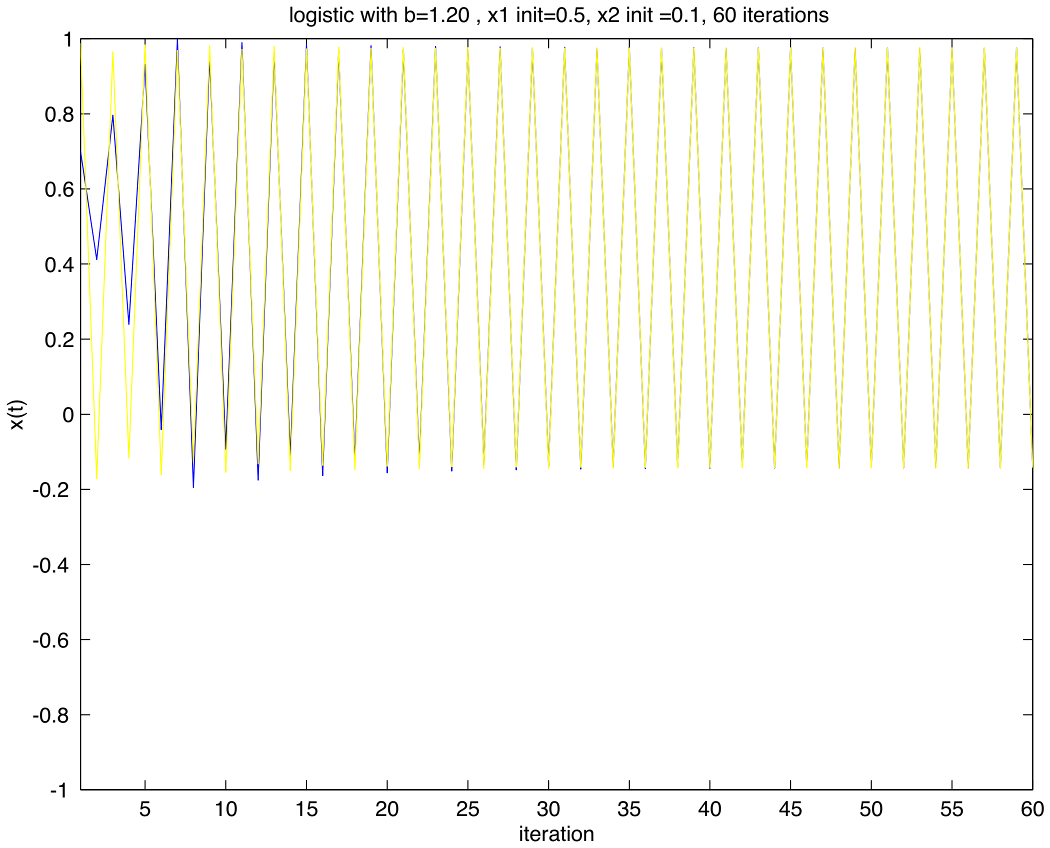


Fig. 14. Time series or *orbit* of the map $x_{t+1} = 1 - bx_t^2$ with x =0.5 and x =0.1 initial states overlaid. The bifurcation parameter b set to 1.2, leading to a stable period 2 attractor for any initial condition.

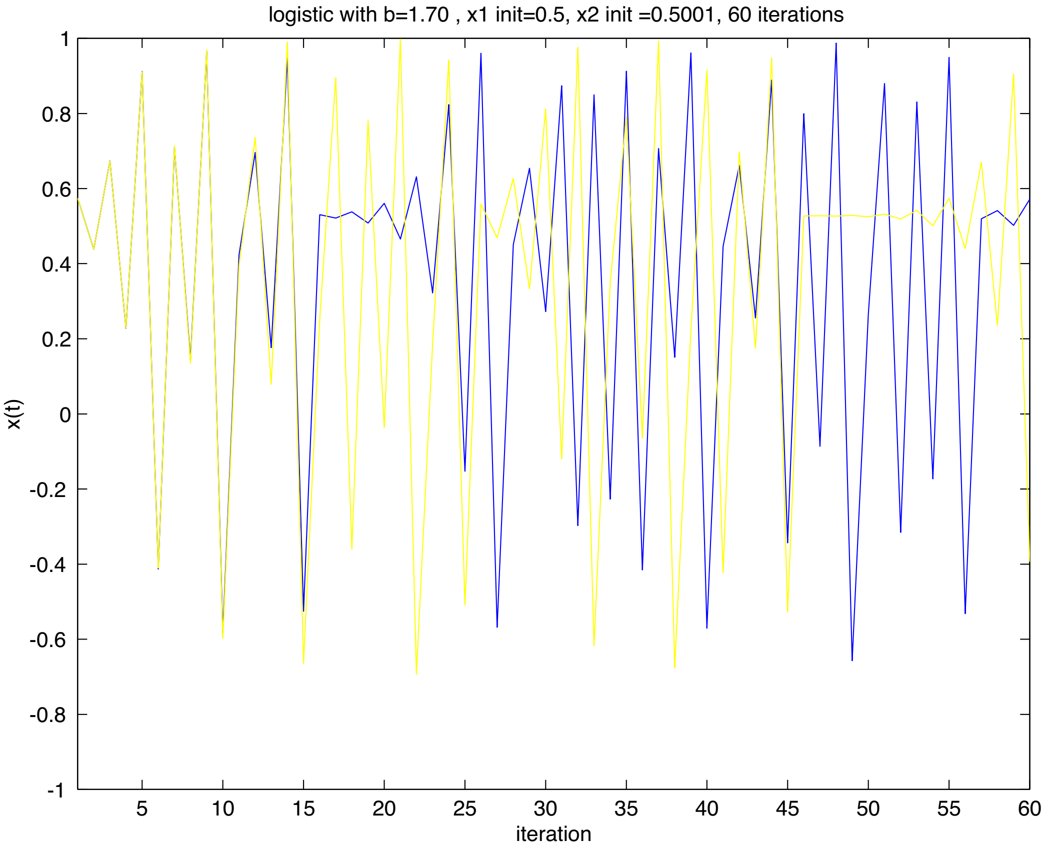


Fig. 15. Time series of initial states x1=0.5, x2=0.5001 superimposed with bifurcation parameter b=1.7, beyond the transition to chaos at b=1.544. The separation of initial conditions differing by .001 illustrates the phenomena of divergence of orbits of nearby initial conditions. The Lyapunov exponent is a measure of divergence (positive exponent) or convergence (negative); the plots here indicate divergent dynamics, while the previous figure is convergent.

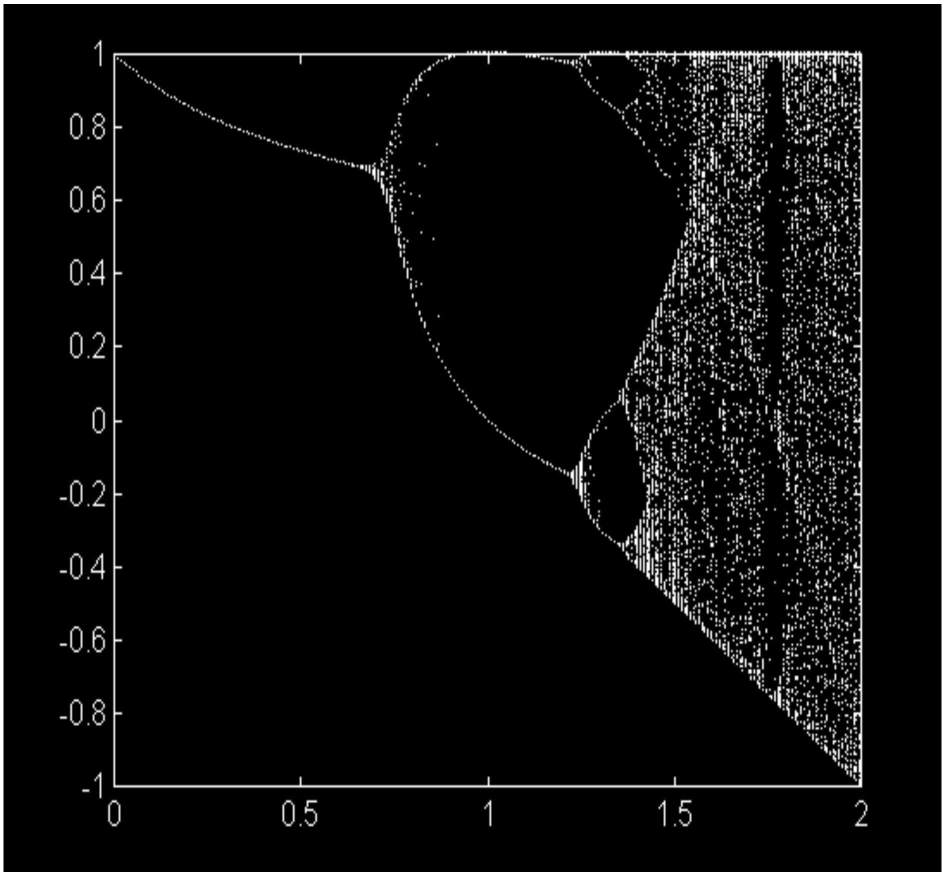


Fig. 16. Bifurcation tree showing asymptotic states of attractors as the parameter *b* is increased. 512 random initial points are chosen for each **b** value and the logistic map is iterated for 100 time steps before plotting. Where multiple state s points exist for the given *b*, a periodic, noisy (unstable) periodic or chaotic attractor is present; the actual state values are cycling between these y axis points as shown in the time series.

Depending on the value of the bifurcation control parameter *b,* the attractor state sequence may be a single state (fixed point), periodic oscillation between a few states (limit cycle), or a pseudo-random visitation of the state space points but within a bounded area (strange attractor, chaotic attractor). Each of the attractor types can be considered as a **phase** or **phase regime** of the dynamics, analogous to thermodynamic phase in classical physical systems. These phase regimes are bounded by **critical values** of the control parameters. When a control parameter is modulated to cross a point where attractors appear or disappear, and in particular change their **topological structure**, the crossing event is known as **bifurcation**. Bifurcations between qualitatively different regions of phase space, such as crossing the transition from limit cycles to chaotic behavior, are termed **phase transitions**.

Bifurcations occur as stability is lost for the attractor. This manifests in slow convergence time, and an increasing number of trajectories which lead away from the attractor. As the bifurcation parameter is increased into the chaotic regime, the unstable periodic orbits remain a controlling influence on the dynamics, effectively forming a "skeleton" for the dynamics.

The evolution or motion of a chaotic attractor in one dimension can be understood as cycling between sets of **unstable periodic orbits** (UPO); the emerging theory of control in low dimensional chaotic systems depends on analytically identifying such UPOs, and applying perturbations to the system to suppress chaos (Barreto, Kostelich et al. 1995). The persistence of such UPOs in chaotic behavior may have implications for the probability of reaching a particular state during the transient evolution of a system, or on the temporal statistics of a time series. This area of dynamics, particularly regarding transients, is not well characterized at the time of writing.

In non-biological systems, bifurcation parameters are typically constant or slowly changing with respect to the equations of motion. It is possible, and assumed by many researchers, that rapid bifurcation is a key aspect of the performance of biological systems.

SPACES, DIMENSIONS, MAPPINGS

At this point in the discussion, we must revisit the notions of space and dimension which have already been introduced, albeit in the context of cognitive theories of similarity as a space of features. Since I will return to that idea, but must use the term space in the dynamics context, the distinction should be made clear.

In the definition of a map given above, I emphasized the discrete nature of the process by using the term state. However, much of the theory of nonlinear dynamics – and more generally topology, of which it is a branch – is formulated in terms of continuous **spaces.** Indeed, the underlying space must be metric by the same criteria described earlier. When referring to the evolution of dynamic variables in $\Re^{n;}$ the dynamics literature normally uses **phase space**, indicating the space of the mapping dynamics.

Space is also encountered in the context of spatially extend systems or networks – here, it has essentially its commonplace meaning, with oscillating computational units or cells located in $\Re^1$ (a line or ring) or $\Re^2$ (a lattice or torus). Networks can also simply be defined on an arbitrary topology or graph, without reference to any embedding in real space.

The term **dimension** must also be revisited. In the context of metric spaces, dimensions typically refer to some measurement or feature, with objects represented as a point (or perhaps as a subspace) in the space. For hidden layer neural network representation spaces, the dimensions may be less directly related to the input; recall that the output units of each RBF classifier in the Chorus of Prototypes system correspond to a dimension in the space of prototypes. In Edelman's terminology, these can be distinguished as proximal (representation) and distal (feature) spaces. In dynamics and oscillatory neural network theory, dimension typically refers to the number of state variables in a system of coupled equations, such as the number of units in a **field** of identical connected units in a spatially extended system. There are additional **measured dimensions** of the orbit itself, characterizing the dynamics and information flow (Grassberger 1991).

We will also be interested in synchronization of oscillating units. In this case, there may be a **structural network dimension** (the number of units) and an **effective dimension**. When a network of coupled oscillators is fully synchronized, all units evolve in parallel and thus behave *effectively* as a one dimensional system.

In the system described here, all of these senses of dimension must come up at least briefly. In equations, I will follow the following notational convention: the variable N will be used for the network dimension, while the variable k will refer to the dimensionality of the representation space. Effective dimensions will be mentioned but will not require standard notation.

HIGHER DIMENSIONAL SYSTEMS: SPATIOTEMPORAL CHAOS

The logistic map introduced above is a typical discrete time nonlinear dynamical system with a single state variable. However, this formalism can be extended to networks of coupled nodes (cells, units) where each node has a real valued state. Such networks are known as **coupled maps.** When coupled maps are arranged in a regular spatial array, the term seen most often in current literature is coupled map lattice, introduced by Kaneko (Kaneko 1989). Other investigators have referred to similar structured spatial systems of nonlinear elements as cellular neural networks[20] (Chua and Yang 1988), fractal chaos networks (Perez and Massotte 1987), cellular dynamical systems (Abraham, Corliss et al. 1991). Due to the iterative or feedback network topology on each node, these systems may be considered as **recurrent neural networks with non-monotonic or bifurcating units**, and some investigators have described work in those terms (Farhat and del Moral Hernandez 1996). Coupled map models with local unit dynamics at the transition to chaos (known as the Feigenbaum accumulation point) have been designated as Feigenbaum networks (Carvalho, R. et al. 1999), and were investigated for utility in pattern recognition tasks by the present author under the name chaotic reaction diffusion networks (DeMaris 1995). Related systems with discrete state values, discrete time, and typically boolean mappings (transition functions) are known as *cellular automata*; many conceptual tools applicable to coupled map systems have been addressed in the cellular automata literature (Wolfram 1986). Some investigators have directly analyzed transformations between the two system types (Chate and Manneville 1989).

---

[20] Cellular neural networks (CNN) have been defined in a more general way to encompass both continuous and discrete time systems; the term is more likely to appear in engineering literature (especially circuit theory). For discrete CNN and CML I find no clear distinction apart from terminology.

For higher-dimensional and coupled map systems, the network state is a vector of the states of the constituent nodes. The network attractor, if one can be said to exist by virtue of sufficient coupling, is a **sequence** or **image** of this vector. For spatially structured dynamical networks, a spatial pattern formation behavior[21] at the level of the entire network is evident, emerging from the cooperative and competitive interactions between the patterns and dynamics in the coupled nodes.

This network-level pattern formation may be tuned by controlling the phase regime of the individual nodes, the number of connections between nodes (neighborhood size), the ratio of excitatory to inhibitory connections (for suitable activation functions), or the coupling strength between nodes.

The original papers on spatial pattern formation in locally coupled map lattices (Kaneko 1989); (Kaneko and Tsuda 1994) introduced many visualization techniques and correlation measures to characterize the rich behavior in various parameter regimes. In general, the long time behavior investigated by Kaneko is not applicable to the system described here, and as argued in the neuroscience review, is probably not applicable to rapidly developing perceptual processes. The present work investigates and uses only brief transients (10-16 iterations), while Kaneko's original simulations of locally coupled maps examine the dynamics after 10,000 steps, omitting any consideration of the transients.

The term attractor is sometimes used at the network level, but is generally less useful for spatial lattice systems with weak or local coupling, where oscillations and competition between **clusters** (oscillation modes, sites in the same state, attractor, or basin) form **dynamic patterns**. The **evolution** of a network from an initial state under relatively low coupling results in an organization in which patterns of continuing activity between interacting cells are spatially bounded by "frozen" areas, in which the neighborhood interactions reach a stable state. The local active areas are referred to as **domains**, while the frozen separating regions are domain boundaries. In a sense, the network organizes itself into sub-networks, with the activity pattern in a domain more conventionally related to the definition of an attractor.[22]

Systems with strong random coupling or global coupling, in contrast, can be shown to reach an attractor which may be equivalent to a one dimensional map; this **synchronization** process is taken up later. **Intermittency**, or **chaotic itinerancy**, is a phenomenon appearing in a small, weakly chaotic region of the parameter space, in which the dynamic behavior is a blending or linking of unstable periodic attractors existing in isolation in the more ordered regions of parameter space. Over time, individual periodic attractors are separated by sequences of intermittent chaotic transitions. This particular dynamics has been proposed by Tsuda as a supporting mechanism for binding (Tsuda 1992) and may offer advantages as an associative memory

---

[21] Kaneko uses the term spatial bifurcations for formation of separated islands or domains; I would prefer to reserve that term for dynamical scenarios in which spatial patterns actually influence local or global bifurcation parameters.

[22] The concept of a domain as used by Kaneko applies only when the lattice has been iterated for many (> 10,000) generations, so that transients have died out.

which overcomes the limitations of previous parallel distributed processing networks with respect to the issue of **compositionality** (the recovery of bound or integrated features in a composite memory).

Various network or neighborhood topologies have been reported in the literature on spatio-temporal chaotic systems. Network nodes may be **locally coupled** to adjacent nodes, **diffusively coupled** to a small region of the lattice with connection strength weighted by distance, **globally coupled** to every node, **randomly coupled** to a non-spatially localized set of neighbors, or some blending of these conditions. The coupling may take any functional form, including diffusive (multiplicative), difference or Laplacian coupling, or time varying functions. Coupling may be **symmetric** or **asymmetric** between adjacent units; it may be **homogeneous** over the spatial extent of a lattice, or **inhomogeneous**. The terms uniform and regular have also been used to denote homogeneous bifurcation and/or coupling.

Coupled Map Lattices

A coupled map lattice (CML) is a dynamical system with discrete time, an extended field of state variables in discrete space, and continuous state. Of course, we approximate continuous states with floating point values in map computations, so strictly speaking chaotic attractors must all in reality be periodic with very high period. Kaneko (Kaneko 1993) describes the generic CML modeling process for a physical system as follows:

1. Choose a set of field variables on a lattice. Typically these variables represent macroscopic (distributed) qualities, such as temperature, fluid velocity field, local concentration of a chemical substance, or in our case neuron pulse density in a local population.
2. Decompose a process into independent units, such as convection, reaction, diffusion, etc.
3. Replace each unit by the simplest possible parallel dynamics on a lattice, consisting of a transformation function at each lattice point or a coupling term among suitably chosen neighbors.
4. Carry out each process successively. In the present model, this means that at each iteration, a diffusion step is performed, then a reaction step.

The logistic map was originally chosen for its well understood properties rather than any explicit biological motivation[23] for that *particular* nonlinear function. The equation is known to be numerically stable when the state (and perturbations or forcings) are maintained within certain bounds, and the critical points where bifurcations occur are known. The behavior of logistic maps in a toroidal lattice with both local and global diffusive couplings has been extensively studied by Kaneko. These studies investigated

---

[23] However, the range of the asymmetric logistic map [-1 to +1] could be appropriate for modeling a process of *modulation about a background average frequency*. Some investigators of IT cortex present evidence that modulations of the background rate in individual units (Richmond et al.) or populations (Gochin et al.) predict the stimulus present. The zero value would be the background rate.

random initial conditions at each site. The pilot study leading to this dissertation (DeMaris 1995) was possibly the first to explore the dynamics and resulting distributions of a diffusively coupled logistic lattice *with structured spatial data* supplied as initial conditions.

The specific steps used for image processing in the system developed here are as follows. Note that the complete sequence of steps here is one iteration imbedded in one stage of the larger computation.

DIFFUSIVE COUPLING STEP

$$S_d(x,y) = (1-c)S_t(x,y) +$$

$$\frac{c}{4}\left[S_t(x,y+1) + S_t(x,y-1) + S_t(x+1,y) + S_t(x-1,y)\right]$$

where:d is the intermediate diffusion array, t is the current time step, x, y are the spatial indices of the pixel array S at the center of the diffusion neighborhood, S is the state variable at each site of the shape array, and c is the coupling constant restricted to the range <0.0 to 1.0>.

The diffusion or averaging step is implemented as an 2-D filter with the convolution kernel:

$$\begin{bmatrix} 0 & c/4 & 0 \\ c/4 & 0 & c/4 \\ 0 & c/4 & 0 \end{bmatrix}$$

where c is the coupling constant .

Diffusion is followed by a squashing step, which insures that the state remains bounded in the stable domain of the reaction step as described below:

$$S_t(x,y) = (1-c)S_t(x,y) + d_t(x,y)$$

LOGISTIC MAP STEP

The second computational unit applied in each time step is the logistic map:

$$S_{t+1}(x,y) = 1 - b(S_t(x,y))^2$$

where *S, t, x, and y* are as stated above and where:*b* is the bifurcation parameter, restricted to the open interval (0.0,2.0).S is restricted to the open interval (-1.0, 1.0).

The dynamics of the logistic equation are such that given any initial state, after some transient number of iterations the system will reach a steady state attractor which is fixed (low b), periodic with increasing cycle length and transient length, or chaotic (higher b values). As long as the initial input states S representing an image are bounded as described above the system will be numerically stable. In the periodic regime, the iterations required to converge to the attractor vary from 1 to 100 or more; longer convergence times are observed in the vicinity of the critical (bifurcation) points. This phenomena can be seen in diagrams in the next chapter illustrating the evolutions of transients as the bifurcation parameter b is scaled.

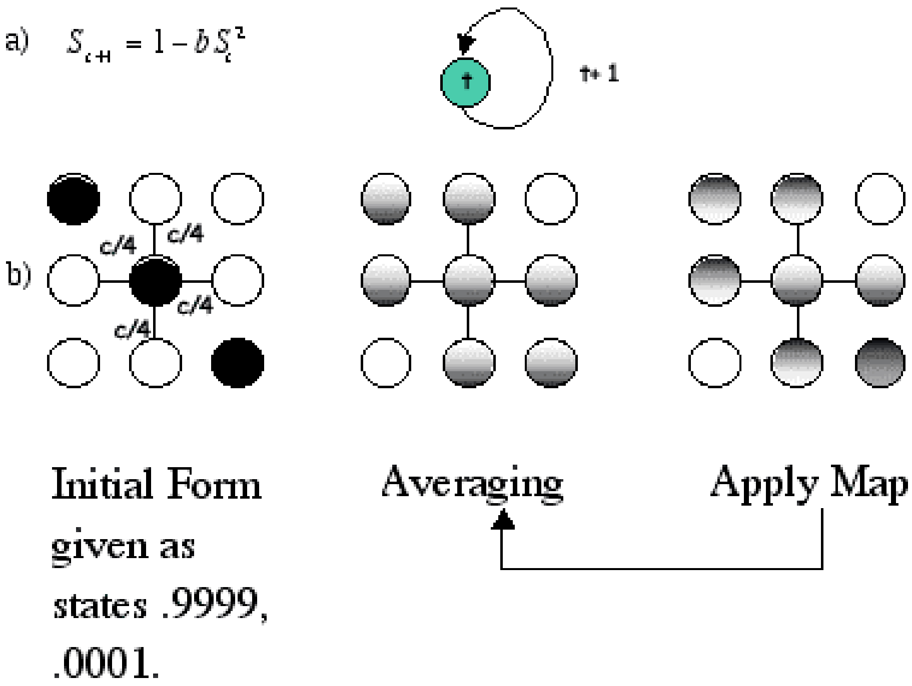


Fig. 17. A pictorial illustration of the CML architecture. The logistic equation is executed in each node of the CML. Each node implements an *iterative* or *recurrent* computation, illustrated by the loop on the node illustrated. b) A regular spatial lattice of such nodes is connected to nearest neighbors, with the NSEW edges labeled averaging a scaled fraction of the state value of their neighbors at each iteration of the lattice. The illustration is intended to show how a contour of black “zeros” in a background of white “ones” is diffused by the averaging process, then transformed by the map. The sequential application of both operations is one iteration of the coupled map process.

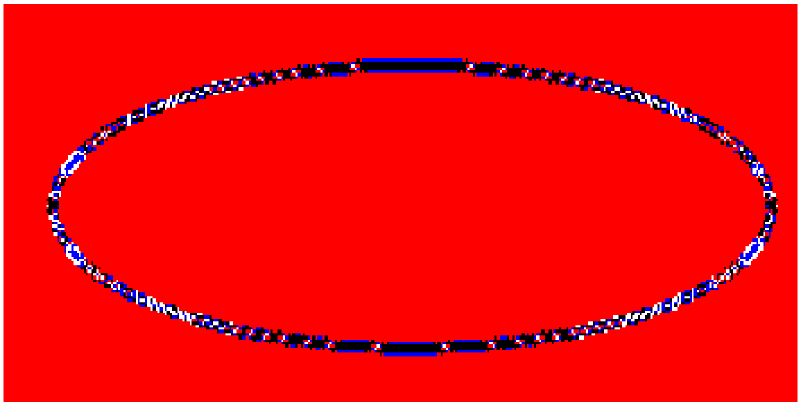


Fig. 18. A snapshot in time of an evolving coupled map lattice. An ellipse form with values .0001 on a background field of .9999 given as an initial state is evolved through the diffusively coupled CML process described above; a cyclic color map highlights distinctions between the state values in the interval [-1,1]. The thickness of the ellipse indicates the propagation of the diffusion wavefront, while the local spatial structures show the effects of coupling and map operators to produce characteristic local statistics for a particular curvature region.

Synchronization in Coupled Map Lattices

Several types of synchronization may occur in CML systems, given the different phase regimes that the individual component maps and collective can operate in. If individual units reach fixed points, the only possibility is totally synchronized, possibly in different domains or clusters.

If individual units exhibit periodic oscillation, due to their local bifurcation parameters or strong coupling, they may exhibit phase clustering, with P clusters corresponding to the possible phase offsets in limit cycles of length P relative to time t modulo P. If all units are in the same phase, they are defined as totally synchronized.

If individual units exhibit more complex (chaotic or itinerate) behavior, they may be still be totally synchronized. Apart from such total synchronization, a variety of synchronization types and clustering identification techniques are beginning to be noted by investigators.

Partial synchronization is defined (Maistrenko, Popovych et al. 2000) in contrast to total synchronization. For a set of units coupled in some graph, total synchronization is defined as the case in which

$\left|s_t^i - s_t^j\right| \to 0, t \to \infty$The system of N units (N dimensional network) effectively operates as a one dimensional map when synchronized.

Partial synchronization, then is the situation where some units obey the synchronization condition above while others do not. The network may be characterized by clusters C1, C2, ...Cn < N, where N is the number of units. These may be spatially segregated into contiguous oscillating *domains*.

For Maistrenko and colleagues, the emphasis is on clustering of units which reach an asymptotic state sequence which falls short of full synchronization. In contrast, another type of partial synchronization will be the focus of the rest of this thesis. Over the time course of a dynamical evolution from an inhomogeneous initial state towards synchronization, in either the full or partial sense of Maistrenko, we can consider the distribution during the transient state as exhibiting partial synchronization.

If such partial synchronization is measured by sampling the dynamics (or simply halting them, if such control is available), the differing rates of convergence of local configurations to the asymptotic distribution can be used in information processing.

Clustering Phenomena in Globally Coupled Chaotic Maps

The dynamics of a coupled, spatially extended system of maps can no longer be visualized as a bifurcation tree (the pitchfork diagram show above); instead, a phase regime plane is used, where each point corresponds to a family of attractors of the same type or mixtures of attractor types, and bounded regions correspond to phase regimes in the entire network state space. In the more complex regimes described later, mixtures of the simple attractor types (such as cycles and intermittency) may co-exist in physical space, or move in the physical space of the network as traveling waves.

A schematic plot of the control space for a network of **globally coupled** logistic map nodes, where each node is coupled to a mean field or average of all nodes at each time step, is shown below. The axes of the plot correspond to local bifurcation parameters, and coupling parameters between nodes. While the boundaries between the network phase regimes are simple in this depiction, they can normally be very complex and intermingled even for single map units, (e.g., the appearance of a period 3 window surrounded by chaos can be seen in the bifurcation tree).

The structure of bifurcations is known as the **route to chaos** of a chaotic nonlinear system. The individual logistic map at each site, with no coupling, cycles through a **period doubling limit cycle cascade,** reaching the chaotic regimes at a critical value *b*. The addition of spatial arrangemetnts and coupling to the low dimensional dynamics picture complicates the description of dynamical structure. The intermediate regimes for locally coupled maps are considered to have "frozen random" pattern selection behavior in which **domain boundaries** form. Higher couplings produce larger domains and ultimately a pattern formation behavior.

The globally coupled map has a toroidal collapse route to chaos. In this route, the boundaries between the network phase regimes are monotonic with respect to the control or parameter space. In globally coupled maps, domains are more unstable and clustering is the dominant phenomenon.

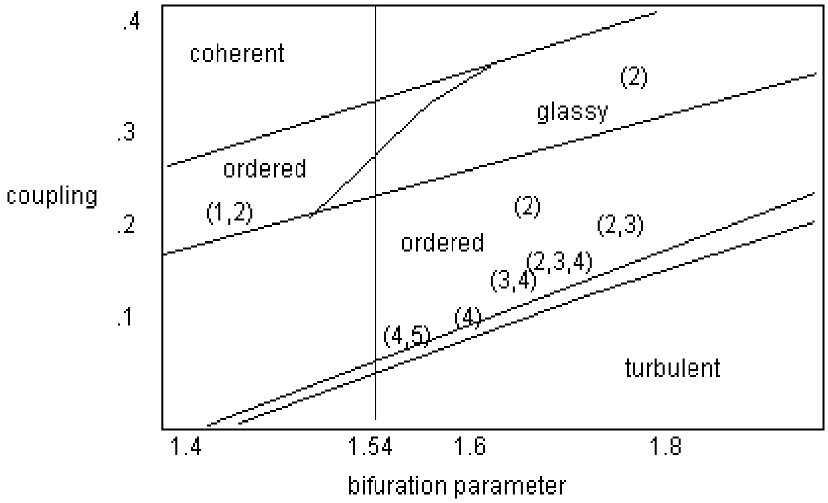


Fig. 19. The control plane and resulting *phase regimes* for a globally coupled map lattice, after (Kaneko 1990). The line at 1.54 is the critical point (transition to chaos) for an uncoupled map. The coherent or synchronized regime indicates that all sites exhibit the same orbit. The numbers in the ordered and glassy regimes refer to the number of dominant clusters, defined as those with basin volume more than 10%. A cluster is a set of lattice sites oscillating in the same attractor, though not necessarily with the same phase. Note that the even when individual sites would be chaotic, strong coupling can enforce coherence and complex behavior.

THE SOCA DESYNCHRONIZATION-SYNCHRONIZATION CYCLE: A TIME-VARYING CML

Having assembled the conceptual tools underlying the network at the heart of the hybrid pattern recognition system, I now describe the small extension of the classical coupled map lattice. This extension, motivated by the experimental observations of changing slow wave potentials and correlations (Bressler 1995), is shown in the next chapter to increase recognition performance (both the recognition rate and average recognition time) achieved in bounded iteration counts for a recurrent network. This key extension is to use time-varying parameters, alternating between opposing epochs of desychronization and synchronization.

Once again, the reader's attention must be called on to note a subtle shift in th meaning of synchronization, as it is used in a graph theoretical and statistical perspective on dynamics; this will be explained in more detail in the following section. In this graph theoretical perspective, desynchronization (broadening of the distributions) by loosely coupled chaotic dynamics is followed by partial synchronization with more strongly coupled chaotic dynamics.

The two meanings of synchronization are closely related in a high dimensional system. In my use of coupled maps for shape representation, the initial conditions are confined to a small subspace of the phase space; this is synchronization in the new sense. The chaotic dynamics, even in a strongly coupled system, will desynchronize in the sense that for some transient period a range of time series will emerge through the neighboorhood actions, whose correlations depend on the initial conditions but also diverge in a pattern dependent fashion. When time series synchronization occurs – either through temporary fluctuations in the transients, or through the gradual formation of synchronized domain strutures – we will see the concentration of the system state in a subspace again. This concentration may be apparent in instantaneous measurement of the system state, or by examining time averaged occupancies of subspaces.

These operations take place in a homogeneous, orientation-sensitive array of recurrent logistics maps. By homogeneous, I mean that the bifurcation and coupling parameters are uniform across the entire array. By orientation-sensitive, I refer to the spatial asymmetry in the coupling kernel (i.e. NSEW neighbors coupled, diagonal neighbors uncoupled).

Expressed in algorithmic form, the procedure is:

```
procedure synchronizationOpponentNetwork
 image = threshold(downSample(readImage)))
 for iterations = 1 to t1
    diffuseImage = filter2D(couplingMatrix,image)
    image = logisticMap(diffuseImage,b1,c1);
 end // Desynchronize stage
 for iterations = 1 to t2
     diffuseImage = filter2D(CouplingMatrix,image)
     image = logisticMap(diffuseImage,b2,c2);
 end // Partial Synchronize stage
end  // procedure
```

Because of the variation in divergence and convergence times, a specific set of bifurcation, coupling, and iteration time parameters {b1, c1, t1, b2, c2, t2} has a characteristic response to any given image or family of images. Each image can be considered as a set of overlapping initial configurations of size t1+t2; by the end of the Soca process above, information about local configurations from a window of size t1+t2 is contained in each unit (pixel in the processing array). The set of initial configurations comprising one image may be highly synchronizing for those parameters, while another image may be less so less so.

The intuition behind the network operation is that images in some category are considered as *productions* of a *stochastic language* on an alphabet α whose symbols are local pixel configurations. We seek parameters for the first (desynchronizing) stage which, for this language, have the appropriate divergence rate matching parameter determined characteristics of the second (synchronizing) stage. The second stage must have characteristically avoided regions of state space and state transtitions such that images in the category will *converge near a characteristic sparse distribution*.

Similar images should result in similar output distributions, and the inherent characteristics of coupled chaotic systems – divergence of nearby states with time under low coupling, but convergence to synchronized or partially synchronized states with high coupling – offer a potential computational framework.

This computational framework can be viewed in terms of dynamical recognizers reviewed earlier. Alternatively, it can be described as a generalization of a problem in graph theory known as the *road problem;* it also has similarities to Markov chains and probabilistic finite state automata. These descriptive frameworks are briefly described in the next chapter on representation and learning.

This interaction between the specific initial configurations on a shape boundary and the dynamics is an example of *cooperative processing*. Cooperative phenomena, and particularly pattern formation processes, are distinguished by Palm (Palm 1982) by the following criteria:

1. The phenomena arises from the interaction of a large number of similar components.
2. The laws governing the interaction of the components are local.
3. The dynamical laws are the same for all components.
4. The local dynamics should not contain the "symmetries" of the global pattern evolving through the local interactions. It is usually difficult to predict what pattern will arise from the local interactions.

The essence of the activity flow in the present network is cooperative processing mediated by the *opposed processes* of desynchronization and synchronization, hence it is designated as *synchronization opponent cooperative activity*, or **Soca network**[24].

The table below indicates the particular network parameters investigated here, selected from an expanded *network parameter space* of possible single layer CML systems.

Table 3. Network Design Choices for the Soca System

	Input	Coupling	Bifurcation	Readout
Structure	All cells	Local Diffusive	Logistic	partition cell occupancy of all states in diffusion wavefront,
Spatial	Homogeneous	Homogeneous	Homogeneous	All states
Temporal	Initial condition (one shot)	Variable "opponent" stages	Variable (opponent stages)	Instantaneous

The entries in the table above represent parameters of network design using CML derived systems. The set of choices shown are *fixed constraints*, within which an evolutionary search proceeds to discover solutions to an object recognition problem. A more ambitious evolutionary search could choose to optimize networks choosing different alternatives for some or all of those parameters.

MACROSTATE VARIABLES AND MEDIUM SCALE NETWORK MODELS

Before proceeding to experimental methods and data, I now return to the general topic of modeling biological networks with oscillators and synchronization phenomena. The following sections serve two goals. The first is to highlight the history and justification of such mathematical systems as neural models. Second, I mean to survey

---

[24] The name is inspired in part by the admonition of Walter Freeman that neuroscientists need to learn to dance.

closely related work in networks oriented toward some specific perceptual modeling or pattern recognition function.

As noted earlier, connectionist models have typically assumed rate coding and monotonic activation functions, which correspond closely to the common assumptions of transfer functions for single neurons, and to assumptions about multiplicative weighting as the essential operator controlling network information flow. Some investigators are more cautious, noting that units in connectionist networks may correspond to larger structures. For example, questions raised regarding the reliability of single neurons for rate coding suggest that a sigmoidal activation function unit might be better interpreted as the average behavior of many parallel detectors converging on a family of readout neurons (Softky and Koch 1994).

An alternative to modeling single neurons or putative "parallel averaged networks" is to more directly model large scale dynamics, consisting of thousands of neurons. This has historically been the domain of *statistical mechanical* (Amari 1974) and oscillatory models. One of the first such oscillatory models still in relatively wide use [25] was developed by Wilson and Cowan (Wilson and Cowan 1972); (Cowan 1974). Continuous variables represent activity levels of excitatory and inhibitory sub-populations, rather than activity of single neurons.

There is one crucial difference between the Wilson-Cowan (WC) model and the use of maps to model oscillatory brain dynamics; while both maps and the WC model can generate chaotic time series and be coupled in spatial aggregates, the WC model also has "resting", non-excited states. Since a map with chaotic control parameters is chaotic for any input, there is no equivalent rest state. It would be possible to introduce additional nonstationarities in the model, with a baseline fixed point attractors (and corresponding bifurcation parameter) designated as a rest state. Input shifts the bifurcation parameter to leave the fixed point state, perhaps with some decay to the resting state. In the modeling here, I address this by simply ignoring "background" states beyond a diffusion wavefront region of interest in the evolving pattern by omitting the highest bin count when gathering the statistics. Given the current algorithmic "back end" recognition process, the use of histogram functions intended for image processing would require similar suppression of the rest state values even if a more complex input coupling and evolution dynamic were used, without really contributing to the essence of the project.

When modeling physical or psychological phenomena with spatially extended (field) dynamics, it is common for each variable in a field (i.e. each unit in a coupled map lattice) to represent a quantity associated with an aggregate of microscopic units. This kind of representation, originating in statistical mechanics or fluid dynamics, is known as a macrostate variable. Temperature or instantaneous velocity of a fluid, for example, are macrostate variables in the study of fluids. In neural modeling, the macrostate variables are quantities like ensemble activation (average spike train frequency of all units), temporal phase distribution, or ensemble average frequency (pulse train density or spikes / unit time measured over the whole ensemble). Parameters in

---

[25] Cowan cites earlier work by Buerle (1957) and Griffith (1963) as historically important in this research stream; I have not encountered recent work citing them.

equations may also be considered as macrostate parameters, indirectly capturing quantitative effects of distributed aggregates like neurotransmitter fluxes, and excitatory-inhibitory ratios of neural sub-populations, or distributions of delay times which induce bifurcations in large networks.

Earlier the distinction was made between state variables and control parameters, with the latter consisting of bifurcation and coupling parameters. In large scale models, control parameters - especially bifurcation variables – generally will not have a direct map to single parameters in the underlying micro-circuit model. In fact, bifurcation variables can be interpreted as subsuming the effects of widely separated neural tissue, such as cortical regions and sub-cortical nuclei, or long range connections between cortical regions.

As noted earlier, the trend in interpretation of EEG signals and evoked potential responses is as a signature of large scale coordination and control processes between the regions that brain imaging indicates cooperatively produce computations. The sudden step function change in network parameters above should be interpreted as an example of such large scale control, where staged or periodic volleys from cooperating cortical or subcortical regions effect this rapid change in bifurcation and coupling parameters.

The systems described here should be regarded as spatiotemporal cooperative systems acting on vectors in **pulse density** space, with the step-function changes in bifurcation and coupling parameters representing slower control dynamics implemented by modulation from separate sub-populations. The control dynamics of bifurcation and coupling at the population level supplement the traditional neuron level control dynamics of gating, inhibition, and feature selection. These population level control dynamics may be more easily correlated with MEG and EEG observables than the traditional control dynamics acting at more local scales. In turn, they may produce neuron level observables such as modulations in correlations of neurons engaged in a processing task. I will return to this subject in the final discussion section.

The next section addresses the question of how the use of chaotic maps is justified in terms of standard neuron models, without addressing questions of learning.

SYNCHRONIZATION PHENOMENA AT MULTIPLE SCALES

If a map is to be regarded as representing the aggregate behavior of a large system of neurons, it becomes clear that we must explain oscillatory and synchronization phenomena operating at multiple scales of the brain, from micro-circuits of a few neurons to large networks. This task has been addressed by several investigators. In one such study, Wennekers and Pasemann investigate coupled pools of sigmoidal activation neurons with random diffusive coupling and a probability distribution of coupling strengths. They found that for appropriate parameters the temporal behavior of the whole system can be described by a single, low-dimensional equation (Wennekers and Pasemann 1996). This is proven true asymptotically for long times and system sizes approaching infinity. Even for fairly small networks (N=50 for two interconnected pools) with 50% standard deviation in coupling strengths, a similar bifurcation structure between the low dimensional system and the full network (average activation of all micro-circuit neurons) is seen. The network bifurcation structure has a period doubling cascade leading to chaos, similar to the logistic map used here, but with the possibility of multiple coexisting attractors for any given coupling between pools. Their work situates itself as an extension of seminal work by Anninos, Harth and colleagues (Anninos, Beek et al. 1970); (Harth, Csermely et al. 1970), and Palm (Palm 1982), which demonstrated that *threshold modulation* in random networks leads to mean activity input-output curves described by a single humped function.

In a series of studies, these earlier workers performed discrete time simulations of random networks of threshold neurons of mixed excitatory and inhibitory types. Parametric curves (reproduced below) of activity levels α resulted from studies of the networks which varied the parameters shown on the following page along with a typical activity transfer function.

Network parameters affecting activity curve

h percentage of inhibitory cells
μ^+ average number of synapses of an excitatory cell
μ^- average number of synapses of an inhibitory cell
η threshold of all neurons (number of excitatory synapses to fire)

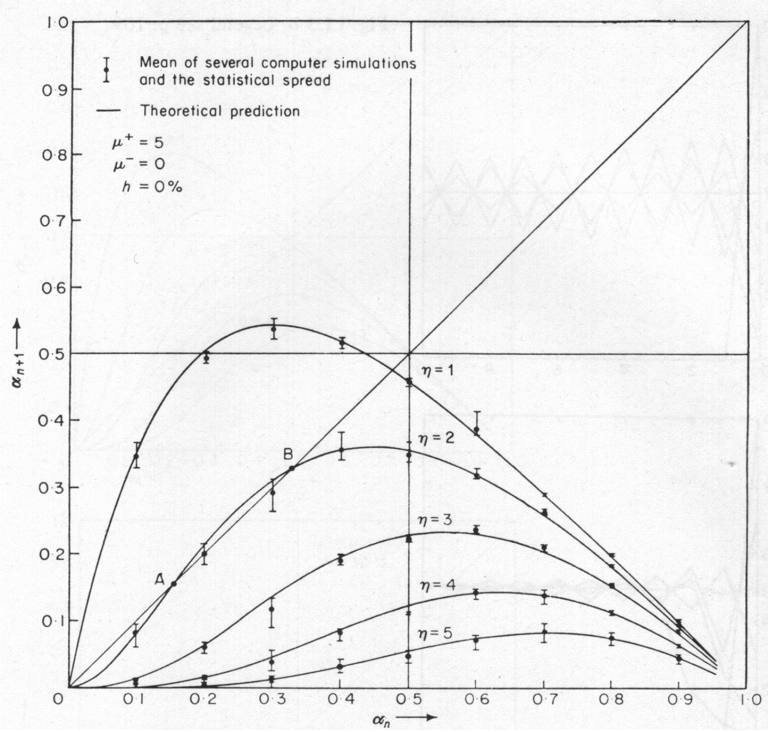


Fig. 20. Activity at time step t+1 vs. t. Note that the form of these curves, are single humped maps – the same input output structure governing the logistic map. This illustrates one scenario to produce the input-output form, characteristic time series and bifurcation structures by micro-circuit models. Changing excitatory-inhibitory ratios and the firing threshold act as the bifurcation parameters of this model. From Anninos, P. A., B. Beek, et al. (1970). "Dynamics of Neural Structures." Journal of Theoretical Biology 26: 121-148. Reproduced with permission of Academic Press.

BIOLOGICALLY PLAUSIBLE COUPLING VALUES

A question that should be raised when thinking about CML systems as models of medium or large scale neural dynamics is what, if any restrictions on the range of coupling parameters should be considered as biologically realistic?

One position would be to consider anatomical data on the proportion of lateral connections to recurrent connections in mini-columns as a bound, resulting in a smaller upper bound (perhaps .03) than is typically found by evolutionary learning in my experimental work. I would argue that such a restriction is unjustified, because the underlying micro-circuit dynamics which regulate synchronization may not require large numbers of independent connections, instead relying on more subtle mechanisms. For example, spike doublet firing mechanisms have been implicated in long range synchronization (Traub, Whittington et al. 1997). Thus the coupling strength should be understood in terms of ratios along the continuum achievable by any such synchronizing mechanisms. hese will of course depend on a non-zero anatomical connection density, but need not bear any linear relationship with the lateral connection density.

REVIEW OF ANALYTIC RESULTS ON SYNCHRONIZATION AND ENSEMBLE DENSITIES OF MAPS

The previous sections have covered the basic definitions needed as a basis for understanding the dynamical network function for representation forming representations and for pattern recognition. I will now briefly describe recent theoretical developments which hold promise for future work on more direct computations of required parameters, or for establishing bounds on the applicability of the technique. The algebraic treatment of graphs, known as spectral methods, has been extensively developed since origninating in 1972 (Donath and Hoffman 1972) for graph partitioning, with application to integrated circuit design. Building on these techniques, Wu has established several theorems on the lower bounds of coupling required for full synchronization on various topologies of coupled map lattices (Wu 1998).

These results apply when the bifurcation and coupling parameters are homogeneous or uniform. Three cases are treated.

For globally coupled logistic maps, coupling to the mean field exceeding a threshold $|1-c| < \frac{1}{b}$ guarantees synchronization.

For coupled maps in which each of N units is symmetrically coupled to k neighbors with coupling $\frac{1}{k}$, the system almost always synchronizes for large k.

For coupled maps with uniform symmetric coupling on a connected graph of n nodes, a region in the (b, c) parameter space results in full synchronization. The region is bounded by the constraints

$|1-cm| < \frac{1}{\sup_x \|Df\|}$, $|1-c\alpha| < \frac{1}{\sup_x \|Df\|}$ where α is the algebraic connectivity of the graph (smallest nonzero eigenvalue of the graph Laplacian matrix, Df is the

Jacobian matrix of the map function *f* maximum divergence over nearby points in the map domain, and m = min(n, $2\mathrm{deg}_{\mathrm{max}}$) for the graph.

Another result due to Wu is that for uniform discrete time maps with homogeneous *additive* coupling (Wu 1999). For the case in which the connection topology matrix has zero row sums, the dynamics of the coupled system after synchronization is proven equivalent to some parameterization of the *uncoupled* map. This result may have implications for efficient encodings which are generated from a high dimensional system, suggesting that an encoding or matching process could generate the effective dynamics with a one dimensional system. The proof does not hold however, for diffusively coupled maps used in this thesis.

Synchronization per se is not a goal for the following description of a pattern recognition system; on the contrary, in order to form a representation space, complete synchronization must be avoided or the representation space becomes effectively one dimensional, with little separation among object representations likely to occur even if sampled prior to complete synchronization. One possibly use for these theorems would be to set constraints for any learning or search process, since coupling values guaranteed to synchronize may be unsuitable for the purposes of forming representations, if it occurs rapidly.

THE TIME COURSE OF EVOLVING DISTRIBUTIONS IN ENSEMBLES

As noted, most work in dynamical systems focuses on the long time equilibrium behavior of systems, in contrast to their transient behavior. One recent exception is work on the evolution of densities (distributions) in ensembles of identical maps (Driebe 1999).

Driebe’s work takes the statistical mechanical viewpoint, studying the distribution of states for an ensemble of identical maps over a distribution of initial conditions. While the density concepts and the methods developed are also applicable to equilibrium states, the work is notable in its emphasis on distributions during the transient evolution toward the limiting equilibrium distribution.

Densities may be measured instantaneously over an ensemble, or over time. Given the latter emphasis, complex motion (orbits) in low-dimensional chaotic systems is naturally described by densities. Density means the occupancy of a region of phase space, i.e. the fraction of the ensemble in a particular subset of the domain of the function. Initial nonequilibrium density represents an ensemble of uncoupled maps with different initial conditions (or perhaps just uncertainty about the initial conditions).

Behavior of orbits in a map and densities in an ensemble of maps may be strikingly different. A typical orbit in a chaotic system looks qualitatively similar forward and backward in time, even if the dynamics is non-invertible. In contrast evolution of density is usually obviously time oriented. Thus density evolution is not reducible to trajectories.

For non-invertible systems with chaotic trajectories, evolution of densities will show regular behavior; for systems with regular (i.e. periodic) trajectories, the density will mirror the orbit level. For chaotic maps, nearby trajectories diverge, while initially different densities, over an ensemble of identical maps converge. Instantaneous densities of low dimensional systems may rapidly (e.g. only a few iterations) approach the equilibrium density.

The evolution of densities is described by the Frobenius-Perron operator. The spectral decomposition of this operator can be used to compute decay rates of correlation modes (peaks in the Fourier transform corresponding to poles in the complex frequency plane). The *spectrum* of an operator is the set (discrete or continuous) of eigenvalues of the operator under consideration acting in a specified functional space, and is different for different functional spaces. For uncoupled maps the distribution of a random ensemble can be computed exactly at each time step.

For the purposes of the discussion here, the main message is the *rapid evolution of distributions* toward an equilibrium value for chaotic dynamics; however, it is not entirely clear if this rapid evolution applies to *coupled* map systems with highly structured distributions. The relatively small number of lattice sites, strong peaks in the initial distribution, and the constant perturbation from equilibrium states due to coupling clearly result in stronger peaks in the transient distribution. If the density approach can be extended to coupled maps (perhaps building on Wu's approach to produce an equivalent single map) it may be possible to more directly construct a classifier for a particular initial distribution characterizing some family of patterns to be recognized. For now, adaptive learning methods, such as I use in this thesis, seem to be the only practical approach.

SPATIALLY EXTENDED DYNAMICS, TRANSIENTS, AND SYNCHRONIZATION: NOTES ON THE LITERATURE

In this section I will survey additional literature exploring nonlinear oscillation dynamics in vision, pattern recognition and other engineering tasks, but which is somewhat tangential to the main thread of shape representation and similarity.

Conceptual Ties with Cellular Automata Literature

The fields of cellular automata (CA) and random boolean networks were the better established "parent" disciplines which spawned work on coupled maps (Wolfram 1986). Both share discrete time and space iteration, with most work employing synchronous update at all cells. CA and coupled maps systems differ only in that CA typically have boolean or small integer state variables, with boolean transition functions. Coupled maps use on or more real-valued numbers as state values, and use algebraic or piecewise-linear functions as update rules. *Transients* in cellular automata have been studied more extensively than in continuous dynamical systems or coupled maps; methods for creating appropriate structure of the attractor basins (i.e. the transients leading to an attractor) have been derived by Wuensche, but these computations have no obvious mapping to biological dynamics (Wuensche 1996).

The concept of *time-varying spatially extended dynamics* was proposed by Wolfram in the context of cellular automata (Wolfram 1986). A slow lattice controls the rules governing the update of sites on a fast lattice. Wolfram informally describes several strategies for pattern classification with such systems, covering some of the same ground described earlier by Rosenfeld (Rosenfeld 1979).

Pattern formation phenomena are the major object of study in CA, but are typically not considered as synchronization; however, results (reviewed in an earlier section) on problem solving strategies involving regular domains (Hordijk, Crutchfield et al. 1998), (Mitchell, Crutchfield et al. 1996) may be possible to recast in terms of synchronization. The essential difference is that *n-blocks* (words or spatial configurations) in CA play the role that *discrete phase space intervals in a single cell* play in coupled map lattices. Measures on block statistics and correlations replace measures on occupancy of phase space regions. Cooperative pattern formation processes are the essential characteristic in both systems.

Pattern Processing in Coupled Maps

Work on the processing of spatial patterns by arrays of chaotic units has been performed by Farhat and del Moral Hernandez (Farhat and del Moral Hernandez 1996). The standard symmetrical logistic map formulation is used for units, with the state variable interpreted as phase in the interval $[0,2\pi]$. They interpret the map as a model of spike processing in a single neuron, in contrast to a large scale network as in most other work reviewed in this section. One notable aspect of this work is that the coupling function between cells is itself nonlinear; two variants of coupling are proposed. One is an exponential function of input, the other a series quantization thresholds against this exponential function. Quantization (binning) results in a loss of smoothness in the characteristic pitchfork bifurcation diagram of the coupled maps, producing instead constant values until bifurcation points. The quantization is interpreted as different neurotransmitter release characteristics, associated with different presynaptic activation levels.

In their demonstration of pattern processing, piecewise linear activation values are applied as bifurcation parameters to an input logistic ring, which is coupled to a second processing layer via the nonlinear scheme above. Coupling between elements is homogenous. It is shown that after long transients (1700 cycles in one example) for some input patterns the dynamics may collapse to clusters of periodic attractors. The number of clusters is much smaller than the array, i.e. 6-7 clusters in an ring of 100 chaotic units. It is suggested that this convergence to clusters of periodic attractors for “coherent input” may be interpreted as recognition and classification of the input, while inputs which are not recognized remain incoherent.

As noted in my review, research in IT cortex, the putative site of object level feature recognition, has not turned up obvious periodic oscillatory dynamics fitting this hypothesis at the single neuron level, but neurons could be participating in larger scale aperiodic oscillatory dynamics. However, the large number of iterations required seems inconsistent with rapid processing. The correspondence of an iteration cycle with particular micro-circuit parameters is not developed in the paper, but even if it corresponds to recurrent processes in dendritic spike processing 1700 iterations seems a heavy burden to justify biologically.

A fully connected network of quadratic maps with period doubling route to chaos has been studied by Carvalho and coworkers (Carvalho, R. et al. 1999). A baseline bifurcation state at the critical transition to chaos is chosen, with a correlation learning rule based on the response to input. They note that uncoupled dynamics at this transition, while not chaotic in the + Lyapunov sense, consists of an infinite number of unstable periodic orbits; when coupling is induced between units, some of these are stabilized, resulting in a characteristic distribution over a set of high period orbits for the learned input pattern. This stabililization is observed for a fixed initial condition prior to presentation of patterns and is measured after 2×10^4 time steps.

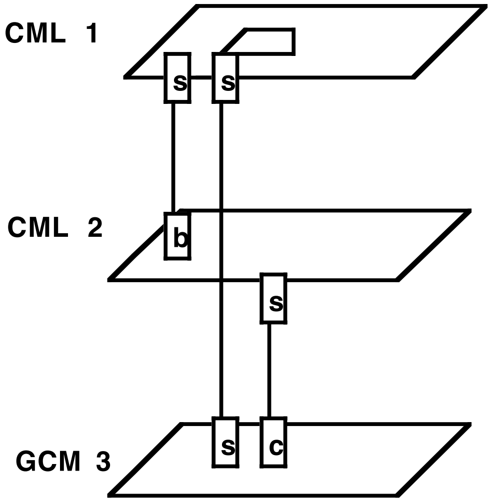
Robert Gregson has, over the course of many years, explored nonlinear models of psychophysical phenomenon and collected the results in two monographs (Gregson 1988; Gregson 1995). Some of this work has utilized spatially extended or *field* models, designated as (nxn) Γ models. These are notable in the present context because many of the studies also deal with low numbers of iterations, thus are essentially exploring the computational correspondence of transient phenomena with psychophysical events. Also, Gregson introduces the notion of *cascades*, a set of n recursions in a lattice; the output of this system is fed back to the input for an "outer loop" of some number of iterations. In this model, the initial "stimulus" variables are not the state variables but rather gain values affecting the evolution of an autonomous complex variable; in the outer loop, the output of one such cascade is used to control the gain in a subsequent cascade. Gain in the n-Γ system serves as a bifurcation parameter, so the system as a whole is non-stationary and effectively *auto-bifurcating* in my own terminology introduced in a previous thesis (DeMaris 1995). That specific kinds of computations are effected by bifurcation changes *on a slow scale relative to evolution equations* is a major commonality with the present model [26]. Gregson has modeled spatial vision phenomenon such as the Muller-Lyer illusion, using total iteration counts under 100. He makes many points which I arrived at independently; that nonlinear evolution equations are a "total system analogue", rather than corresponding to any local (retinal or cortical) neural sheet. Also, in contrast to earlier field theories (Ratliff 1965), there is no reliance on opposed excitatory and inhibitory influences. Like the network dynamics explored here, Gregson notes that coupling connections between these nonlinear field units have no obvious interpretation as excitatory or inhibitory.

The phenomena of ambiguous perceptions has been of great interest since the earliest days of visual psychology. The spontaneous switching of the images such as the Necker cube is clearly a dynamical phenomena, and the apparent instabilities might be expected to shed light on perceptual processes. The literature on Necker cube psychophysics details interactions between eye movements and switching events, as well as interactions between scale, orientation, and the distribution of switching times. In an

---

[26] Because of this emphasis on transients, Gregson's approach must be acknowledged as a key precursor of my work, though the original impulses for my investigations came from other work outlined here; due to terminology differences I only realized the similarity of CML with his "cascades and fields" approach after personal communication with T. Henmi, comparing our respective work on Muller-Lyer illusions and attempting to combine aspects of both.

earlier study motivated by these complexities, I modeled the cooperative formation of monocular depth fields and attentional foci using a multi-layer CML model, mixing two locally coupled lattices with a globally coupled lattice. Scale changes in the cube led to changes Scale changes in the cube led to changes in the distribution shape conforming to psychophysical trends. A low dimensional dynamical model by Kelso et al. previously linked interactions between reversal rates and distribution shapes, but did not provide details of interactions between spatial forms and coupling, or account for attentional correlates (Kelso, Case et al. 1995).



Input layer

Primal sketch figure is injected to a local diffusively coupled CML. The state values at this site drive 2 other layers.

Attention layer

Form mediated fluctuations between chaotic and periodic oscillations generate spatially localized attention events and control organization of depth field layer by coupling to depth layer.

Gestalt (depth field) layer

Coupling modified during attention events, influencing spatial distribution of clusters.

Fig. 21. Schematic of a network modeling formation of monocular depth fields with multiple CML layers. The labeling on the states indicate the nature of couplings between layers; i.e. state fluctations in layer 1 influence bifurcation parameter in layer 2. From DeMaris, D. (1998). Pattern formation in spatially extended nonlinear systems: toward a foundation for meaning in symbolic forms. !st. Intl. Conf. on Anticipatory Systems, Liege, American Institute of Physics. Reproduced with permission of Lawrence Erlbaum Associates.

Segmentation and Synchronization in Coupled Oscillator and Coupled Map Systems

Coupled oscillator and coupled map systems have by now been investigated as models of several perceptual tasks and phenomena. Segmentation, the task of identifying object boundaries and separating a scene into object tagged regions, has received the most attention. A few of these studies are relevant because of their intrinsic interest in the context of computing with synchronization processes, and also in that practical work in **computerized scene analysis** demands that several tasks be solved nearly simultaneously, such as the discrimination of object boundaries and the recognition of the objects demarcated. Successful segregation of objects is assumed in the computational work here, but in future engineering systems or more complete biological models these tasks may be combined; happily it appears that the synchronization framework, and chaotic synchronization in particular can support both tasks.

Peréz applied self-organizing responses at the local cell level in an inhomogeneous coupled map lattice model, using the evolution statistics of a "spin" order parameter as a discriminator of irregular structures in a regular background in a semiconductor defect recognition task (Perez 1988). By self organizing, I mean that the bifurcation parameter at each time step is a function of the state of each cell; this is further regulated by a local spin-correlation measure, which is essentially a measure of synchrony. Spin is defined as a difference measure from one time step to the next; positive increments are spin up, negative spin down. Thus neighboring cells with similar time series derivatives are stabilized, leading to segmented spatial activity domains corresponding to manufactured shapes regions and defects.

Price et al. (Price, Wambacq et al. 1993) used a coupled map lattice forced by damped sinusoidal modulation of a sample image for texture segregation, with results comparable to other approaches. Particular strengths claimed for the technique are relative insensitivity to the dynamic range and contrast of the original signal, avoiding the tuning usually associated with adaptive filter approaches, and the ability to effectively overcome the classic paradox of region segmentation in noise: how to smooth noise without blurring essential features. More recently, two groups have presented work on segmentation with continuous formulations coupled oscillator models. A system based on simplified Wilson-Cowan Oscillators with local spatial coupling has been developed by Campbell and Wang, with a particular emphasis on fast-synchronization, demonstrating that a network of several hundred oscillators in a one dimensional chain can be entrained in a single cycle (Campbell and Wang 1996). Coupling is dynamic, so that once groups are synchronized the coupling disappears. A special long range connection "global separator" unit acts on all local units to desynchronize oscillations by adjusting their parameters; this effect can be overcome by sufficiently strong local diffusive coupling, giving rise a sequence of activations of objects in connected regions. The state variables of the system are activity levels, with different objects represented by phase separation in time, with some clusters active while others are silent. Up to nine objects could be represented and separated by the system.

Another group has focused on overcoming what they term the "Synchrony - Desynchrony" dilemma, resulting from conflicting requirements for synchrony of oscillators coding the same object and desynchronization between clusters coding for different groups (Zhao, Macau et al. 2000). Like Campbell and Wang, they used Wilson-Cowan oscillators, but with Laplacian (2^{nd} derivative) coupling between oscillators, and with the parameters of the system such that *chaotic* oscillations result.

They claim that this allows effectively unlimited segmentation. It does so, however, by increasing the complexity of read-out of the encoded segmentation. Their procedure for identifying objects in the oscillating field involves observing which sets of oscillators visit a particular phase space region (Poincaré section) simultaneously. Since the chaotic trajectories of different clusters may coincidentally cross the section simultaneously, a decision on the assignment of groups is not made until 3-4 such simultaneous visits are made within a particular time interval. By that time, sensitive dependence is presumed to have separated the trajectories sufficiently.

Finally, I highlight another study using the coupled map lattice formalism to perform segmentation which goes further than most in seeking psychological plausibility by fitting psychophysical data on ambiguous perceptions (van Leeuwen, Styvers et al. 1997). As part of that project, a numerical study of C_{crit} , the critical coupling value leading to synchronization between two maps over a range of random initial conditions, against the bifurcation parameter was performed. The relation is not strictly monotonic, but does generally show increasing C_{crit} for increasing b parameter. C_{crit} Values in the range .16 to .25 appear in the chaotic bifurcation regime.

In the network of van Leeuwen et al., the presence of a signal in the input field reduces the bifurcation parameter of the map to the minimum of the specified range. The background state of the network is uncorrelated, chaotic oscillatory activity.

Coupling is adaptive to a smoothed difference measure between coupled nodes in this model, with weights scaled by a sigmoidal function of the difference function up to a maximum coupling. Therefore, spatiotemporal patterns of synchrony are achieved, with varying stability depending on the parameterization. With the addition of on axis directional preference to weight adjustments, switching between alternative Gestalt organizations is in evidence, with the distribution of switching times qualitatively matching psychological data. It is noted that this distribution is obtained with only one free parameter. A numerical study determined the critical coupling values leading to convergence over a range of random initial conditions vs. the bifurcation parameter. The relation is not strictly monotonic, but does generally show increasing C_{crit} for increasing b parameter. Values in the range .16 to .25 appear in the fully chaotic regime.

SUMMARY

The evolution of states in arrays of coupled discrete oscillators is a rich source of phenomena, ranging from attractors of various types, synchronization and clustering, complex transient structures, and spatial pattern formation. For the diffusive coupling and low iteration counts used in the present work, a few simple trends are evident. Increasing the coupling across a lattice decreases the effective dimension and nonlinearity; units which would be chaotic if uncoupled will become synchronized chaotic (for high b parameter) or even periodic with high coupling. The dynamics of response to structured inputs, used as either initial conditions or to modulate bifurcation or coupling parameters, is relatively unexplored.

Chaotic dynamics can be produced in small circuit neural models and larger ensembles through a variety of underlying pathways. Given the aperiodic, stimulus linked rate modulations and changes in correlation seen in biological networks, to study the dynamics of coupled chaotic systems seems a natural direction for neural modeling.

A researcher familiar only with the well known principles of *low dimensional chaos* – the resemblance of chaos to noise, the sensitivity to initial conditions - might dismiss the relevance of coupled recurrent chaotic systems as a model of neural processing. The added complexity of spatial interactions and coupling, however, can push a CML system either towards linearity (i.e. regarding the temporal or instantaneous statistical response to input), or may provide the substrate for very complex computations, such that correlations between input patterns and measures of the system response become useful tools for neural system design.

To date, very little research has gone beyond pure dynamics studies to perceptual modeling or pattern recognition with chaotic or periodic oscillatory systems. Segmentation is the most well studied area, and I have reviewed several recent contributions from other investigators.

In the next chapter I investigate the ability of oscillatory systems to rapidly form responses to spatial forms. I begin with the parametric study of transients in coupled logistic maps to spatial forms, and ultimately demonstrate a system for recognizing 3 dimensional objects from their 2 dimensional silhouettes. The demonstration shows that if the assumptions of place coding are abandoned, coupled map systems can serve as the physical substrate for algorithmic approaches ranging from the classical (e.g. metric spaces) to more modern (e.g. view based normalization).

Chapter 5: Representation, Learning, and Recognition

The aim in this chapter is to discuss how the dynamics of coupled maps, with their intrinsically complex transformation of an input pattern, contribute to the forming of representations. Given a scheme for representation, which I have argued is plausible from both biological and classical pattern recognition viewpoints, an evolutionary learning process is outlined which is intended to achieve that representation.

A few concepts from formal language theory, Markov chains, and probabilistic finite state automata are used in the following discussion without formal introduction. I have included appendices outlining basic concepts in these fields.

CODING AND REPRESENTATION

In neuroscience and neural network theory, the words coding and encoding are used in a variety of senses, often different than their usage in communications engineering. In the latter, coding normally indicates that for some data stream, a set of p symbols is replaced with q symbols in an alternate, usually smaller alphabet – the code. This is the *encoding* process, undertaken for a variety of purposes including data compression, error detection and correction, or satisfaction of certain constraints imposed by the operating characteristics of the channel or the decoder. Codes are normally *invertible* by some decoding process, indicating that the initial stream can be recovered exactly.

In both natural and biological neural networks studies, representation is a more appropriate descriptor of the transformation process than coding. However, If the representation involves a reduction in dimension from an initial encoding of some input, but can reconstruct the original input, we can properly speak of it as an invertible code.

In biological studies, the term code is often used to indicate that an external observer is able to discriminate between a small set of stimuli by observing a spike train; the neuronal discharges are then thought to realize a one-to-one mapping from a set of input states to distinct output states. This fits nicely the hypothesis of neurons serving as detectors, sampling and coding a single variable, and has allowed a good deal of analysis of coding in low level vision. It is often less clear in what sense the stimulus relates to a *complete alphabet*[27], a set capable of representing all recognizable stimuli, and whether some downstream receiving neuron uses the spike rate as a code, even when some form of correlation exists between the rate and the stimulus. Moreover, the previous discussion on non-classical receptive fields should caution us on accepting that assumption without careful scrutiny, and to consider alternate interpretations of the spike trains.

---

[27] The exception is when the stimulus set is a complete basis set, such as the Walsh functions used by Richmond et al.. Then all possible input stimuli of a given resolution are covered; this is a desirable property.

For example, there are indications from the auditory system that neurons code more efficiently for natural stimuli, due to nonlinearities in their response which match the correlation structure of natural stimuli (Rieke, Bodnar et al. 1995). This indicates that the code is not a linear combination of some basis set.

PARTITION CELLS AND SYMBOLIC DYNAMICS

The idea of coding sketched in the previous section presupposes discrete symbols and alphabets. How are we to bridge the gap between the real valued states and complex temporal evolution of the logistic map units and the world of discrete symbols? The field of **symbolic dynamics** has been developed to address precisely this issue.

The orbit T of a map specifies a sequence of points in the phase space. Suppose that we quantize the phase space into a number of equally sized regions. Now, any point in phase space is identified with exactly one such region, known as a **partition cell.** These cells can have labels *corresponding* to an alphabet ***A*.**

Every (partial) orbit T in this dynamical system can be considered as a sequence of elements or a **word** in the language ***A****. The set of all sequences generated by the system T, along with a given partitioning is a language ***L***.

This simple coarse-graining idea allows complex dynamical systems to operate in the symbolic realms handled by formal language systems, automata theory, and by neural networks utilizing local coding. As one example, symbolic dynamics gives an alternative "negative" procedure for defining a formal language: a language may be defined simply by enumerating forbidden words, i.e. sub-sequences that may not appear in the output of a dynamical system.

In neural networks operating within the localist, rate coding framework, *activation on a particular output unit* is typically interpreted as some form of coding. For high dimensional systems of coupled maps (model oscillators), the natural equivalent is to consider the *population (occupancy) of a partition cell* as an activation state. By normalizing this occupancy by the total number of cells, a distribution over all maps can be used as an output vector with real values. This allows even those sites inhabiting phase space intervals corresponding to *reduced* firing rates to play a role in coding.

Recall the distinction between Marr's concepts of the computational and algorithmic levels of description in a vision task. We can now see that a computational level process - the view interpolation and normalization strategy of Poggio and Edelman described earlier – can be remapped into a different algorithmic level, with correspondingly different assumptions about the underlying primitive physical elements. In the algorithm applied in this study, computation and coding arises from synchronization processes, resulting in coding across partition cells in the state space.

The underlying units of activity corresponding to each lattice sites are ensembles of neurons which cooperate to produce the synchronization process, at both local and distributed scales. Modulations in bifurcation and coupling, at the CML level of abstraction, arise from the interaction of underlying time varying micro-circuit parameters, or the influence of distributed modules whose rhythmic impulses affect these changes. To study the problem algorithmically, however, these underlying causes need not be specified in detail; the dynamics and state flows are sufficient to capture the essential operation.

FORMING REPRESENTATION SPACES WITH PARTITION CELLS

We have now covered the relevant material to introduce an essential feature distinguishing the approach to pattern recognition developed here from the other methods reviewed. *The dimensions in which representations are formed and distances between objects are computed are defined by partition cells in the phase space.* The representation of an object as a point in this space is based on some statistical measurement of the trajectories occupying particular subspaces. This is in contrast to dimensions corresponding to individual features as in classical statistical approaches, or associated with individual units as in feed-forward neural networks.

In particular, the present algorithm uses instantaneous statistics sampled after the dynamical evolution with two stages corresponding to two parameter sets. Alternative measurement strategies could be envisioned; for example, one might measure the overall residence time in particular partition cells.

SYMBOLIC DYNAMICS AND SYNCHRONIZATION: THE ROAD PROBLEM AND GENERALIZATIONS

At this point, I want to consider the problem to be solved by the second (synchronization) stage of the Soca network. As described earlier, this consists of sharpening the distribution (e.g. approaching synchronization into one or more clusters) over the partition cells so that *spatial configurations of initial conditions corresponding to objects in the same category* map to identical or very close distributions, taking the partition cells as dimensions of a metric space.

It seems that this problem structure corresponds to a generalization of an open problem in graph theory known as the *Road Problem* (Lind and Marcus 1995) . The general problem has been open since 1970, while some special cases have been solved (O'Brian 1981). Since progress on this problem might in turn enhance the prospects for more rigorous treatment of bounds on the pattern recognition capability of the Soca network, a brief description seems in order.

The Road Problem can be understood by a colorful analogy. A lost driver headed for Austin calls the auto club for directions; and happily following the directions (consisting of a sequence of intersections and turns), she reaches her destination. But the driver realizes she never told the auto club where her starting point was! This is only possible if the highway network were labeled in such a way that the same sequence of instructions would lead a driver from *any city* to Austin at the same time. Clearly, this is true only for a special subset of graphs and road labels. If a graph has such a structure, the sequence of labels leading from any state to the terminal state is known as the *synchronizing word* of the graph. The formal statement of the Road Problem is restricted to graphs with out degree two:

*If **G** is a directed graph with out-degree 2, is strongly connected, and is aperiodic, there is a road coloring (set of arc labels) with a synchronizing word.*

How is this related to the partial synchronization task? Suppose that the states resulting from the desynchronization stage of the network cover all k partition cells in the state space. If we could construct a state flow graph and transitions such that all states converge to a single partition cell at time t, we would have an exact correspondence to the road problem with $\binom{k}{4}$ colors. The states of the graph correspond to k partitions on the state space, with 4 neighbors in space requiring arc labels with $\binom{k}{4}$ colors.

However, convergence to a single state implies total synchronization, and both intuitively and experimentally we will find that it is difficult to distinguish initial conditions for totally or largely synchronized lattices. Everything looks alike for strong synchronization, and objects will tend to "collide". Instead, we need to generalize the problem to reach *a particular distribution over the partition cells at time t*, or at least to approach such a distribution within distance ε over the space defined by the partition cells.

If we want to avoid mismatches with these classifiers based on partial synchronization, what limits must be obtained. Let $dmin(c_1,c_2)$ be the minimum distance between the mean distributions of any two exemplars of all pairs of classes. Let ε_m be maximum distance of any view's signature distribution to the mean of its class distribution. To avoid collisions in a nearest neighbor search, the normalization process must produce

$$\varepsilon_m < dmin(c_1,c_2)/ 2$$

An analogy for this generalized problem, consider this "All Industries Traveling Salesman's Regional Convention" problem. We want to have the salesman, distributed around a graph of cities, converge to some favorable set of cities; if we are convention organizers, we would like to optimize the distribution to match the conference hall size in each city. We cannot allow each industry to all choose the same set of cities.

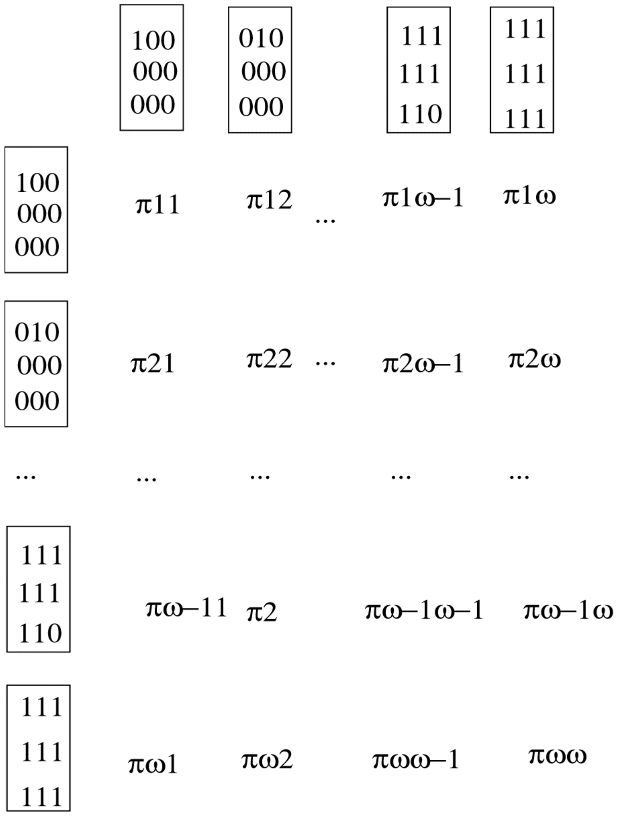
If we wanted a single network with a labeling of arcs and synchronizing word which sends all salesmen in all industries to the correct city, we have a challenging problem; if we want this to occur in some bounded number of iteratations, the challenge increases. The simpler problem is to send all salesmen (lattice sites) in a particular industry (object) to a satisfactory set of cities (partition cells), while trying to handle the "overbooking" problem through rules limiting the conference size and encouraging many small conferences. Even this seems considerably more complex than the original road problem.

DYNAMICAL RECOGNIZERS AND PICTURE LANGUAGES

I have already briefly discussed the idea of a dynamical recognizer and mentioned that previous investigators have considered images as spatial extensions of formal languages. Languages can be described deterministically with grammars; in contrast, the productions of a language are often described with the tools of communications theory.

While the earlier discussion of dynamical recognizers, was strictly in terms of deterministic processing, the use of distributions for representation and decision process recalls the well developed field of *statistical language learning*. While perhaps not strictly necessary, I will describe the deterministic computation involved in representation and recognition in a statistical framework – it seems almost necessary for the purposes of gaining intuition, and it is possible that techniques from that field may be used to improve the present system.

Communications theory allows statistical characterizationof a stream in terms of probabilities of emitting words, or sub-blocks. Enlarging the *memory* or *anticipation* (e.g. including higher order statistics) window of a stream allows more accurate estimation of the probabilities. In a one dimensional data stream, **co-occurrence probabilities** between 2 adjacent symbols (digrams), 3 adjacent symbols, and generally n-blocks can be measured from the stream. A model of such a Markov process, using such higher order higher order probabilities, can generate streams with increasing fidelity to streams generated from a grammar. A large literature on construction of such chains for decision processes exists (Rabiner and Juang 1986); (Charniak 1993).



	100 000 000	010 000 000		111 111 110	111 111 111
100 000 000	$\pi 11$	$\pi 12$	...	$\pi 1\omega-1$	$\pi 1\omega$
010 000 000	$\pi 21$	$\pi 22$	...	$\pi 2\omega-1$	$\pi 2\omega$
...	...	...		...	...
111 111 110	$\pi\omega-11$	$\pi 2$		$\pi\omega-1\omega-1$	$\pi\omega-1\omega$
111 111 111	$\pi\omega 1$	$\pi\omega 2$		$\pi\omega\omega-1$	$\pi\omega\omega$

Fig. 22. A simple stochastic language description of an image or family of images. A matrix of blocks consisting of possible pixel configurations and probabilities that 2 such configurations are adjacent.

In a spatially extended system, the use of neighborhoods and iterations allows the recognition process to compute on the basis of a hierarchy of co-occurance statistics. After the first iteration, each cell (unit, pixel) has first order (2 block) information on the surrounding 4 on axis neighbors; after the second iteration, information from the overlapping 2 block statistics of surrounding 12 cells is included, after the third, information on the second order statistics of 24 cells is available, etc. The first order statistics condition the state flow of higher order statistics.

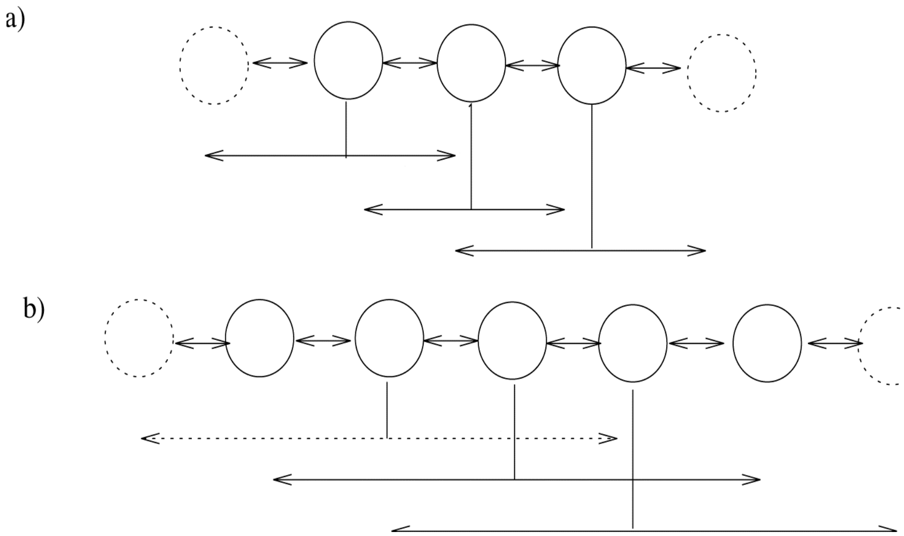


Fig. 23. The integration of higher order statistics over time. a) A one dimensional illustration of the increasing window size. After the first iteration, a 3-block has been analyzed. b) After the second iteration, each cell contains information about a 5-block, conditioned by the overlapping 3-blocks embedded in it. The result is a bottom-up interpretation of information at coarser spatial scales. Information integrated from the highest order statistics and spatial scales occurs in the synchronizing stage of the dynamics.

The end result of the dynamical evolution through two stages and coarse graining to k partitions is that each cell maps a point in a very high dimensional space to a lower one, limited by the precision required to encode the number of partition cells k. Since the images are given as binary images, each cell *after dynamical evolution* has indirect information about local statistics of various orders up to t1+t2, and the entire lattice has such information based on the original image in space of possible configurations with dimension 2^N. After iteration the value is quantized to a point in range 2^k. The occupancy statistics of all such cells are gathered to some precision, which need not be higher than $\log_2 N$, where N is the number of image pixels (and lattice cells). For values used in this study (k=64,N=5626) a dimension reduction of 49:1 is achievable for a family of 7 75x75 pixel images [28].

In the experimental work described later, the partition cell concept is only applied at the end of the dynamics, after sampling. However, in order to visualize how this dynamical system can *compute* similarity, it may help to consider the *intermediate* states as coarse grained, i.e. consisting of transitions between states defined by the

---

[28] The dimension reduction is computed as $7 * 75^2 / (64 * \log_2 75^2)$

partition cells. For chaotic dynamics, nearby points will by definition diverge, winding up in different partition cells. The resulting state transitions can be considered as similar to a *Markov chain*, where the transition probabilities between nodes corresponding to blocks of various sizes. The spatial dependence between cells and the deterministic dynamics result in the mapping being rather forced; it is unclear whether methods for constructing Markov chains from decision processes would be directly applicable.

The following diagrams illustrates schematically the processing of the Soca network serving as a classifier or recognizer of a picture, with the latter considered as a stochastic language. The network state flow is illustrated as a *nonstationary probabiliistic finite state automata* (PFSA), with each state corresponding to one of the partition cells.

The class learning problem consists of finding the two opponent stage parameters which label the edges and cells so as to map productions of the language (i.e. members of a class, projections of a 3D object) to the same distribution over the state transition graph at time t=t1+t2. Each arc is labeled with a transition probability to another state, when a site in that state is surrounded by an input word. The number of arcs between any two states is $\binom{k}{n}$ where *k* is the number of partition cells and *n* the number of neighbors.

The use of Markov chains on partition cells was recognized as a bridge between low dimensional dynamics and information processing models by Nicolis (Nicolis 1986) and Grassberger (Grassberger 1988); however, the Markov chain formalism cannot account for the spatial coupling of the CML. The extension to PFSA, by augmenting the transitions with *input words*, models the neighborhood-determined state transitions. (The use of an initial distribution of states is unusual for PFSA, but standard for Markov chains in distributed systems). To compute the next time step occupancy O in each partition cell i

$$O_i = \sum_{i=1}^{k} \sum_{w=1}^{w*} \pi_{iw} O_{iw}$$

where k is the number of partitions, w is the word label on each arc, w* is the last of $\binom{k}{n}$ combinations for neighborhood size n, and π the probabilites for each input word state pair.

The Soca network parameters {*b1,c1,t1,b2,c2,t2*) and the number of partitions k form a concise specification of this PFSA graph. Of course, the two stages correspond to a relabeling of the transition probabilities, so the label on each arc is of the form

$$\begin{Bmatrix} \pi 1_{ij} w_m, t < t1 \\ \pi 2_{ij} w_{m,} t < t2 \end{Bmatrix}$$

While a PFSA model can be deduced from a CML, the two are not equivalent. The final state distributions computed by the two systems will diverge for images with equal first order statistics but differing in 2^{nd} or higher order statistics. This PFSA formalism is presented only for the purposes of visualization, as it is clearly an inexact mapping; it has less discrimination power than the actual CML. The actual CML is influenced by increasing order statistics (i.e. block adjencies) through time, while the PFSA over partition cells is limited to 1rst order spatial statistics at each step. The diagram on the following page illustrates the probabilistic, neighborhood dependent state flow between partition cells.

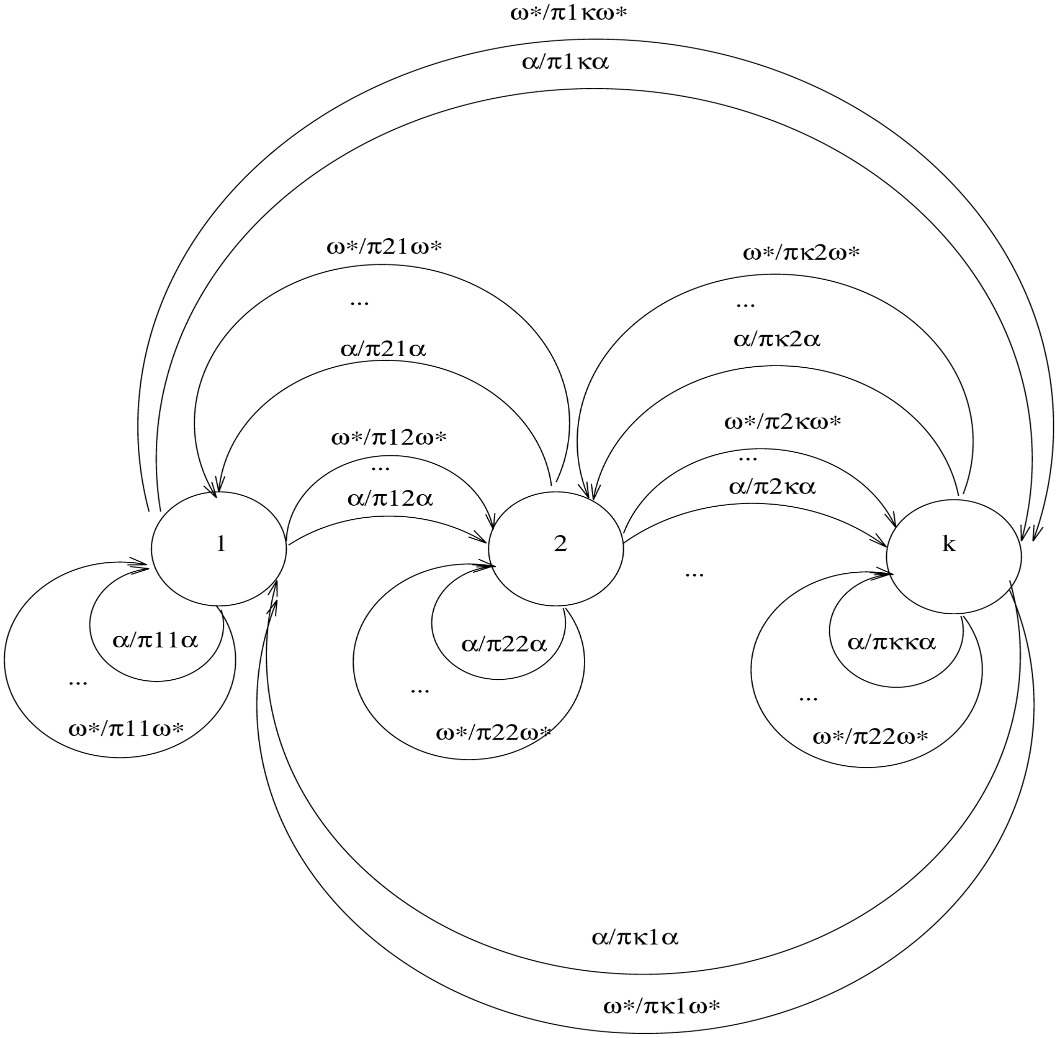


Fig. 24. The statistical operation of CML dynamics can be visualized as (and partially approximated by) a probablistic finite state automata. Each state in the PFSA corresponds to a particular partition cell, or interval of the phase space. Each arc is labeled by the probability of transition from source to sink state given the presence of an input word from alphabet {α,β, ... ω*}, Probabilities π for each input configuration and state are given on the arc; for simplicity, only one of the two stages in the synchronization opponent network is shown.

In summary, progress on mathematical frontiers including:

1. the computation of transient distributions for ensembles in coupled map lattices
2. coupling bounds leading to synchronization
3. the Road problem in automata theory

could result in the replacement of a heuristic search algorithm with a more direct solution, or to provable bounds on the power or limitations of the algorithms.

In the absence of such developments, improvements in the learning process beyond the present work might result from consideration of these issues. It is fairly easy to see, for example, that the state flow graph implicit in the second (synchronizing) stage must have paths of roughly equal length leading from the states occupied at the end of the first (desynchronizing) stage. It is less easy to see how to translate this to effective contraints on network parameters.

EVOLUTIONARY DESIGN OF SYNCHRONIZATION OPPONENT CLASSIFIERS

The mountain – Buddha's body
The torrent - his preaching.
Last night, eighty-four thousand poems.
How, how make them understand?

Sotoba, a Chinese layman (Translation by Lucien Stryk)

Evolutionary computing is a form of search intended to mimic aspects of biological evolution. A candidate solution to some problem is conceived of as a *phenotype*. A representation which can construct such a solution is the genotype, a collection of genes which code for individual parameters of the solution. Search takes place by forming a population of genotypes, evaluating the *fitness* of the phenotypes they specify according to some objective function, and letting the genotype material of the winners of this *selection* process assume a greater proportion of the population in subsequent generation. To reduce the likelihood of finding *locally optimal* solutions, the conservative action of selection is balanced by exploratory behavior in the form of mutation and recombination operators acting on the genotype.

Neural networks are generally specified by their topology, the activation functions of the units, connection strengths or weights, and any coordination or leaning structures (Balakrishnan and Honar 1995). Any or all of these may be subject to construction by an evolutionary learning process, but often known characteristics of the problem will constrain the network architecture. The learning process may fix some values, or may put constraints on some values (i.e. unit thresholds), while the problem specific values are set in specific.

Genotype representations represent an encoding of network parameters. This encoding may be direct, or indirect. A mapping of the genotype directly to weights in a fixed topology would be direct; a mapping of the genotype to a set of rewrite rules which must be recursively applied to specify a network would be indirect. The network here specifies bifurcation, coupling and number of iterations with a fixed, regular lattice interconnect topology; this is a clear example of direct encoding. In this example, the constraints on ranges that may appear in the genotype are set manually to rule out certain uninteresting parameter ranges.

Typically, the fitness function is a direct measure of the performance of the phenotype on the engineering task. In the present work, that approach is used in one toy problem, namely the search for parameters producing a quasi-metric space for a family of parametric curves.

However, for more complex tasks complete testing of candidate solutions may be impractical, at least for portions of the learning process, due to the computational expense associated with full testing. For example, the more challenging experimental task described in the next section is to identify particular objects from a set of similar objects, based on observing a single view after training on two or more views. Evaluating performance on the final task involves not only testing that all views of the object are recognized as the same object, but also that other objects are not mis-classified. Since the computation of the candidate solution is a relatively expensive nonlinear filtering operation, it is impractical to apply it to the full database of views, or even to a sub-sample to evaluate misclassification, during the evaluation of 3000 candidates phenotypes.

Instead, the performance on part of the task (recognizing different views of the same object) is emphasized, while certain *constraints on the output or encoding* produced by the network phenotype are incorporated into the objective function.

In addition, a reproduction strategy was chosen which preserves the best phenotype in each generation, guaranteeing a monotonically improving score.

Implementation and Performance Details

The current Matlab implementation of the evolutionary learning algorithm runs 30 generations of 100 genotypes. For the major object recognition task, learning the best representation for each object in the database (by using all 7 views) used 40-50 minutes/object on a 400 MHz PowerPC 750, 1M cache, 100 MHz system bus. The majority of time in learning (85%) is spent in the image convolutions of the diffusive step of the CML dynamics. The time variability results from branching around the computation of some objective function terms when a basic synchronization criteria is not met.

Each stage, or *attractor frame* consists of a triple $\{b, c, s\}$, where s is number of iterations in the stage. Two such frames are applied in turn. During evolutionary search for the parameter sets, the first stage is constrained to 2-6 iterations, the second constrained to 2-9 stages. In the evolved solutions, the total iterations required to create the representation space ranges from 6 to 14 iterations, with a mean of 11.3 iterations for the 39 objects in a 3D object recognition task. Each object in this set is 75x75 pixels, plus padding proportional to the number of iterations, to prevent array boundary diffusion statistics from contaminating the statistics. Thus each typical network phenotype evaluated represents roughly 77000 (40000+ 30000 * 11) multiply operations for the convolution and nonlinear map operations.

The simplex reproduction strategy (Karr 1991) governed the transmission of successful genes. It is well suited for situations where the time complexity of evaluating the gene is much greater than the processing associated with the evolutionary strategy, including complex image processing operators (Brigger 1995). In the simplex strategy, each adjacent pair of genotypes in the pool is evaluated as a local tournament, with the winner promoted to the next generation. The overall best-fit genotype in each generation is saved in element 1 of the genotype array, and is protected from mutation; this insures that the fitness is monotonically increasing and effectively makes the learning procedure a form of gradient descent. All winners of the pairwise tournaments are subjected to crossover mutation; given an effective rate of crossover of .5 for the whole population. The losing genotypes are replaced with copies of the best genotype from the last generation. These replacements are subject to mutation at rates given in the table in the experiments section.

The evolution parameters in these scenarios are the bifurcation parameters (b1, b2), the coupling parameters (c1, c2) and the iteration count of each of two dynamical synchronization opponent stages (t1, t2). Bifurcation and coupling are represented as doubles, the number of iterations as integers. The number of partitions k used to record the instantaneous occupancy at time t1 + t2 is also a critical value in the network performance; the values used in the work here were 256 for the curve evolution experiments and 64 for the object recognition experiments. This system parameter was not subjected to evolution, but held constant within a family of experiments.

Mutation was applied to the genotype copied from the best in generation to the pairwise competition losers, and to the winners after crossover. The mutation probabilities are set independently for each parameter as shown. Constraints on the evolution of parameters, as well as a general form of the desired solution (i.e. two dynamical "phases"), were provided as assumptions for this study.

The use of such constraints is perhaps unusual for evolutionary computing, and might be considered undesirable in the sense of biasing the solution. I argue that such constraints are justified here for at least two reasons. The general hypothesis being explored here was that a metric space could be created based on the dynamical flow, and solutions not conforming to this were not desired. Without such constraints – chiefly the lower bound on bifurcation parameter – solutions would rapidly emerge in the fixed point parameter regime which were clearly based on the size of the boundary. The high coupling regime, which can also lead rapidly to a fully synchronized fixed point, was similarly avoided. Such solutions could be avoided by adding terms to the fitness function; exactly this approach is taken in the next section, where excessive synchronization was found to degrade the performance of an object recognition system. Avoiding the regimes directly by restricting the parameters improves the search efficiency.

The learning algorithm is summarized below in pseudocode. The actual code uses an object-oriented control flow, but here I adopt a procedural style in the interests of clarity.

The variable *viewSet* is a set of related *images* (e.g. exemplars of a parametric curve, or different views of an object rotated in depth). The complete set of such families is the *objectWorld*. Variable *population* is an array of genotypes {b1,c1,t1,b2,c2,t2} and a field fitness to record the result of evaluating the fitness function.

```
procedure learndb
  for viewSet in objectWorld
    population = initializeGenotypeRandom;
    for generations = 1 to maxGenerations;
      generation(viewSet, population);
      population = simplexReproduction(population)
    end // generations
   objectDB(viewSet) = normalizedHist(lattice(viewSet, population(1))
   genotype(viewSet) = population(1);
  end  // viewSet
end learndb

procedure generation
   for all genotypes in population
         for each image in viewSet
          shapeDB(image)= normalizedHist(CML (genotype,image))
         end image
         genotype.fitness = fitness(shapeDB(image))
   end genotype
end generation
```

```
procedure simplexReproduction(population)
  bestfitGenotype = find genotype with best fitness in population;
  if fitness of bestFitGenotype < currentBest, replace;
  strongHalf = winners of pairwise competition;
  apply crossover operator on winners;
  copy currentBest to losers;
  apply mutation to copies of currentBest;
  return population;
end simplexReproduction
```

Summary of Evolutionary Learning Method

A evolutionary learning framework for image processing problems was developed to train a two stage coupled-map lattice network using a computational strategy known as synchronization opponent cooperative action. The framework allows substitution of objective functions depending on the specific image processing task. Two tasks, described in the next part of the thesis, both require the formation of representation spaces whose dimensions correspond to partition cells in the network phase space. The objective functions differ slightly, however.

In *learning to order examples* from families of parametric curves, direct scoring on the position of points in the representation space is used. In *learning to identify an object* in a set, *direct evaluation* on a stimulus equivalence measure is combined with *indirect* evaluation of inter-object discrimination during the learning process. The representations emerge from these constraints, thus the learning should be considered as *unsupervised*.

The objective functions used vary depending on the exact task being performed, and are consequently described in the appropriate context in the following experimental work chapter; some details on the organization of Soca software package are given in the appendix.

An example of the evolutionary learning procedure for one task, the recognition of paperclip objects rotated in depth, is illustrated in the figure below:

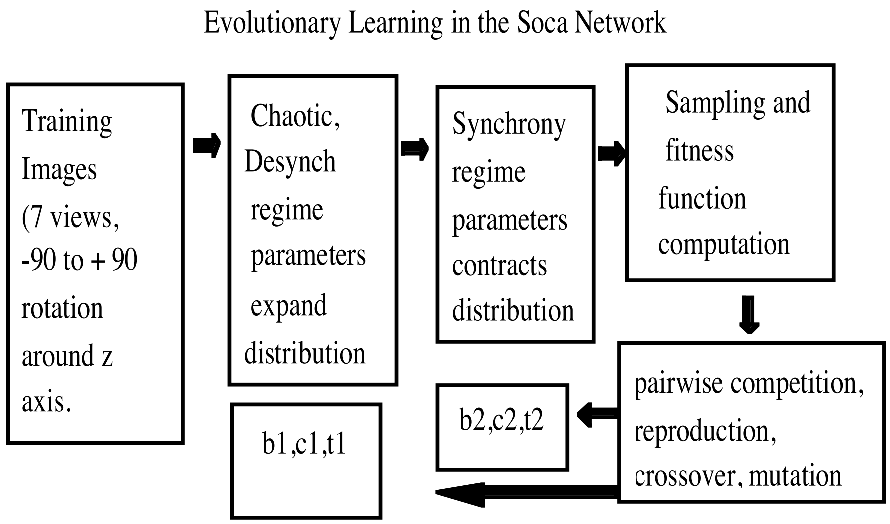


Fig. 25. Binary images provide the initial condition for a network of recurrent logistic map units. The network is homogeneous in space, with bifurcation and coupling parameters between cells identical but time varying. The final result of training is a set of six parameters (genotype) and a distribution of state values obtained after a small number of time steps.

RECOGNITION WITH AN ENSEMBLE OF DYNAMICAL RECOGNIZERS

In this section, I describe in detail how the *recognition* process functions when the Soca network concept is applied to a pattern recognition task. The task and results are described in detail in the next chapter.

Recall the description of a dynamical recognizer given in the first chapter. A recurrent dynamical system was shown by Pollack to function, after training, as a recognizer for a language by reaching an accept or reject state after iterating over a bit string. In that case, the network functions much like a finite state automata. Each bit in the string is a bit in an input stream, with the next state computed from the input and the current state. Given an appropriate partitioning and labeling of the network's continuous state, the input output relation could be preserved and a finite state graph produced for a regular language.

The situation here is more complex, due to the higher dimensions and the 2-D spatial arrangement. Each cell in the CML functions like the one-dimensional recognizer in Pollack's scheme, with the following differences:

1. Rather than sequencing over sequence of raw states, at time t the cell receives a set of 4 inputs processed t-1 times by neighbors. Only cells in the first iteration receive raw input.
2. No cell can reach a decision on its own; the representation is distributed over the population of all units.

In Pollack's recognizer, when the last bit in a pattern is processed the recognition process is complete. Here, due to the parallel spatial processing, all bits are processed in the first iteration and every subsequent iteration; the number of iterations until recognition is given as part of input to the recognizer. In the Soca network, the *iteration counts are part of the genotype.* If recognition in a fixed number of time steps were desirable, this could have been given as a constraint, but the effectiveness of the search process under that constraint remains to be studied.

The nearest neighbor recognition process after creating a database for all of the objects, in pseudocode form, is as follows:

```
procedure bestMatch(image, genotypes, binMeans)
    for each genotype in database
      hist = dlattice(image, genotype);
      if (thisDistance(hist)  < bestDistance)
        return  index(genotype);
      else bestDistance = thisDistance;
end // bestMatch
```

In an alternative testing paradigm, two views of objects are presented; the algorithm applies the best match procedure as given above; if the same object is selected, the trial is a match, otherwise a no-match response. This allows the possibility of false positive and false negative judgements.

Chapter 6: Computational Experiments

EXPERIMENT GROUP 1: SAMPLED TRANSIENTS OF COUPLED MAPS

Dynamical systems theory inherits from analytical mathematics an emphasis on asymptotic states, even though analytic solutions remains elusive for nonlinear, and particularly high dimensional nonlinear systems. Accordingly the literature on the time evolution of transients in essentially non-existent, though certain phenomena have been noted. For example, most treatments of bifurcations in period-doubling maps mention the phenomena of *critical slowing down*, in which the average time to reach the attractor increases in the vicinity of a branching critical point, where the attractor loses stability.

In order to study the evolution of transients in maps, several sets of simulations on asymmetric logistic maps were performed. These encompass both uncoupled, coupled, and two stage *synchronization opponent* systems. In some, a fixed initial condition is explored over a range of the {b (bifurcation), c (coupling)} parameter space. In others, an image generated by a parameterized function is given as an initial condition to the lattice. The image is systematically varied while one dynamical parameter is changed. Animations of the resulting distributions of states over the parameter and input space help to understand whether such dynamical behavior can be used in classification systems, and what limitations to expect.

In most cases these simulations were perfomed after the curve ordering experiments discussed in the next section; the parameters used were chosen from the set found to produce interesting behavior (i.e. the mapping of a curve family to distributions supporting ordering the curves by occupancy statistics defining a point in a partition cell metric space). The number of iterations in each study is low (5-9 iterations), relative to both standard practice in studying CML evolution from random initial conditions, and relative to the 10-20 iteration counts for recurrent neural ensembles to be considered biologically plausible. The lower numbers are chosen to emphasize the possible behaviors in a single processing stage in a two stage system.

When the term *final state* is used below, it simply means an arbitrary sampling time in the transient stage of an ongoing dynamical evolution, and does *not* imply that any asymptotic state or attractor has been reached. This sampling time is one of the parameters evolved by evolutionary search for the other experiments in shape similarity and object recognition. The sampling time is, rather, implicit in the sum of the two stage iteration parameters {t1 ,t2} of the Soca network.

Convergence Time Distribution for a Single Logistic Map

For constructing engineering systems and modeling biological processes, it is essential to study the rapid evolution of networks of chaotic units (i.e. the behavior in the attractor basin *transients)*. The best known measure of convergence and divergence for a particular instantiation (bifurcation parameter) of the map, Lyapunov exponents, is defined for times approaching infinity (Ingraham 1991). However, for computing with transients, knowledge or this number may in itself be inadequate; to estimate lower or uppers bounds on the number of iterations expected to play a role in computations, we can only perform numerical studies on the short term convergence behavior for uncoupled or coupled maps. The following diagrams illustrate convergence times for two bifurcation parameter values in the period 2 limit cycle regime. For the purposes of constructing representations during the transients, it is crucial that the coupled systems should not converge too rapidly from random initial conditions; hence, it is encouraging that such a large range of absolute convergence times are possible in the ensemble.

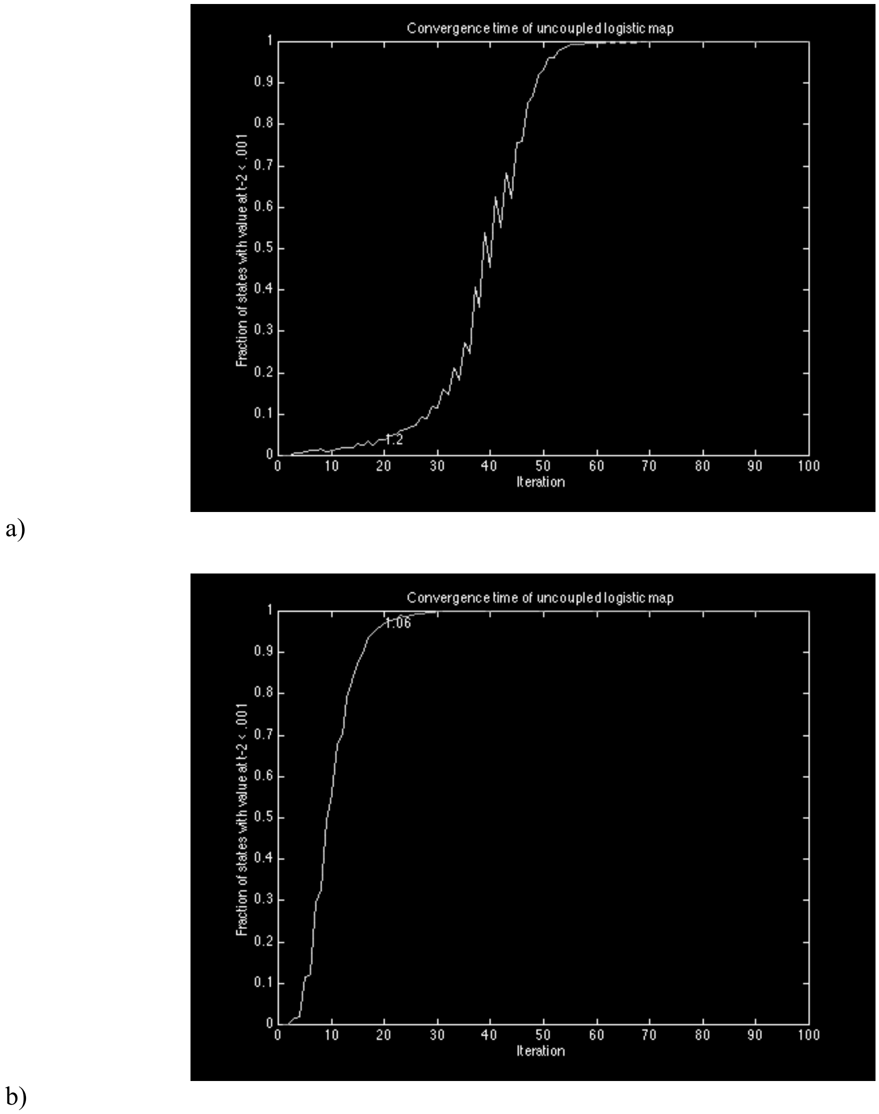


Fig. 26. Convergence times for an ensemble of 1000 initial state pairs.uniformly distributed from the interval [-1,1]. The number of iterations to reach $x_t - x_{t-2} < .001$ is plotted. a) average convergence time for b=1.2 b) average convergence time for b=1.06. the steep slope and lower absolute time to reach any given fractional threshold indicates a greater negative Lyapunov exponent.

Synchronization Time for Two Coupled Maps

The previous section illustrated the time for a single map to reach stable values of period 2 oscillations, until the period 2 oscillations break down. In this experiment, we examine two coupled map units with the same logistic equation but symmetric coupling. Between iterations, a diffusive coupling step adds to each map’s state a fraction of the other map's state, scaled to avoid exceeding the stable domain of the map.

$$
\begin{aligned}
x1_{t+1} &= 1 - b((1-c)(x1_t + c * x2_t)^2 \\
x2_{t+1} &= 1 - b((1-c)(x2_t + c * x1_t)^2
\end{aligned}
$$

Here, the time for two coupled maps orbits to approach to one another to some distnce ε from an ensmble of random initial conditions (x1-x2 <.01) is examined for a coupling value (.38) which achieves synchronization until strong chaos is induced in each map. The bifurcation parameters are varied from sub-critical 1.5 to the maximum 2.0. Illustrated below are two values of b, first sub-critical (i.e. below the uncoupled transition) and the second beyond the uncoupled transition to chaos. While the maps may not be fully synchronized at this first approach, they certainly are *not* synchronized prior to this time.

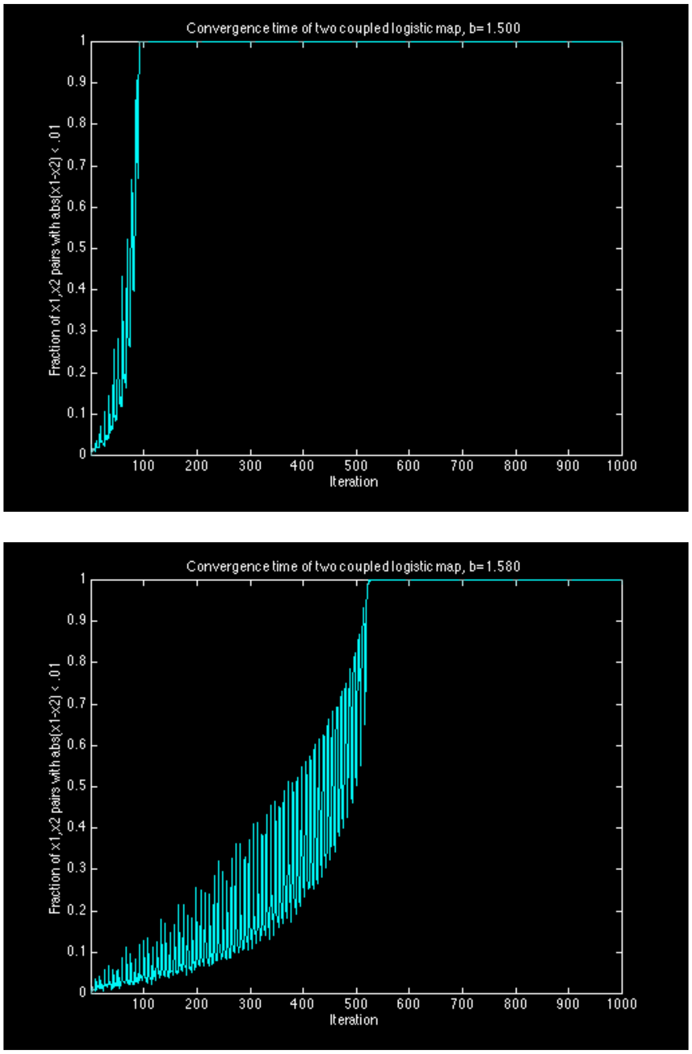


Fig. 27. First approach time for a) b=1.5 , c=.38 b) b=1.58, c=.38. The transition to chaos for uncoupled maps is approximately 1.54.

System of Two Coupled Discrete Maps: Basic Behavior As Bifurcation And Coupling Parameters Are Varied

Time Evolution of Single Map in a System of Two Coupled Logistic Maps with Constant Parameters

Two maps are iterated from initial conditions .0001 and .9999. Each map iterates the logistic equation

$$x_{t+1} = 1 - bx_t^{\,2}$$

Between iterations, a diffusive coupling step adds to each map's state a fraction of the other map's state, scaled to avoid exceeding the stable domain of the map.

$$x1_{t+1} = 1 - b((1-c)(x1_t + c * x2_t)^2$$

$$x2_{t+1} = 1 - b((1-c)(x2_t + c * x1_t)^2$$

This animated surface shows the time evolution of the .9999 map, with the axis indicating fixed time parameters for bifurcation and coupling. In this early transient portion of the map's orbit, the variation between adjacent points in the parameter space is seen to be relatively smooth. With increasing time, the surface will become irregular as the bifurcations implicit in the dynamics separate the orbits into attractor basins. Note that the procedure is carried out on a 200 x 200 matrix of states, with each matrix element value computed from corresponding matrices of *b* and *c* values.

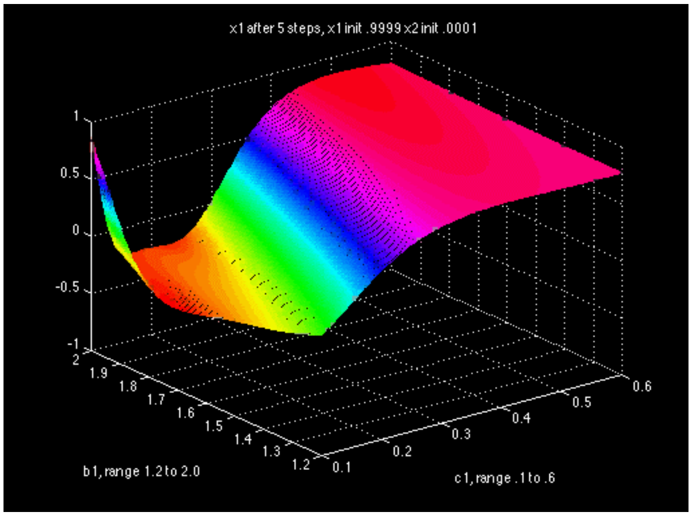


Fig. 28. A snapshot of the parametric variation in the time evolution of one coupled map cell after 5 steps, a typical single stage iteration count used in subsequent experiments for the formation of representation spaces. The smoothness of the surface suggests that network operation should be robust in the face of small variations in the parameters.

Final State in First Map of Two Map, Two Stage Constant Parameter System as Initial Conditions Vary

To investigate whether the particular initial conditions chosen to translate binary images into the logistic map phase space plays a large role in the outcome, we examine the state of one of two coupled maps over a range of bifurcation parameters when the initial conditions are varied to decrease the initial distance between the maps. The dynamics here are representative of the synchronization opponent style, with the second stage held constant (b2=1.3694, c2=0.19134, 4 iterations) while the first stage b and c are varied for 5 iterations. (The second stage values were drawn from parameters found to produce a distribution supporting metric distance functions matching a parametric curve).

We can see that the large separation in initial conditions plays a significant role in creating a complex surface in the transient state, supporting. adaptive behavior by delaying rapid synchronization.

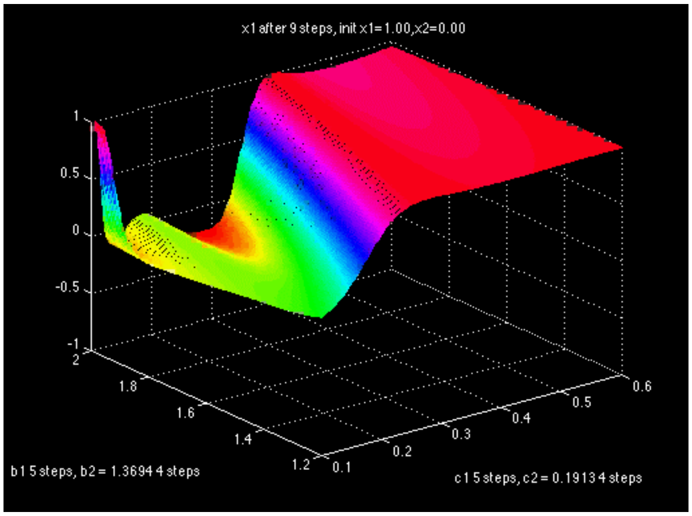


Fig. 29. Sampled orbit at 9 iterations with initial conditions 1 mapped to 1.0, 0 mapped to 0.0001. b2=1.3694, c2=0.19134 for 4 iterations, b1 and c1 as plotted for 5 iterations.

When the initial states are not well separated, synchronization occurs rapidly for most c values, thus only a small coupling parameter range results in any separation, thus the ability of the sampled distribution to construct state flows mapping curves or arbitrary shapes to a metric output distribution is reduced.

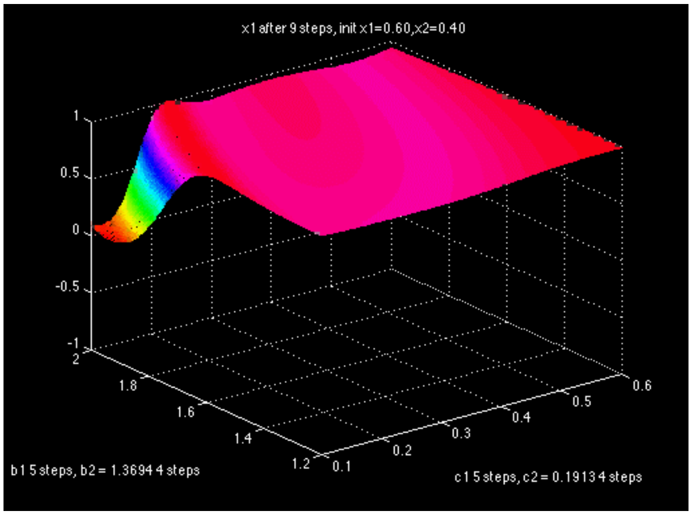


Fig. 30. Sampled orbit at 9 iterations with initial conditions 1 mapped to 0.6, 0 mapped to 0.4.

Effect Of Coupling And Initial Condition Separation On Final State In A Two Stage Parameter Cycle

High coupling values result in the eventual convergence or synchronization of maps, even when their individual bifurcation parameters would result in chaotic trajectories.

These two animations show the effect of relatively weak and strong coupling on two maps, for a two-stage dynamics. The surface shown is the difference between the final states after 5 steps of c1 and b1, 4 steps of b2 and c2.

In this case, the bifurcation parameter in the second stage is a fixed value (b2= 1.3694), while the first stage parameters range over the values shown. Each subsequent frame in the animation shows a different set of initial conditions, with the difference between x1 and x2 decreasing. The first frame corresponds to the values actually used in the subsequent experimental work. The b2 value chosen is one found in the evolutionary computation process for similarity ordering.

Difference Between Final States of Two Unit Synchronization Opponent System With Medium Strength Coupling (c2=.1913)

Note that the complexity of the surface shape in the low c, high b regime supports the ability of the lattice dynamics to preserve a linear relationship between the generating parameter producing an initial distribution of adjacent states and the final state. High values of c1 are seen to synchronize in 9 steps, even with low c2 coupling. Bifurcations values are as plotted with b1=b2.

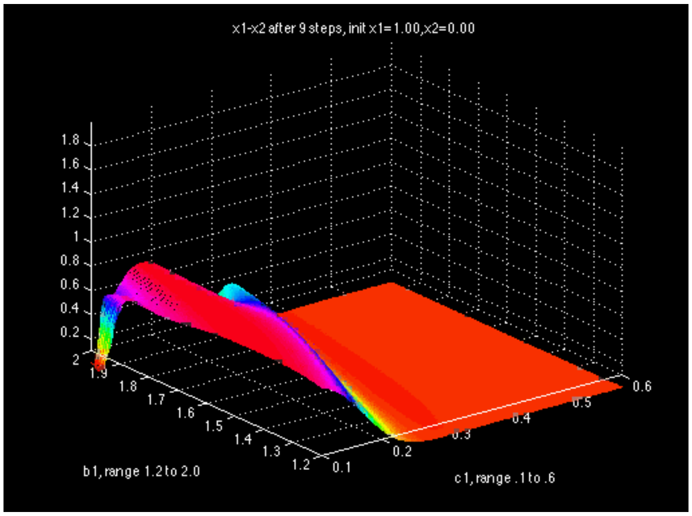


Fig. 31. Difference between two coupled logistic maps in synchronization opponent stage with b2= 1.3694, c2=.1913 and b1, c1 as plotted.

Difference Between Final States of Two Unit Synchronization Opponent Sytem With Stronger Coupling (c2 = .3913)

With stronger coupling in the second stage (steps 6-9) the final states are nearly synchronized for all values, whether the initial states are separated or not.
Obviously no distinctions can be made in this regime, after only 9 time steps.

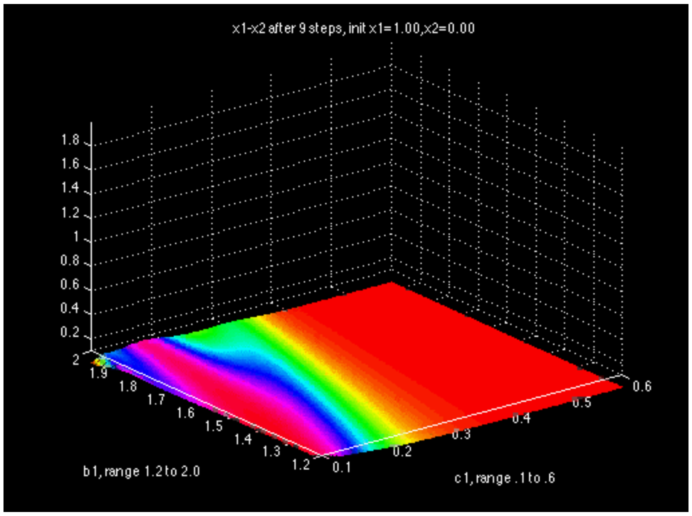


Fig. 32. Difference between two coupled logistic maps in synchronization opponent stage with b2= 1.3694, c2=.3913 and b1, c1 as plotted.

Sampled Transient Evolution in Lattices with Varying Parameters, Random and Structured Initial Conditions

Final State Distribution After 8 Iterations Varying b and c with Random Initial Conditions

A 200 x 200 matrix is initialized with random values uniformly drawn from the set {0.0001, 0.9999}. After 8 iterations for a particular pair of b and c values, the instantaneous distribution of states is measured with 256 partition cell bins. The procedure is repeated 20 times with new random matrices, and the average population at each {b, c} pair is recorded. Each frame of the animation shows a family of 256 bin histograms ranging across the bifurcation value b for a constant value of c. For each b value, the entire distribution is plotted (i.e. any bins with equal population would be overlaid by the last drawn bin). The first frame illustrates that with no coupling, the two initial states evolve independently to state values determined by the b parameter; with the different colors indicating the region of phase space visited in the 8th iterate.

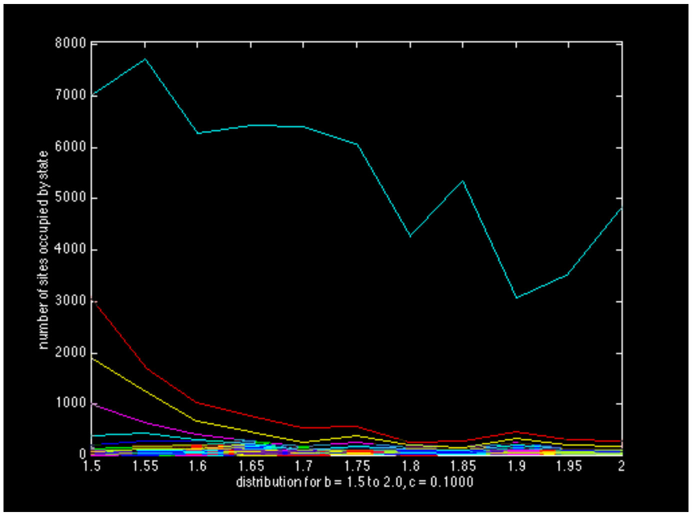


Fig. 33. Distribution of states after 8 iterations for spatially random matrix with site initial conditions chosen from the set {1.000, .0001}.

A few trends are evident. Low coupling with bifurcation values just above the transition to chaos of the uncoupled map (b = 1.54) produce distributions with several major components well separated. The smooth evolution of the modes as bifurcation is varied - for example, with c=.1, b ranging from 1.6 to 1.7 - is somewhat similar to that produced by the response of a system to variations in a spatial form.

This suggests that joint statistics of two interacting networks - one driven by high contrast spatial noise in such a parameter range, coupled to another with a response which linearly tracks the 2-D projection of a 3-D object - might serve to achieve a more nearly invariant response, while retaining metric properties as studied in the next section. Learning to recognize variants might consist of tuning the bifurcation parameters in the randomly driven subnetwork to offset the natural distribution changes associated with different views. This approach is not pursued further in the present work.

Final State Distributions with 45° Line Initial Condition Varying b and c Parameters

A 45 degree line is rendered, with the resulting image matrix mapped to values .0001 and .9999 in the lattice. The coupling between cells is increased between frames.

The x axis shows the bifurcation parameter b ranging from 1.5 to 2.0, while the y axis shows the population of each partition cell (histogram bin) for that value of b.

Each bin is plotted in the same color across the range of b values. Note that there are regions with multiple high population modes (i.e. c=0.1, b = 1.84 or 1.87), while others have a single low population mode with broad dispersion over the remaining bins.

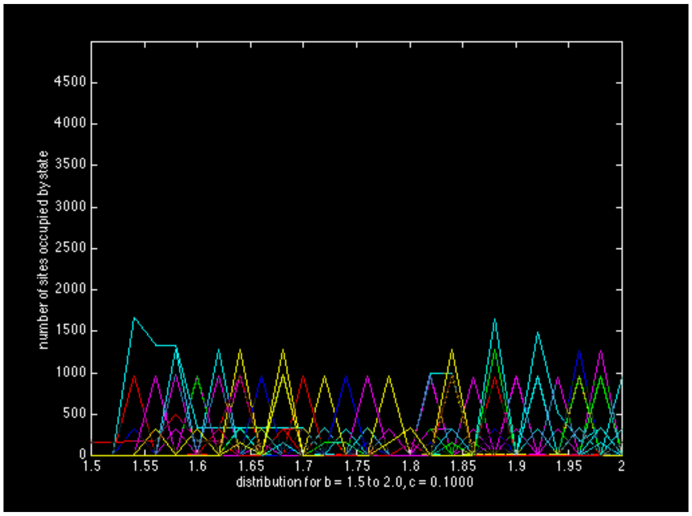


Fig. 34. The state distributions after 8 iterations of a single stage CML with a 45 degree line (line value 0.999 in a background of .0001) used as the initial condition. c=0.1

Final State Distributions with Fixed c varying b and Linear Initial Condition Rotated 0° to 45°

An interesting value of c (0.5 with some strong response modes) is selected from the previous experiment; holding that c value constant, each frame of the animation shows the distribution over the bifurcation parameter b as the slope of a plotted line (mapped to the .0001 and .9999 values) is varied.

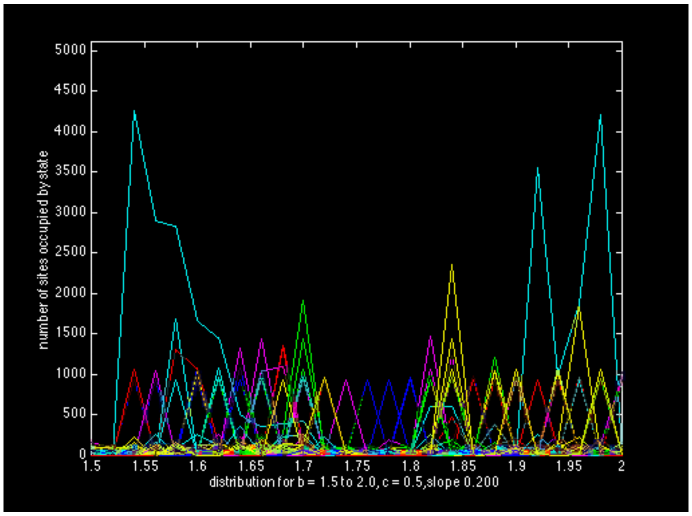


Fig. 35. Distributions after 8 iterations of a single dynamics stage with constant parameter c= 0.5, as the orientation of a diagonal line initial condition is varied from 0-45° between frames. The line shown in this animation frame has slope 0.2.

Discussion

The surfaces formed by sampled transients of logistic maps across the parameter control plane are rather smooth. This has both benefits and drawbacks in the engineering application of sampled transients. It suggests that smooth changes in a curve family or image can be matched by adjusting the CML control parameters to shift some distribution smoothly between bins. (This will be developed further in the next section). On the other hand, the simplicity and smoothness of the sampled surface might suggests that perhaps only simple shapes could be tracked by a network with spatially homogenous bifurcation and coupling parameters.

However, the smoothness in response of a single coupled site does not necessarily translate to simplicity in the network transformation of a form. The experiments with a 45 degree line and a line of varying orientation show abrupt changes in the distributions as dynamical parameters are varied.

EXPERIMENT GROUP 2: ORDERING OF PARAMETRIC CURVES

Data Selection and Methods

There is no universally accepted metric for similarity of shapes (or objects in general), though many have been explored as reviewed in the background material. For this reason, the next step in this study of spatial computations via partial synchronization focused on a restricted class of shapes where a metric is well defined: that of parametric curves. For a set of curves generated from an equation with one free parameter, the distance in the perceptual space of the resulting forms is assumed to vary linearly with distance between the generating parameter. This implies an *ordering* of the images generated for each curve. For each curve, I generated a set of 6 exemplars over a range of the parameters, captured a bitmap of the plot, performed binary to floating point conversion and scaled the values of the resulting array to fit the domain of the map.

The fitness function employed to evolve the parameter sets was simply to match the *order* of the distance between curves to the order expected from the curve's generating parameter. I employ a simple distance metric based on the post-evolution state statistics for the entire shape, with no information about local adjacency statistics (co-occurrence matrices) used in the evaluation of distance between the curves.

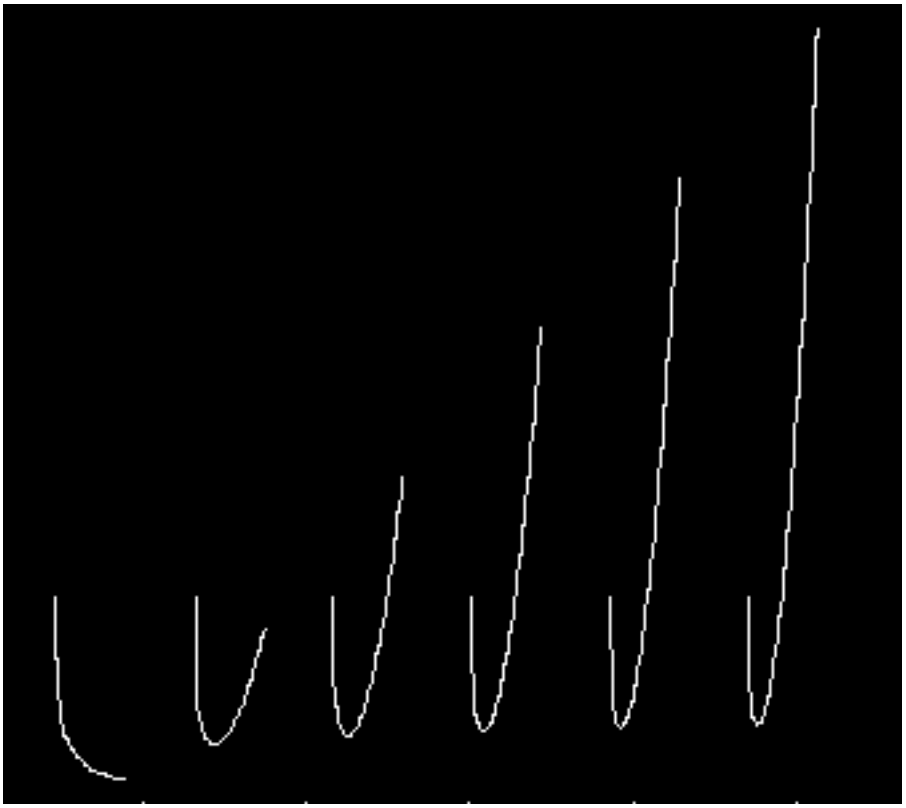


Fig. 36. Exemplar images for one of the parametric curves used in this set of experiments. The images are plots of the equation $y = c * x^2 - \log x$, with parameter $c = 0, 1.5, \ldots 7.5$.

The evolutionary programming algorithm, described in the last section, searches for network parameters effective in ordering the parametric shapes according to their generating parameters. A learning trial consisted of 50 generations, with a phenotype pool of 50 individuals per generation, or 2500 individuals. Perfect orderings were found for all curves with this simple termination strategy.

The fitness function employed to evolve the parameter sets was simply to match the expected ordering shown below. Each curve was represented by six variants, thus the expected ordering of distance was represented by the matrix:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 5 & 6 & 5 \\ 3 & 3 & 2 & 3 & 4 & 4 \\ 4 & 4 & 1 & 6 & 3 & 3 \\ 5 & 5 & 5 & 2 & 2 & 2 \\ 6 & 6 & 6 & 1 & 1 & 1 \end{bmatrix}$$

Pairwise distances from each curve were computed with a weighted Euclidean metric on the space defined by k=256 partition cells. The resulting distance array was sorted, and distance replaced by the corresponding exemplar number, resulting in a matrix suitable for element by element comparison with the order matrix above. A perfect score of zero indicated that the distances obtained between the distribution vectors were identical to the matrix above.

Given that the distances between adjacent neighbors were theoretically equal, mismatch scores were computed according to row sums (1, 2, 3+4, 4+5, 6), with respective weights (10, 3,2, 3). The first and last rows were weighted more highly, given that values in rows 3,4, and 4,5 could be swapped. The high weight on the first row ruled out degenerate highly synchronized solutions, with a single partition cell occupied. This would essentially create a representation space based on size alone, which the genetic algorithm exploited prior to this adjustment.

Mutation probabilities for each network parameter and constraints on ranges are given in the following table. While the mutation probabilities would be high for some reproduction strategies, the simplex method always preserves the best individual of the previous generation and higher mutation rates serve to balance this conservative tendency.

Table 4. Mutation probabilities and ranges for each network parameter.

parameter	mutation probability	Range
b1	.40	1.2-2.0
c1	.25	0.1-0.7
t1	.30	2-6
b2	.40	1.2-2.0
c2	.25	0.1-0.7
t2	. 30	2-6

Experimental Results

Summarizing, I set out to examine the statistics of families of curves after a dynamical evolution in coupled map lattices. I observing that *two independent stages of CML parameters* in the dynamical transients produced distributions which served as a space with metric properties, with the partition cells of the dynamics serving as dimensions of the space. The study set out to answer the questions:

1. What trends are seen in the parameterization of the time varying coupled maps for assessing similarity, given bounds on the number of iterations? I know of no analytical procedure for computing the parameters, so any such optima would be discovered via evolutionary programming. If there were a single optimal value, perhaps the transition to chaos, the network should choose the same parameters for both stages.

2. Is there a single optimum parameter set for all such curve families, or would different curve families result in unique optimum parameters?

The answer to the first question is that the general principle of phase space expansion and contraction proposed in an earlier study appears valid; such solutions were consistently found through adaptive search. For one curve (astroid), a solution with all chaotic-regime local dynamics and high coupling was found. Network dynamics do not cluster around the transition to chaos for uncoupled maps, though this value should be shifted by coupling. This adds support to the earlier refutation by counterexample (Mitchell, Hraber et al. 1993) of optimistic claims for the universal utility of transition–to-chaos dynamics.

The answer to the second and perhaps more interesting question was negative. Unique parameter sets were found for different curves. An explicit attempt at evolutionary search for *a single parameter set* which would produce the proper ordering *for all three curves* was not successful, using the same evolutionary parameters and synchronization opponent stages which successfully ordered each individual curve family.

Table 5. Evolved parameter sets for four curve families

curve	b1	c1	t1	b2	c2	t2
ellipse	1.8671	0.1101	6	1.6392	0.3623	8
ellipse 5 cycles	1.7834	0.5515	5	1.9654	0.3604	2
astroid	1.7084	0.4546	3	0.4937	0.1974	5
pursuit	1.6739	0.2899	2	0.7567	0.3416	5

The table above lists the evolved parameter sets which served to correctly order each family, indicating that the partition cells form a quasi-metric space. Six exemplars were used to train the network. The critical point (i.e. the transition to chaos) for an uncoupled map is b=1.542.

It is reasonable to ask whether any common features characterize the selected parameters and their distributions. I measured the distributions of the resulting images after the CML transformation for parameters with winning and losing fitness, and observed a trend towad higher co-occurence entropies in the selected distributions.

To verify whether the parameter sets generated smoothly evolving distance or perhaps only random points which happened to match the exemplars, I generated a larger family for the curves and plotted the distributions. Typical distributions and co-occurence entropies for some curves are shown in the following plots.

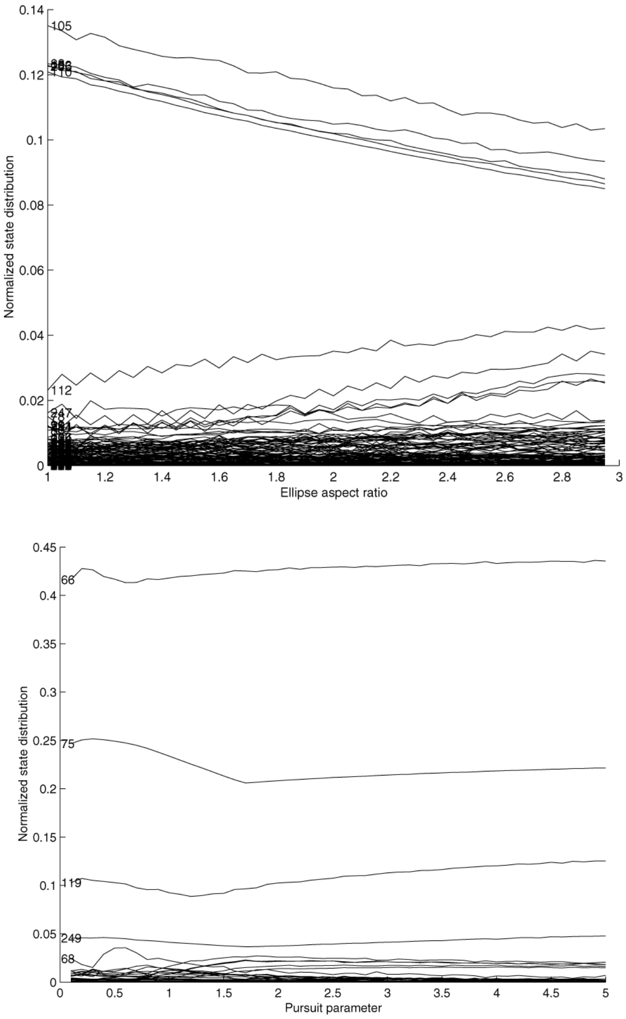
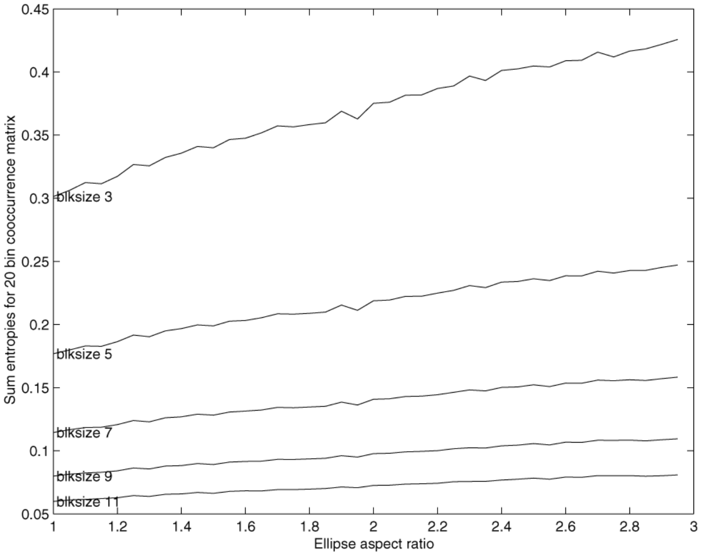
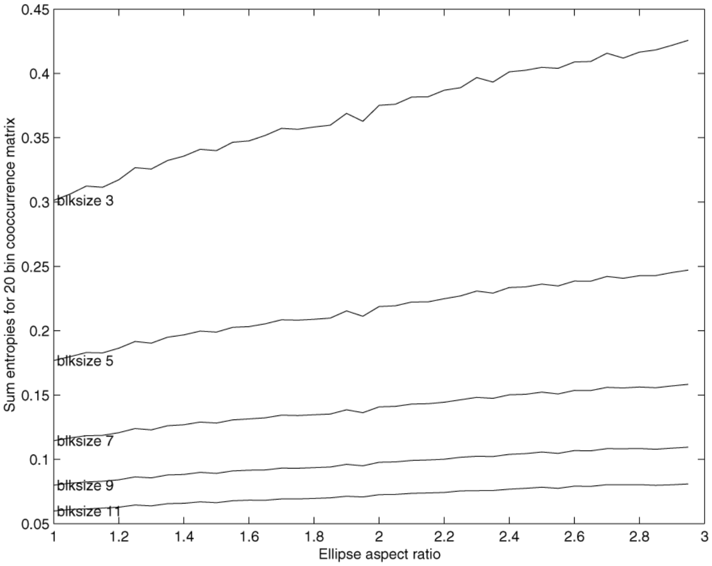


Fig. 37. Normalized state distributions from successful parameter sets on rank order task are plotted against the free parameter for a) the ellipse b) pursuit curve. Each plotted line is the fraction of states in a particular bin (partition cell); numbers are bin labels 1-256. Note that different bins are found for each curve, corresponding to different clusterings in the representation space. As the shapes evolve, the distribution changes might be described as a mode transfer.



a)



b)

Fig. 38. Sum of entropies plots. a) Sum of entropies of 20-bin co-occurrence matrix for increasing block size on ellipse with winning parameters; the entropy computation is as in (Del Bimbo 1999), but scaled by 1/(pixels/block). b) Sum of entropies for the *worst* score on the ellipse ordering problem in the first generation.

As noted earlier, the two processing stages could be viewed as a prototype of a cyclic change in bifurcation parameters based on slow wave rhythms seen in biological systems. Accordingly, I also examined whether an extended cycle (five "outer loops" of two stages) would still allow learning of parameters producing an ordering state distribution for the ellipse shape family. It was able to do so, but learning produced a different parameter set for this scenario. However, it is unclear whether cyclic changes are required or more realistic; readout could occur after the first (slow) cycle, and biological performance constraints indicate that this must be true if recurrence is used in the computation.

The *sensitivity* of the solution to changes in the bifurcation and coupling parameters was investigated for the ellipse curve. The fitness function was evaluated after perturbing each of the base parameters by the increments shown, while holding the others constant. The b1 parameter was very sensitive, with a change of +/- .0001 degrading the solution so that exemplars were incorrectly ordered when using the distributions as a distance vector. Changes in b2, c1, and c2 of +/ .0005 produced little degradation, preserving the order. The b1 parameter, typically operating in the chaotic regime, is presumably responsible for most of the entropy production, while the b2 parameter and high coupling accounts for the smoothness.

Table 6. Sensitivity analysis for the ellipse ordering.

parameter	delta	fitness (+ / -)
b1	+/- .00005	0 / 0
	+/- .0001	30 / 30
	+/- .0005	15 / 50
	+/- .0010	43 / 61
c1	+/- .00005	0 / 0
	+/- .0001	0 / 0
	+/- .0005	0 / 0
	+/- .0010	5 / 0
	+/- .0100	54 / 15
t1	+/- 1	55 / 68
b2	+/- .00005	0 / 0
	+/- .0001	0 / 0
	+/-.0005	15 / 0
	+/- .0010	5 / 0
	+/- .0100	30 / 54
c2	+/- .00005	0 / 0
	+/- .0001	0 / 0
	+/- .0005	0 / 0
	+/- .0010	0 / 40
	+/- .0100	20 / 54
t2	+/- 1	64 / 42

Discussion

Network parameters are found produce high entropy but smoothly evolving state distributions which serve to order curves in a metric space defined over partition cells of the network phase space. I interpret the *corresponding parameters* as the memory trace for the category corresponding to each curve, and the distribution produced as an intermediate computational state to be subject to some comparision to memory. The transformation of arbitrarily detailed images to partition cell bin statistics constitutes feature selection and dimension reduction. I have demonstrated the basic ability of the chaotic-periodic transient scenario to produce a state distribution with metric properties, allowing it to be used as a similarity function for curves.

I have developed procedures to compute multi-scale co-occurrence matrices, and examined the entropy statistics at different scales. State distributions found to be effective in rank ordering the members of curve families had higher entropies than the failed parameter sets.

Entropy measures are commonly used in statistical pattern recognition as a measure of the effectiveness (in terms of error minimization) of feature subsets (Chen 1973). This may be related to the discovery of high entropy solutions are found in this representation space, which also performs a dimension reduction. This insight was useful in the formation of the objective function used in the next section, where the maximization of a Shannon entropy measure is one of several components used to evaluate the representation for multiple views of an object. However, the maximization of co-occurrence entropies, as measured in this section, was deemed too costly a computation to embed in evolutionary search.

EXPERIMENT GROUP 3: RECOGNITION OF PAPERCLIP OBJECTS ROTATED IN DEPTH

The following set of experiments examines the abiliity of the network to solve the stimulus identity problem, in which different views of an object must be recognized as "similar" to form a category, even though they may differ rather radically. It would not be clear that the ability of the network to make a representation for smoothly evolving objects demonstrated in the last section could be extended to transformations of very different outlines or curves. With an appropriately chosen fitness reasonable performance is obtained on this task, even when a subset of views is shown during training. In addition to recognizing the identity of an object, I will require that this be done against a background of similar distractors.

Selection and Preprocessing of Data

The images used in this set of experiments were designed in the visual psychology lab of M. Tarr. The image set consists of 39[29] *paper clip* objects, with seven views provided for each object rotated in depth. Each object is a chain of 5 cylinders, with a variable joint angle connecting each pair. The views are separated by 30°, ranging from –90° to 90°. The objects were originally developed to answer questions regarding the *recognition by components* or *geon* theory of Biederman (Biedermann 1987). The set consists of four *complexity groups* with 0, 1, 3 or 5 unique geons substituted at some position in the chain.

Similar paperclip objects (corresponding to the low complexity set) have been used in human psychophysics experiments (Bulthoff and Edelman 1992). In these experiments, subjects were trained with motion sequences of 2-D views, giving an impression of a 3-D object through kinetic depth effects. In a two-alternative forced choice task on their object set, with single static views of a target or distractor, the miss rate (failure to indicate a match when the target was shown) averaged 30%, indicating that the task is rather difficult.

---

[29] One object in the last group was duplicated in the original set, hence 39 rather than an even 40.

The specific object set used here was also used in a study by Tarr and colleagues attempting to discriminate between view-based and structural theories (Tarr, Bulthoff et al. 1997). In this study no training period was provided; subjects simply had to judge whether two views shown briefly (200 and 100 ms, separated by a mask stimulus) were the same or different. Under these conditions, the baseline set of shapes (all tubes with no geons inserted) were essentially not recognizable by subjects when presented in other than the training views.

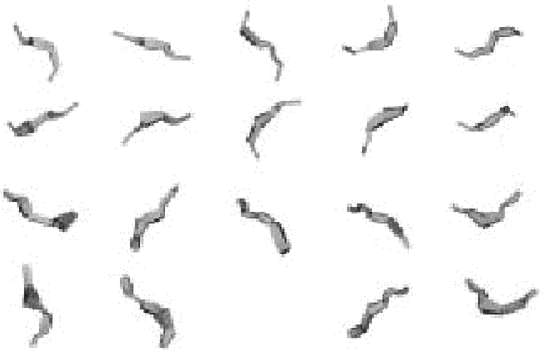


Fig. 39. Single views of a subset of the paperclip + geon objects used for recognition. The base set (not shown) are 5 tubes connected at arbitrary angles. The top two rows are 10 unique objects with 3 geons substituted for arbitrary tubes, while the bottom 2 rows have 5 geons (no tubes). The array of Soca classifiers is required to identify which object is presented, given a single view ranging from +90 to –90 from the 0 orientation shown. In a less demanding task, the network must judge whether two successive images are the same or different objects. From Tarr, M. J., Bulthoff, H. H., Zabinski, M. and Blanz, V. (1997). “To what extent do unique parts influence recognition across changes in viewpoint?” Psychological Science 8(4): 282-289. Reproduced with permission of Blackwell Publishers.

The image set was pre-processed for learning and recognition by the Soca system as follows. The images were provided as 8 bit gray scale, at 300 x 300 pixels. Each image was subsampled two times (i.e. to 75 x 75, thresholded to a binary image,then transformed into floating point as required for the CML iteration, with 0 mapped to .0001 and 1 mapped to .9999. Finally, additional padding sufficient to account for the maximum allowed iterations was added to the image border to prevent boundary effects from the Matlab CML implementation[30] from intruding on the sampled distribution of the object itself.

---

[30] A convolution routine with toroidal boundary conditions would eliminate this step.

Experiments were performed in a staged fashion. Initially, performance was assessed on a small set of objects to experiment with weights in different terms of the objective function described below.

Given a baseline learning algorithm, an experiment to assess the maximum recogntiion performance on a small set (20 objects) was performed by training with all views. The effects of several variations in the learning process were studied with this set. The parameters varied include both weights on terms of the objective function described below, and changes in the evolutionary programming parameters.

A subsequent set of experiments assessed generalization or *view interpolation* performance: what effect does providing fewer training views have on recognition performance? Finally, the same experiments were applied to the full set to get some indication of recognition performance scaling with an increasing number of objects in the visual world.

Learning: Balancing Normalization and Separation

Any effective representation must negotiate the classic clustering vs. separation dillema. In the present case, class members to be clustered are different views of the same object. To associate various views to the same object, their presentation is clustered in time in learning epochs. A representation is formed chiefly on a measure of the CML computation on a single object's views, but with some influence from the representations of objects already learned. The relative influence of these factors is controlled by empirically determined weights in an objective function.

The view normalization concept, in which a classifier network strives to transform all the presented views to a common output is here modified to apply to recurrent networks. (Since the Soca network is a single layer recurrent system, the output is identical with the state variable). Unlike the Chorus feed-forward implementation of view normalization, no particular view is chosen as the canonical one. Instead, the sum of Euclidean distances across *all* view pairs *i, ,j* is taken as the objective function to be minimized:

$$D = \sum_{ij} \sqrt{\left(\sum_{p=1}^{k} v_{i,p} - v_{j,p} \right)^{2}}$$

where v_p are the occupancies of k partition cells for each of the j views.

It is easy for the network to discover *highly synchronized* dynamics which map all views to the same representation; in fact to search more efficiently such solutions are detected, and result in termination of the fitness evaluation prior to the more costly procedure described next. The minimum bifurcation parameters were increased in an attempt to search more efficiently.

To separate categories in pattern recognition applications, it is necessary to tightly cluster the category members (different views of the same object) while separating the means of each category (Duda and Hart 1973). To accomplish this task *without* *explicitly performing the search task as part of the fitness evaluation*, a **cross-entropy** measure was maximized as part of the objective function.

Two different strategies for computing cross entropy were examined. In the first, a reference view (for the genotype under evaluation) was chosen arbitrarily for these computations and used throughout the experiments described here; during the generalization experiments when only two views were used, the previous reference –30 was changed to the –60 view.

After noting that the recognition rate for this entropy reference view often exceeded the overall rate, the strategy was changed to use a mean distribution over all views for both the genotype under evaluation and the stored signature distributions. This appeared to have little impact, with slight gains for small object world but a loss for the full set. More experiments would be required to address this question to a level of statistical significance.

Before describing the objective function in detail, it may be useful to have a look at the desired end result. The following figures illustrate the view normalization process and the effect of the cross entropy term in forcing diverse distributions.

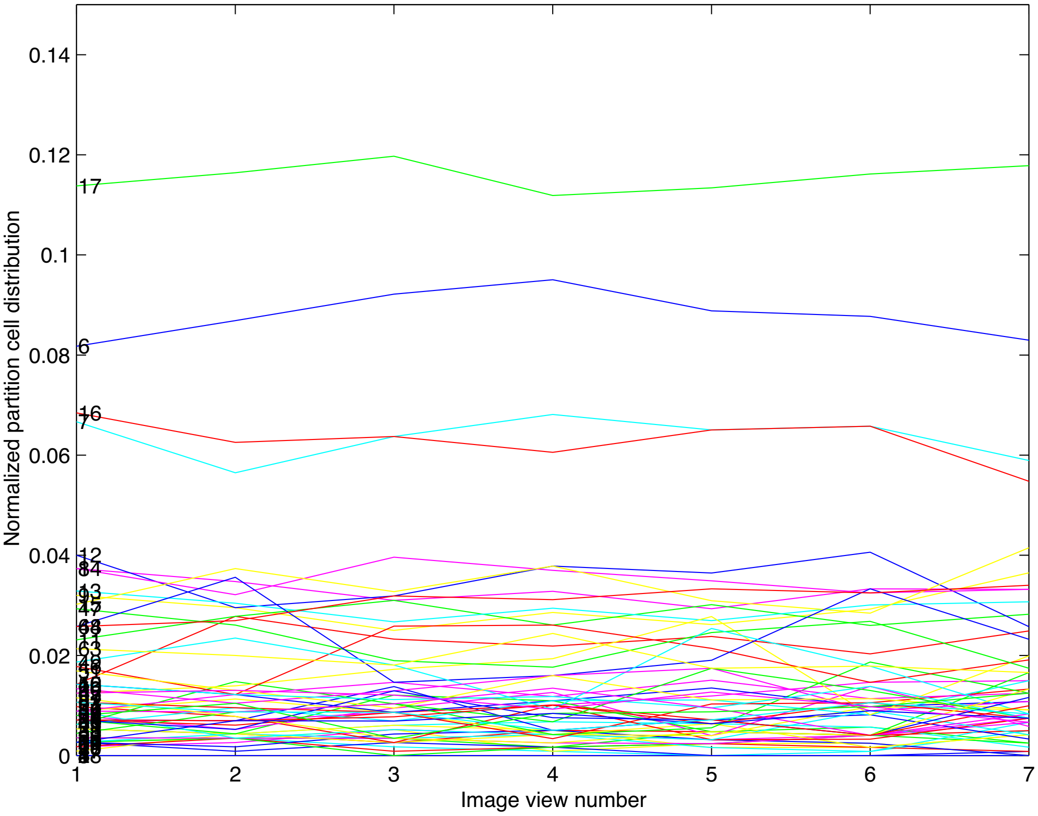


Fig. 40. Distribution of states generated by classifier across six views of an object 5.5 in the Tarr paperclip+ object set. The relatively small variance of each bin's occupancy illustrates the success of the normalization process. The labels on each line are the bin number. The mean value of each bin across all views are used as the signature for comparison during search. This distribution was generated with standard parameters W_d = 20, W_e =2, SyncThresh = .15.

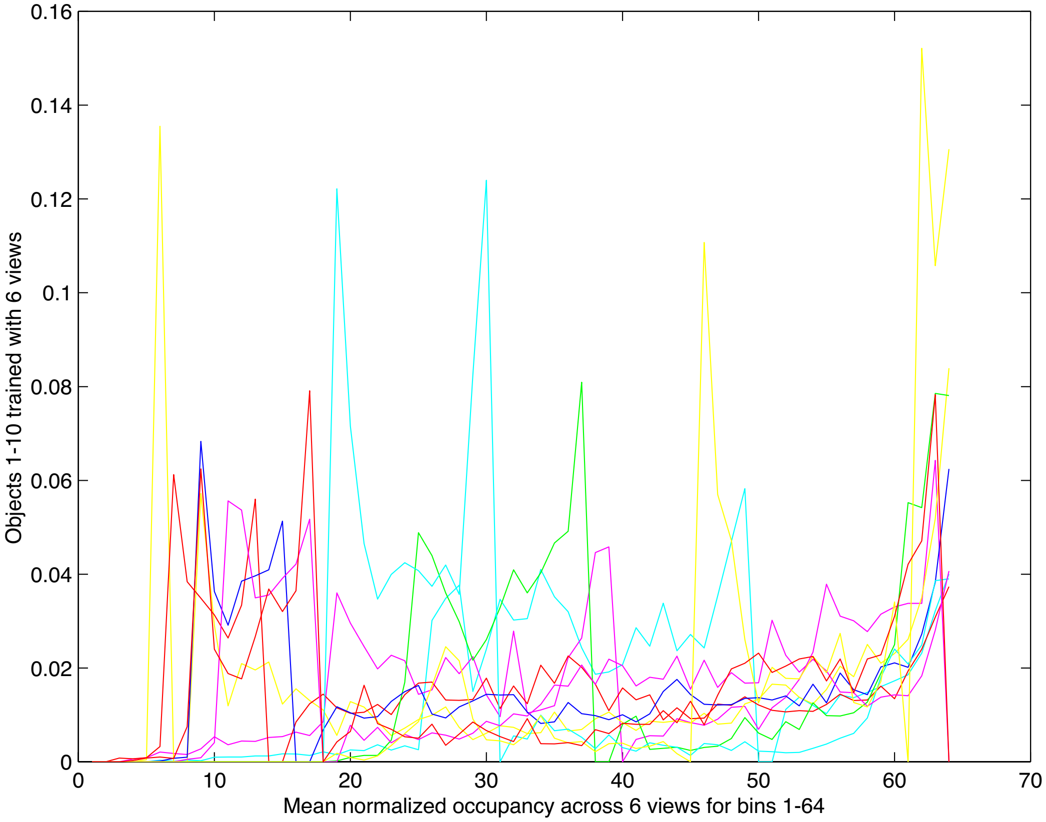


Fig. 41. The phase space dispersion effect of the cross entropy term. Each colored line is the normalized distribution across k=64 partition cells for a single object. The 10 objects comprise the low complexity (tube geons only) group, trained with the standard fitness function weights. Concentration of the population in the upper bins is evident, suggesting that many units may be approaching synchronization or a limiting distribution even in the low number of iterations allowed.

Many errors in earlier recognition experiments resulted from objects whose representation entailed relatively high synchronization, indicated by concentration of the population in one or two bins. Other objects would then tend to map closer to this distribution than to the mean mapping of the “correct” classifiers. It is easy to see that such a distribution is effective on the normalization task – it indicates that the system is approaching a limiting distribution under high coupling, a distribution which may even be uniform *for any input*. This was recognized early in the exploration of the system; to prevent this behavior, a *Shannon entropy* term rewarding broad distributions was added to the objective function.

However, adjusting the relative weights of entropy and normalization terms seemed unable to compensate for the problem. To correct for this tendency, a third term was added which penalized solutions where the maximum bin occupancy exceeded some threshold. After a few trials, summarized in the table below, a value of .15 was chosen. Any distribution whose largest bin population exceeded this threshold had a large penalty (1000) added, ensuring its rejection.

The Objective Function for Balancing Clustering and Separation of Classes

The objective function f, with low values indicating higher fitness, takes the form

$$f = W_d D - W_e(H_c + H_s) + P_{synch}$$

where W_d is the inter-view distance or normalization weight and W_e is the entropy weight; D, the inter-view distance sum for j views is

$$D = \sum_{ij} \sqrt{\left(\sum_{p=1}^{k} v_{i,p} - v_{j,p} \right)^2}$$

where v_p are the occupancies of k partition cells for each of the j views.

H_c is the cross entropy or Kullback-Lieber information measure between the current reference view distribution C and the database of N object distributions (signatures) S with k bins:

$$H_c = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} C_j \log_2 \frac{C_j}{S_{i,j}}$$

H_s , the Shannon entropy of the current signature with k bins is

$$H_s = \sum_{i=1}^{k} S_i \log_2 S_i$$

and P_{synch} = 1000 if max(S_I) > synchPenThresh, 0 otherwise. The parameter synchPenThresh was empirically determined as .15.

The following table indicates results of varying the synchronization penalty threshold on the nearest neighbor match of the 0° view.

Table 7. Error rates for alternative weights on objective function terms

W_d	We	SynchPenThresh	Error Rate %
20	2	.4	45
20	2	,25	24
20	**2**	**.15**	**15**
min	Max	.15	40

The table above is a comparison of effects of different sets of objective function parameter weight on performance of nearest neighbor match with 20 objects and training with all views. The last row indicates that rather than two terms with weights, the objective was formulated as a minimizing the *ratio* distance / entropy. The set W_d=20,W_e=2, SynchPenThresh =.15 was selected as the *standard parameter set* for training in the subsequent set of experiments.

In another investigation of learning parameters, two different mutation styles were evaluated. All results reported here apply a new random value uniformly distributed within the bounds presented earlier. However, standard practice for mutation of real valued parameters is to update values by drawing from a Gaussian distribution centered on the current value (Wright 1991). Using Gaussian mutation (variance = .02) on the CML bifurcation and coupling parameters resulted in an error rate of 29%, significantly worse than the 15% best rate obtained with uniformly distributed mutation. Presumably this is due to the highly nonlinear effect of the parameter values on network performance; however without a more extensive set of trials, possibly exploring a broader range of variances no firm conclusion can be drawn on the relative merits of the two mutation strategies.

Mutation parameters used during learning trials for this task are shown in the following table. Note that mutation is applied only to copies of the best individual from the last generation, which comprises 50% of each new generation, so the rates are effectively half of the quoted number.

Table 8. Mutation rates used for each parameter

parameter	mutation probability	Range
b1	.50	1.4-2.0
c1	.35	0.1-0.7
t1	.30	2-6
b2	.50	1.4-2.0
c2	.35	0.1-0.7
t2	. 30	2-10

Nearest Neighbor Match with Training on All Views

Once a reasonable baseline objective function was arrived at, some variations in learning parameters were examined on a set of 20 objects, consisting of 5 from each complexity group. In this set, all views of the object were used to form the representation. The representation consists of *both* the winning parameter set (i.e. the synchronization opponent genotype) for the objective function described above and the resulting mean distribution over 64 partition cells in the interval <-1,1>. The mean is taken for each cell from the 7 separate distributions computed for each view. The full set of representations, consisting of two arrays *genotype* and *binMeans*, constitute the object recognition database for a particular object world.

In a recognition test suite, each {object, view} pair is used as a target. A set of "hypothesis" distributions is computed by applying all classifiers (synchronization-opponent genotypes) to the target. The Euclidean distance of each hypothesis distribution to each corresponding database mean is computed; the object classifier whose genotype produced the minimum of all such hypothesis distances is selected as the match. Since all classifiers evaluate, performance of the best match process for N objects and corresponding classifiers is of **O**(N) on a serial machine.

Training Order

There was some concern that the best recognition performance achieved in the parameter tuning scenarios could be an artifact of training order, as all members of each 5 member group were trained in sequence, with groups listed in increasing order of complexity. An experiment was performed after scrambling the order of training; this produced an error rate of 18% vs. 15%, suggesting that training order within objects does not have a dramatic influence on the outcome. Once again, a large set of experiments would be needed to draw significant conclusions.

Single Dynamical Stage vs. Synchronization Opponent Stages

A set of learning trials was conducted with iteration time for the second stage fixed at zero, limiting the computation to a *single* dynamical stage (in contrast to the synchronization opponent system with two stages). To allow a fair comparison the single stage was allowed an equivalent maximum number of iteration as the two stage network (16). All seven views were provided for training, on the set of 20 objects. It turns out that the single stage dynamics performs surprisingly well, but with a limited number of trials (10 each for single and two stages) the average recognition rate and the average recognition time over all the classifiers were better for the two stage system. The following table summarizes this preliminary result:

Table 9. Single Stage vs. Two Stage (Soca) Dynamics (Average of 10 trials)

Trial type	Average Recognition Rate	Average Iterations / Classifier
single stage	74.7	15.3
two stages	77.9	12.7

This preliminary result is interesting in at least two ways. Based on the original intuition underlying the network, I expected worse performance from the single stage trials. The fact that the trend in performance is that the difference in scores is not so significant suggests that perhaps the simple probabilistic finite state automata recognizer explanation is adequate and the best possible explanation, without the additional gloss of the subspace synchronization argument.

The difference in average iterations per classifier is interesting, particularly since this is *not* part of the objective function. It is suggestive of a biological evolutionary scenario in which significant incremental peformance and reaction time benefits accrue from the use of non-stationary parameters.

Further study with a large number of trials is required to settle these questions. Ideally, this should be performed after experiments to assess the most effective mutation rates. Early work in progress on mutation rates suggests that the mutation rates used in the thesis produce a larger variance and overall lower mean for the two stage trial.

Random Algorithm vs. Evolutionary Learning

A learning trial was performed with random replacement of all but the best of generation during the reproduction stage, essentially a random algorithm with relaxation to a local minimum. This training procedure resulted an error rate of 29%. This validates the effectiveness of the genetic algorithm with simplex reproduction procedure, which consistently produced results in the range of 15-20% error rates.

Recognition is Not Scale Invariant

A learning trial was performed with the objects sub-sampled to 50% rather than 25% during the search process. The network parameters learned at 25% scale applied to this larger scale object resulted in an error rate of 95%, essentially chance performance.

Recognition of Untrained Views

Having settled on a canonical set of learning parameters, a series of learning trials were run with the number of training views ranging from two to seven, to assess the networks ability to generalize to unseen views. Generalization performance for worlds consisting of 20 and 39 objects, is shown for a number of views ranging from 2-7. Training view subsets started with the pair {-90, -60}, adding additional views in order {-30, 0, 30, 60, 90}. The results of this exhaustive nearest-neighbor match test suite are reported in the tables below and plotted in the figure following the tables. The third column indicates the rate of matching the view designated as the *reference view* in cross-entropy calculations during training, which was expected to be higher. This is indicated as n/a for mean-mean entropy trials.

Table 10. Nearest neighbor match recognition rates (20 objects, ref.-mean cross entropy)

training views, % correct for views from –90 to +90	mean % correct all views	% correct entropy ref.
2 c=[100 100 15 20 10 10 45]	42	100
3 c=[100 100 100 30 10 10 15]	52	100
4 c=[100 100 100 25 15 15 20]	53	100
5 c=[90 95 85 100 80 10 30]	70	85
6 c=[75 75 80 80 80 75 20]	69	80
7 c=[85 80 85 80 85 85 90]	84	85

Table 11. Nearest neighbor match recognition rates (20 objects, mean-mean cross entropy)

training views	mean % correct all views	% correct entropy ref.
2 c=[100 100 35 10 15 15 40]	45	n/a
3 c=[95 95 100 20 35 15 40]	57	n/a
4 c=[90 90 95 85 20 10 15]	58	n/a
5 c=[90 100 100 95 85 5 30]	72	n/a
6 c=[90 95 90 70 95 85 30]	79	n/a
7 c=[85 75 85 75 80 85 80]	81	n/a

Table 12. Nearest neighbor match recognition rates (39 objects, ref.-mean cross entropy).

training views	mean % correct all views	% correct entropy ref. (2)
2 c=[100 100 13 18 8 5 10]	36	100
3 c=[95 95 100 13 8 10 13]	48	95
4 c=[82 92 79 87 13 5 8]	52	92
5 c=[92 87 87 88 85 10 5]	64	87
6 c=[79 77 79 77 74 69 18]	60	77
7 c=[62 69 62 69 61 56 56]	62	69

Table 13. Nearest neighbor match recognition rates (39 objects, mean-mean cross entropy).

training views	mean % correct all views	% correct entropy ref. (2)
2 c=[100 100 15 7.7 7.7 0 13]	34	N/a
3 c=[92 92 92 13 5.1 13 13]	46	N/a
4 c=[90 97 95 90 10 13 5.1]	57	N/a
5 c=[79 82 82 84 64 13 2.6]	58	N/a
6 c=[65 67 72 72 59 56 31]	60	N/a
7 c=[62 60 59 59 69 56 51]	59	N/a

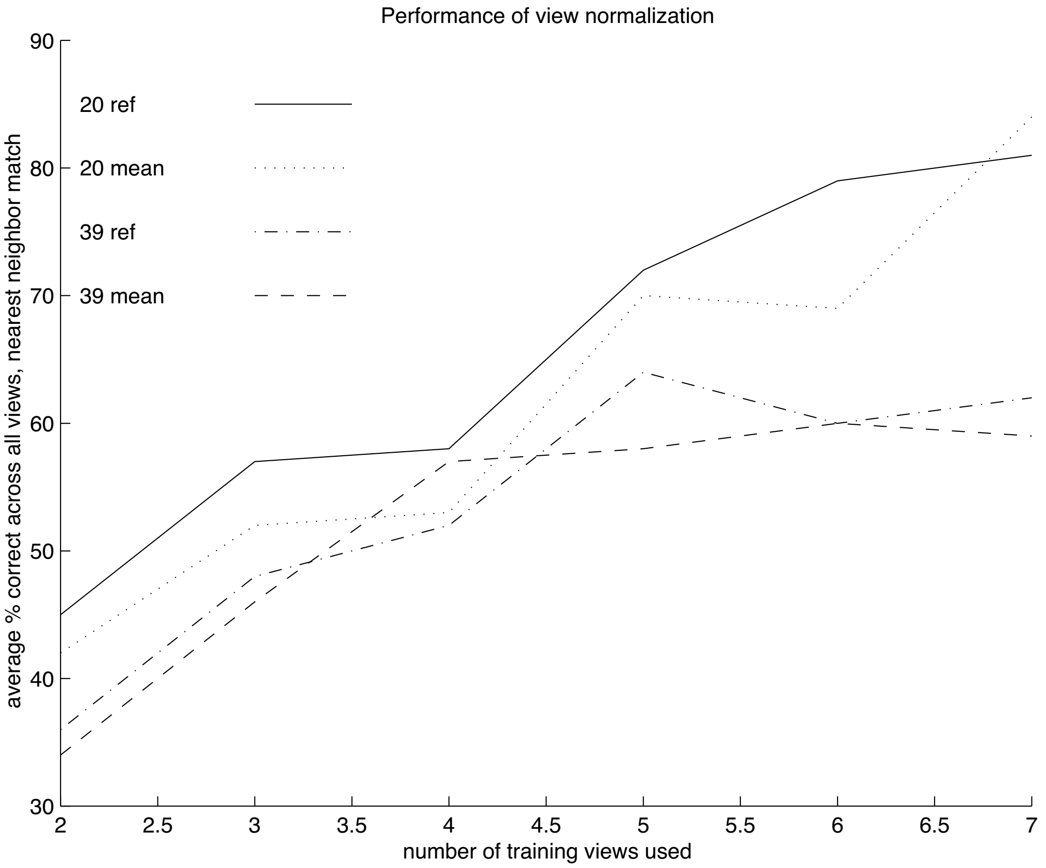


Fig. 42. Summary of nearest neighbor recognition rates with training on 2-7 views, for 20 vs. 39 object worlds, and for reference and mean cross-entropy computations .

For the 39 object world, performance improvement is minimal beyond 4 training views. The curves labeled *ref* computed the cross entropy term of all other object distributions relative to a particularly arbitrary view (-60) during training, while those labeled *mean* computed cross entropy relative to the *mean of all views the current object*. No clear advantage in seen for the mean vs. mean computation style, which I expected after noting that the view used as a reference was rarely misidentified. The mean-mean procedure improved performance slightly for the small set, but degraded it on the full set.

Category Generalization: Training with A Subset Of Objects

During the trials with all views provided for learning, I noted that erroneous matches were predominantly to objects within the same geon complexity class (i.e. tubes + 1, 3 or 5 geons), suggesting that some clustering was taking place in the representation space.

Two learning trials were performed using the standard learning parameters as above, but only using the 5 of the 10 available objects in each group (thus 20 training objects), then applying the nearest neighbor match to the remaining set of 20 objects never seen in training. For the two partitions (learning using first or second 5 of each set), 50% and 55% of the target objects returned matches from the same category. Chance expectation that an individual guess is in the correct category is .25; applying the binomial formula gives a 1% chance of achieving 50% correct assignment by random guesses. In accordance with the previous experiment, fewer errors were made for more complex objects; for the 5 geon group, all test objects were assigned to the correct class.

Given these results, the system appears to form clustering and subordinate level categorization (e.g. which “breed” of species paperclip+) on the basis of complexity. Whether humans would form such categories is less clear and is probably task dependent. Certainly alternatives which cut across complexity based categories are easy to envision, such as the categories ‘Large center with small ends’ and ‘Large ends but small center’. A more explicit examination of clustering is covered below.

Simulating Match / No Match Trials with Random Pairs

In this set of experiments, an alternative paradigm for assessing performance was used to support easier comparison with neuropsychology and psychophysics experiments. The *two-alternative forced choice paradigm* involves a human or animal subject undergoing a period of training on some set of stimuli, then undergoing a brief (sub-second) exposure to a stimulus, followed by a second stimulus. Some distractor mask signal may be used in between the two targets. The subject responds with a match/no-match choice as rapidly as possible. Error rates, in terms of false positives and negatives are computed; reaction times may also be recorded, and used as a measure of confusability between objects.

This paradigm more closely matches the testing strategies undertaken in psychological studies reviewed earlier. Since (to my knowledge) no experiments with the nearest neighbor match paradigm of the previous section have been performed, it is unclear how well monkeys or humans could perform an arbitrary nearest-neighbor match with a large number of paperclip objects, as in the previous sections simulation paradigm. Certainly for special cases of subordinate categories like faces, good performance is possible for a large number of objects.

In the simulated version of the forced choice test paradigm (with the Soca classifier system as subject), 500 trials of two views are drawn from the set of all objects and views. The nearest-neighbor function is computed as above for each view of a pair to determine whether the nearest object is judged as same or different for the pair. An error is counted for false positive or false negative matches, with the total error rate as shown in the table below.

This test scenario was performed for 20 (5 of each 4 groups) and 39 object worlds. In addition, a variant in which the *second object of each pair* is drawn only from *untrained views* is shown.

Table 14. Pair match error rates (20 objects, both drawn from trained or untrained views, 500 trials).

Training views	Error Rate %
7	2.6
6	3.8
5	4.0
4	7.6
3	9.2
2	9.2

Table 15. Pair match error rates (39 objects, both drawn from trained or untrained views, 500 trials).

Training views	Error Rate %
7	4.4
6	6.4
5	6.0
4	6.4
3	7.6
2	10.4

Table 16. Pair match error rates (20 objects, second drawn from untrained views, 500 trials).

Training views	Error Rate %
7	no untrained views
6	8.2
5	10.2
4	7.6
3	8.6
2	8.6

Table 17. Pair match error rates (39 objects, second drawn from untrained views, 500 trials).

Training views	Error Rate %
7	no untrained views
6	4.8
5	4.6
4	3.4
3	3.6
2	4.0

Discussion

The biggest surprise in the pair matching errors is the relatively low error rates in the last set (Table 13), with the first object drawn from trained views and the second object drawn from untrained views. This *increased* the errors when only 20 objects were used made performance nearly independent of the number of training views. Closer examination of the experimental conditions and combinatorics suggest that the explanation lies in the fact that for the alternative condition (i.e. first object drawn from trained or untrained views), the probability of choosing similar viewing angles of different object is higher; in the apparently anomalous condition, the same viewing angles are never compared. The number of “false positives” (judgements that different objects are the same) accounts for most of the higher error rate seen in table 17, when few objects but greater likeliood of similar viewing angles are given. Thus I believe the combination of more (and more complex) objects in the world and non-intersection of viewing angles explains the low error rates in table 18; views which map to the wrong object are simply rare under this condition.

Some degradation of performance on the larger test set was expected, but the amount seen here suggests that some action should be taken. Several modifications which might help in scaling up to a larger object "world" are apparent. Increasing the number of bins (the number of dimensions in the representation space) might help increase the separation between object categories. Increasing the weight of the entropy term relative to normalization might also improve the scaling; the current values were chosen after limited experimentation with the 20 object set. It is possible that the ratio strategy, combined with a constraint on allowable normalization error would be superior. Alternatively, local density in the representation space (Krumhansl 1978) could be used to adjust weights during learning, recognition, or both to reduce false matches.

It may be that the biggest gain in performance from scaling up would be to simply improve the learning process, in particular the normalization or inter-view distance term. Analysis of the errors indicate that the views which are misjudged are typically outliers in a family of similar distributions for the rest of the views, but the system is unable to find a better solution. It is unclear whether this is dynamically not feasible, or due to inefficient learning. Relatively little effort has been spent to date on in the evolutionary search component of the system; some possible improvements are discussed in the last chapter.

More elaborate (and computationally intensive) measures could include

1. using a multi-scale lattice. To implement this would necessitate a higher performance implementation than the current Matlab system, which essentially precludes the use of wider convolution kernels; even with the nearest neighbor kernels, 85% of the computation time in training or search is spent in the diffusion step.
2. using some dynamical process (apart from the localized diffusion currently implemented) to distinguish the location of parts in space.
3. training multiple Soca style classifiers per object and changing the matching process to a majority vote of the top n responses, where n = the number of classifiers per object. This assumes that parameter sets which result in good normalization performance are dense in the parameter space. While there is currently no proof that this is so, some of the experimental results (multiple solutions to the ellipse ordering problem, relative independence of the quality of matching results to the order of objects during learning) suggest that a multiplicity of solutions to particular representation problems exist.

Explaining the Results

The functioning of the Soca network on the view-based recognition by normalization task can be explained on several levels. Clearly if a representation space over dynamical partition cells can be created with perfect normalization and the category boundaries well separated, the system is accounted for in the sense of its input / output performance. In the previous chapter on representation, I proposed a deeper explanation in terms of statistical language recognizers with a state flow resembling a probabilistic finite state automata, but in reality more precise in its discrimination ability.

Partition Cells are Not Spatially Consistent Across Views

Another level of explanation seemed required to assess what the system is "paying attention to" across views and across objects. .To address this question, plots were made of the spatial location of the highest population bins (normalized fraction > .03) across all seven views of an object (paperclip + 5.5). From this and similar plots, it appears that in the course of performing normalization (i.e. maintaining roughly the same distribution across views), there is little correspondence or registration of the partition cell components with particular "parts" or features across views. Instead, the network is able to find dynamics whose underlying state-flow graph reaches the same distribution for the particular starting distribution of initial "words", as they are arranged around the contour in overlapping windows.

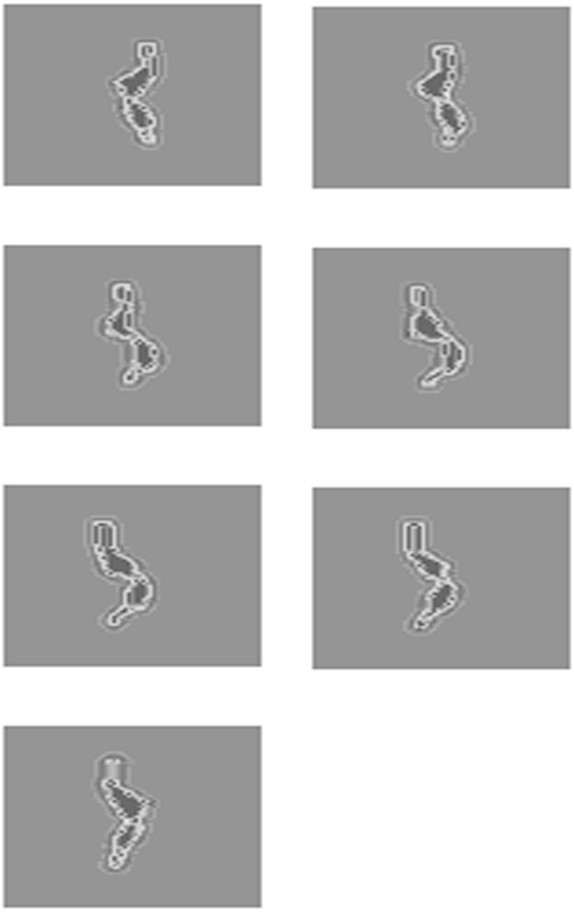


Fig. 43. The spatial locations of high population bins (> .03% of population) of paperclip object 5.5 are shown as a white overlay on the seven views of object 5.5 . Since the maximum occupancy of any partition cell is limited to .15 by the objective function, the light areas must be significant in mapping the object. Note that the entire contour seems to be represented, rather than any particular feature; based on a more detailed examination of locations for different thresholds (which can pick out the highest bin, etc) , there appears to be little consistency of the location of individual bins with particular features across views.

Visualizing Clusters by Multidimensional Scaling of the Representation Space

A second approach to understanding the dynamics was to examine the clustering of the objects in representation space. The following set of images shows the projection in two dimensions, obtained via multi-dimensional scaling (MDS), applied to a distance matrix obtained from the 64 bin representation vectors[31]. The constraints used in the formation of the representation space – minimizing differences across views and maximizing sum of entropies between classifiers – appear to result in a psychologically reasonable clustering and separation, with a few exceptions. No psychophysical testing has been performed, so at this point the reader must judge this by inspection. The requirement for normalization essentially forces the emergence of subordinate level clusters, as the dynamics is simply incapable of equally spacing representations for individual objects in the space under that constraint.

The cluster labeled 1 in the first consists of objects with few extra geons, minimal protrusions; those in cluster 2 have distinctive protruding geons, with several in the center of the chain. Within these clusters (and to some extent across clusters) the pattern appears to be something like the turning angle measure, or oriented run length. There are certainly instances in which I would cluster things differently given free choice to group the most similar object to any target, the paradigm used in the study of (Scasseleti, Alexopoulos et al. 1994). Performing a match – no match task stressing a pre-attentive response may give different results.

---

[31] A measure "stress" indicates the goodness of fit to a particular dimension in the MDS procedure; the stress for mapping to 2 dimensions was 3.4, which is higher than the value 2 which is often taken as a rule of thumb for a good fit. Stress under two was obtained for 3 dimensions.

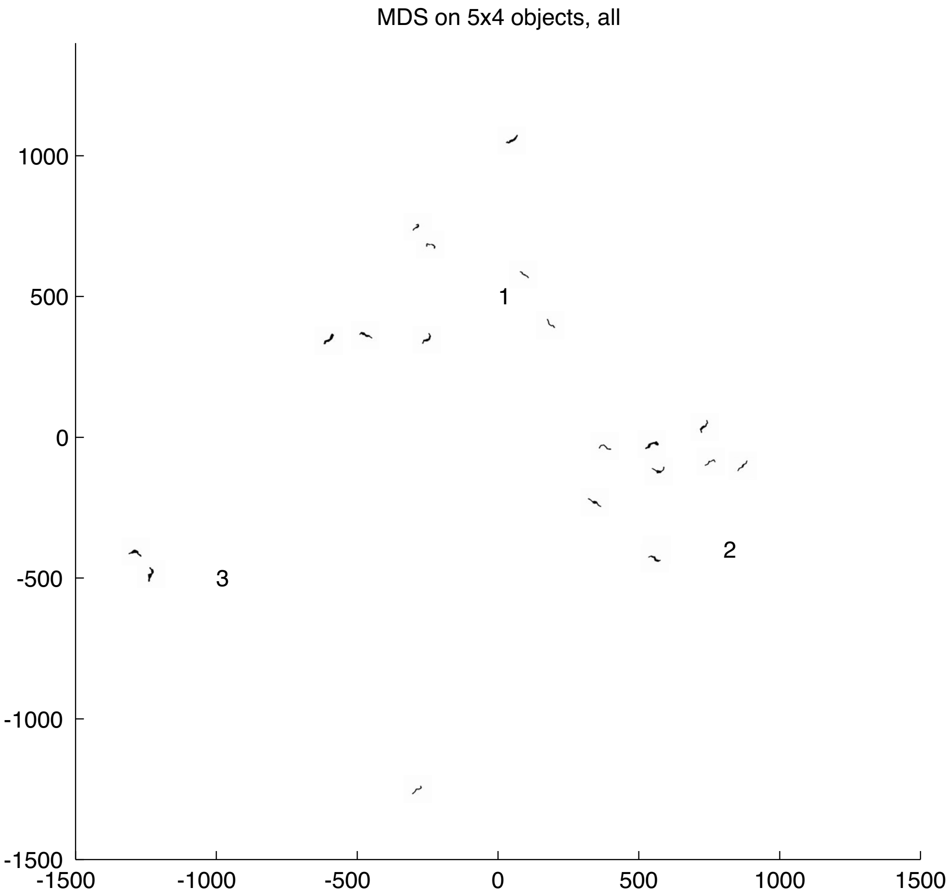
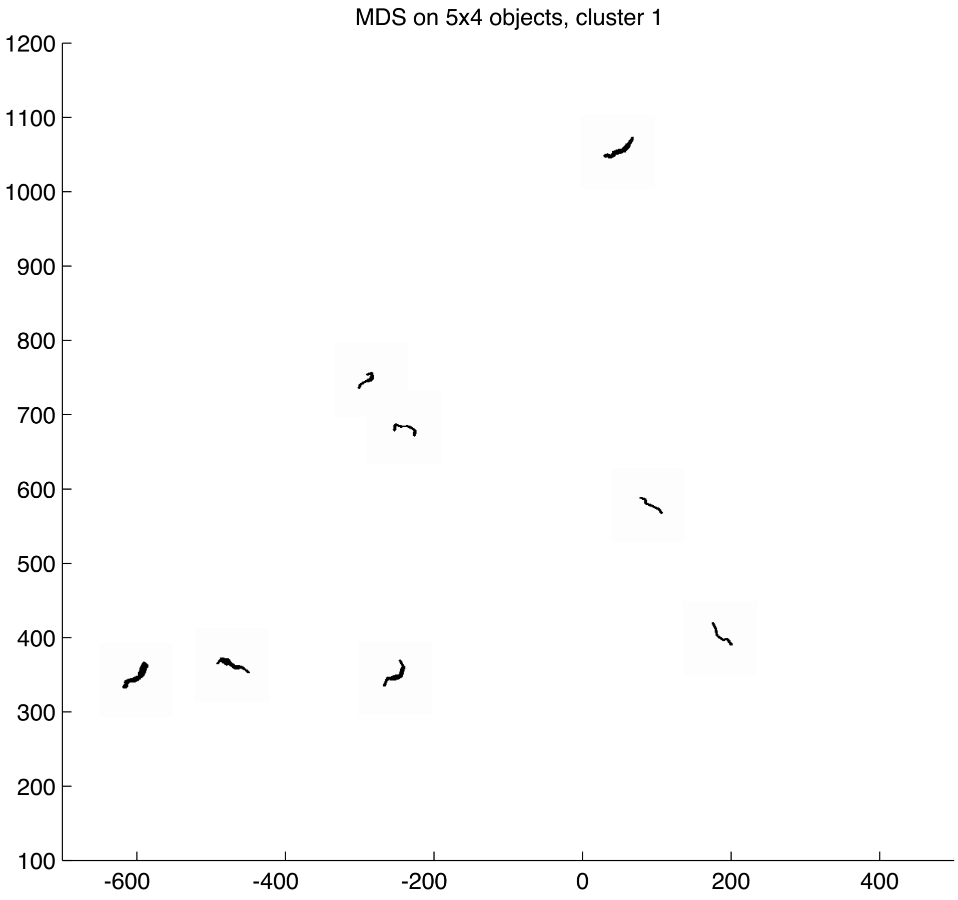


Fig. 44. Projection of distances into two dimensions via multi-dimensional scaling on inter-object distance matrix, with distributions computed for view invariance. Three clusters are seen; clusters one and two are enlarged below.


MDS on 5x4 objects, cluster 1
1200
1100
1000
900
800
700
600
500
400
300
200
100
-600
-400
-200
0
200
400

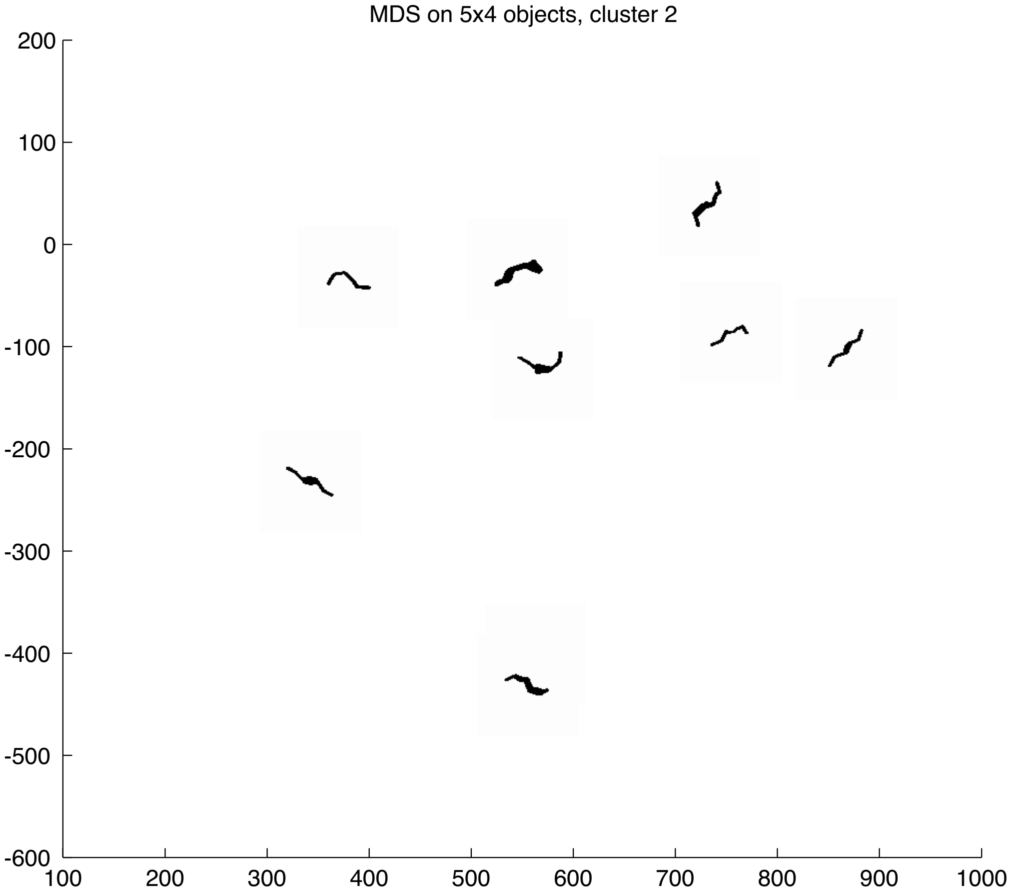


Fig. 45. Zooming in on two clusters in the representation space. The axes are arbitrary distance units determined from inter pair distances in a unit cube.

Comparing The Soca Classifier Approach With Other Systems

Exact comparison with the other systems reviewed earlier is difficult, as the training circumstances reported in each varied slightly. On the whole, the raw information presented for training to the Soca system is less than that provided for Chorus and SEEMORE, consisting of fewer, widely separated views and no shading or color data. The other systems have “front ends” consisting of more than one hundred local feature detectors. Soca, in contrast has no feature front end, only the asymmetric, diagonal-suppressed coupling. We might consider the simple threshold based conversion from shaded color to binary as a primitive front end.

I assume, of course that the non-trivial task of segmentation to produce a primal-sketch style image has been done by previous computational stages. All three systems have only been tested with isolated objects.

Error Rate Measures and Testing Conditions

The Chorus system (Duvdevani-Bar and Edelman 1999); (Edelman 1999), an ensemble of radial basis function (RBF) classifiers, was trained on a visual world consisting of 10 shaded objects readily assigned to categories such as quadrupeds, airplanes and cars. Test views of the objects, separated by 10° increments, were presented in a range of +/- 60° in azimuth and elevation, for a total of 16-17 views for each objects. Of several tests reported, the closest comparison to the present work is the recognition of untrained views of the 10 objects using a winner-take-all algorithm over individual RBF classifiers. In this case, the error rate for Chorus is 10%. Use of an additional competitive layer reduces the error to 7%.

This performance slightly exceeds that of the Soca system, but differences in complexity and information content of the raw data, view separation, and look ahead over the entire object set prior training make judgements about the superiority of either as a computer vision method difficult to assess without further study. Soca consistently exhibited its best performance on more complex objects in its training set, thus might perform better overall in a more complex object set; the effect on performance of providing more closely spaced views is unknown. Testing of Chorus on a set of objects with limited range like the paperclip set has not been performed.

The views selected for training the Chorus classifiers were chosen by an algorithm with access to all views of all objects in order to separate the clusters. No such global view was made available to Soca, which formed representations based on cross-entropy functions of objects representations already formed in an attempt to solve the same problem.

Error rates for a nearest neighbor match performance assessment of Mel's SEEMORE system using only the 79 shape channels alone is reported at 21.3% This rate was based on 12 rotation in depth views separated by around 60°; additional views for scaling were provided. Test target views *excluded* highly foreshortened views. The visual world in SEEMORE consisted of 100 images, with three scales provided for each training view, resulting in 3600 views total[32] if all objects were rigid. It is difficult to project how the performance of Soca or Chorus approaches would scale to that size.

Considering that Soca used more widely separated views, foreshortened views, no shading data, and no selection of preferred training views based on overall world statistics, the performance is nevertheless comparable to the other algorithms. The training procedure is rather arbitrarily truncated, and explores at most 3000 genotypes per classifier; more thorough exploration of the parameter space might improve the process of normalization across views. One could argue that the paperclip discrimination test is

---

[32] 3600 would be the number if all objects were rigid; additional views were provided for nonrigid objects.

the most difficult, given that human error rates range as high as chance for low complexity (tube only) geons with exposure to only single views (Tarr, Bulthoff et al. 1997). Clearly humans are already trained on the everyday objects used in Soca and SEEMORE, and it would be surprising to see substantial error rates for humans in match/no match tasks.

Memory Utilization

For each object represented in the Soca system, a vector of six parameters are stored specifying the dynamics for the two synchronization opponent stages. In addition, a signature corresponding to the means across all views of the sampled distribution after the specified number of iterations is stored. In the Matlab implementation, all floating point values are double precision; thus 48 bytes are used for the parameter set (genotype) and 512 bytes (64 bins x 8 bytes) for the signature. It seems unlikely that performance would suffer greatly by storing single or even lower precision values, since the signatures are not used in iterative calculations. As noted in the learning and representation selection, no precision exceeding $\log_2 \frac{1}{N}$, where N is the number of lattice units, is useful, since the occupancy of a phase space partition cell cannot be less than this unless it is zero.

A table below summarizes the relative memory utilization of these shape recognition systems. Note that the *minimum* error rate for Chorus and SEEMORE is not reported as more training views were available for Chorus and non-shape data was available to SEEMORE. The Chorus rate reflects fewer views (and less memory) to give an error rate comparable to Soca for memory comparison, while the SEEMORE error rate reflects performance and corresponding the memory used for the shape channels only. Both reports featured sufficient information on perfomance with limits on training to partially compensate for these differences in the following comparison.

Memory representations for Chorus involve specification of radial basis function centers, widths, and weights for each node. One center per view is typically used. Error rates depend on the number of views provided; to achieve an error rate of 15%, the data reported for Chorus suggest 5 views would be required, resulting in storage requirements of 2400 / bytes object [33] with double floating point representation of network parameters. For SEEMORE, the appropriate comparison (using only shape channels, representing only rigid objects, and not providing extra scaled versions during training) is 1264 bytes / object.

---

[33] 5 views x 3 (center,width,weight) x 8 bytes/double x 200 receptive field activations

Table 18. Memory utilization and error rates for Chorus, SEEMORE, Soca

System	Memory Bytes/Object	Error Rate %	Objects	Views per object
Chorus	2400	.15*	10	5
SEEMORE	632	.21	100	36
Soca	560	.15	20	7

A NOTE ON RECOGNITION PERFORMANCE

Performance of the recognition algorithm is a O(N), where N is the number of objects. Specifically, it is N*(M+D), where M is the computation time for the CML on the given view for t1+t2 iterations, and D is the time for the distance computation. This is dominated by the CML computation, and profiling shows that 85% of the time is spent in the averaging convolution function *filt2d*. Using 82x82 image arrays for the experiment reported in the next chapters , the Matlab implementation on a 400 MHz PPC 750 (1M cache, 100MHz system bus) resulted in a maximum CML computation time of 0.18 sec.

Chapter 7: Discussion and Summary

It takes two to invent anything. The one makes up combinations; the other chooses, recognizes what he wishes and what is important to him in the mass of things which the other has imparted.
Paul Valéry

The method presented in the previous section was motivated by work in various disciplines. I now turn from details of implementations and experimental results to consider the limitations of this study and opportunities for improvement from the perspective of each discipline, followed by a summary of the relevance of this work for the field and the research openings suggested by the findings here and in the corresponding work surveyed.

LEARNING ALGORITHM AND REPRESENTATION

Several approaches can be considered which might improve the recognition by more effective learning, without changing any other network parameters. One strategy has been designated as a cooperative coevolutionary approach for general optimization (Potter and De Jong 1994) or symbiotic adaptive neuro-evolution (Moriarty and Miikkulainen 1996). This involves maintaining separate populations for subsets of the parameters of some system, and evaluating the fitness based on how randomly chosen individuals from the two pools cooperate to solve the problem. For the Soca network, the two stages could be mapped to different subpopulations, with evaluation based on a random combination of genotypes for each stage. These methods have been demonstrated to improve performance relative to single population approaches. However, given the high degree of interaction between network parameters here and the relatively low number of parameters, the potential gain is unclear.

Incremental learning (Gomez and Miikkulainen 1997) is another approach that promises improvement. Rather than optimize for a complex task from the beginning, the task may be incrementally modified with increasing complexity and an incremental form of learning used to modify the solution. One way to adapt the object recognition task would be to train first for synchronization across views with an entropy constraint, then add the maximum-cross-entropy, then finally test evolve performance on the actual recognition task between objects which generated failures after the second stage.

Using ensembles of classifiers can improve performance, when the errors made by the classifiers are independent (Hansen and Salamon 1990). The use of the cross entropy criterion could be extended to this application by attempting to form new classifiers for the same object which maximize cross entropy from the previous classifiers for the same object.

Alternatively, one might consider abandoning the evolutionary learning paradigm altogether and attempt a more direct control method. The types of optimization criteria employed here (normalization, entropy and cross-entropy measures, and synchronization limits) should be utilized as architectural principles guiding the control design.

I turn now to the current representation strategy. It may be possible that changing the partition size to finer granularity or possibly using non-uniform intervals (bin widths) to match the density of state occupancy in certain regions could increase the recognition performance.

Whether there are inherent benefits or liabilities of the partition cell coding strategy has not been a major focus of this work. Instead, the emphasis has been on demonstrating a column-like assembly with complex, stimulus-driven temporal patterns (in line with recent trends in sensory neurophysiology of the temporal areas) that can serve as the algorithmic and mechanical support for creating population coded metric spaces and normalization in view based recognition.

Benefits may exist; it would be interesting to compare *the ease of learning* this statistical, partition-cell based representation with localized representations, given an equivalent learning procedure such as the genetic or evolutionary learning strategy used here. Apart from the time to find a solution in some learning or search procedure, the density of feasible solutions could also be important. This is particularly true as the number of objects increases and collisions - different objects mapping to the same point in the representation space – becomes an issue. Having many solutions to a task available should make balancing of conflicting objectives easier, and could support more computationally intensive but higher performance classification by creating multiple classifiers per object, with a voting criteria or other resolution mechanism. This remains an area for future work.

SOCA NETWORK CLASSIFIERS AS A COMPUTER VISION TECHNIQUE

If viewed from the perspective of practical computer vision and visual psychology, the study has some limitations. I will note these, together with directions I envision to bring the model into closer correspondence with psychophysical performance levels.

Isolated Common First Order Statistics

The system is currently limited in scale and context sensitivity during processing to a spatial window around each pixel, within which it receives information via spatial diffusion. It is easy to present collections of lines or contours separated by this window distance which, since their "wavefronts" will not interact, will be indistinguishable; see the following figure for an example. At least two possible improvements are foreseen to handle this situation. One would be to directly address multiple scales by introducing a pyramid of CMLs, operating on subsampled binary images or other representations. Another possibility is to relax the current assumption of homogeneity in bifurcation and coupling over the lattice. Rather than choosing exact values, the system could be trained to employ a range of values to be spanned over the space of the lattice.

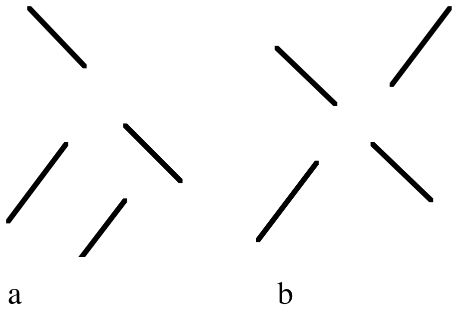


Fig. 46. Two collection of lines which cannot be distinguished by the present architecture. Blurring, pyramid structures, and spatial gradations of control parameters might allow these to be distinguished, but whether learning will perform comparably under those conditions is unknown.

Handling Multiple Scales

As noted earlier, performance with objects subsampled once rather than twice was very poor. It is possible that this could be compensated for by training at multiple scales (the approach taken by Mel with SEEMORE). Alternatively multiple scales, could be built into the dynamics in a number of ways. A pyramid structure could operate in parallel on the data, with various scales uncoupled or possibly coupled.

Ecological Realism and Learning while Behaving

The system is currently limited, in terms of ecological realism, by the separation of learning and recognition epochs. Learning is also assisted by imposing clear temporal boundaries on the presentation of a single object. The extension of such a system to an ecologically realistic environment where learning and recognition occur together in the context of tasks would entail additional control dynamics, the nature of which are not entirely obvious.

Limited Range of Objects

The range of objects studied here is intentionally restricted. In psychological terms the problem addressed is chiefly subordinate categorization, and the refinement of that problem to stimulus identity from multiple viewpoints. Conceivably, basic level categorization arises naturally from the constraints imposed by view based interpolations on the representation space and mappings, as Edelman contends.

For future development of the system it would be desirable to test on one or more standard vision benchmarks. While quasi-standard benchmark sets exist for certain applications (faces, fingerprints, textures), a widely recognized object library is still lacking. One possibility is the library assembled by (Verfaillie and L. 1995), which has the advantages of some existing psychophysical data on view preferences, reaction time trends, and freely available. Human similarity preferences also exist for the large silhouette library used in the IBM QBIC project (Scasseleti, Alexopoulos et al. 1994), but different views are not provided for each object, limiting the direct applicability as a training corpus. One could use portions of the human similarity clusters for each object to train, then assess performance on others.

From Isolated Objects to Scene Analysis

The current limitation to recognition of isolated objects, while common in prototype object recognition systems, is not realistic. Humans can clearly recognize conjunctions of objects, or perform matching tasks with multiple objects in a scene.

As reviewed in chapter 4, many investigators have begun to address the problem of segmentation within the general framework of coupled oscillators. To build a large scale system for scene analysis which combines segmentation, recognition, and search appears to be a feasible goal. The combination of search and segmentation together is of particular interest in a practical system and should be pursued in future work.

As an initial step, one might address the search issue by continuing to learn object representations by isolated presentation, but require matching to occur in the presence of multiple isolated targets.

A more sophisticated recognition of objects imbedded in scenes would require an extension to active vision, a major research area which has not been addressed in the present work. Some sort of scanning behavior, simulating the saccade control and foveation process is envisioned. In earlier unpublished work, I proposed and simulated a dynamical algorithm for choosing stimulus-determined saccade points based on selecting image points which differed from their surround when sampled late in a Soca-like desynchronization - synchronization cycle (DeMaris 1995). These points were typically found adjacent to areas of high information content (e.g. on orthogonal axes near curvature maxima).

To address issues of ecological realism, in particular the fluid nature of mixing learning and recognition in the service of goal-directed behavior, further study of large coordination dynamics is required, combining evidence from EEG/MEG and more micro-level recordings on the same task. Analyzing experimental outcomes in terms of the framework sketched in this thesis for inter-region coordination may prove fruitful. For example, if the separate dynamical stages of the Soca processing cycle play a role in shaping of distributions for memory formation or comparison during search, a mechanism and its dynamical signatures must be identified for initiating and terminating such cycles.

Different possibilities are envisioned. One is to employ intrinsic rhythms, serving as perceptual frames. Alternatively, and perhaps more plausibly, is that behavioral events trigger the initiation of cycles. These events might include actions such as terminations of saccades, or cessation of head movements during gaze orientation.

Improving Recognition Rates on Larger Object Worlds

Before expanding the scope of the problem to segmentation and scene analysis, it would be appropriate to first address the issue of scaling up to larger problems. Performance of the system degraded substantially when the number of objects in the world was increased from 20 to 40. This degradation was more severe in the nearest neighbor recognition of a prototype, less so in the match-to-sample style task. It is possible that identifying an exact match from a large set of similar objects is so difficult that human behavior would degrade similarly, due to increases in the number of indistinguishable foreshortened views or views with similar statistics at all scales. There is a good deal of evidence that for faces, a category where we have impressive discrimination capabilities among large numbers of objects, specialized subsystems are involved.

Several promising strategies are open to exploration. Most of these would entail substantially increased computing burdens, suggesting that a rewrite of at least the core convolution and mapping loops be performed. Given the low iteration counts, roundoff should not be a serious issue; the use of single precision floating point should be explored and should lead to considerable speedup.

The present system is suitable to parallel processing at the level of parallel evaluation of network genotypes in learning and recognition epochs. The local nature of computations in the core algorithms lend themselves to parallel hardware implementations or pipelined computations. Some of the possible algorithmic improvements proposed above would impact the achievable density. Allowing inhomogeneous bifurcation or coupling parameters would entail additional registers or sequencing parameters to computation nodes, for example.

Use of More Complex, Larger and Multi-Scaled Coupling Kernels

It is likely that using a larger, more complex coupling neighborhood (convolution kernel), more partition cells, and possibly non-uniform intervals adapted to the dynamics and the object world would improve recognition performance to some extent, but with a tradeoff in computation time. Of course, the limitation to a single orientation is an artificial limitation that could be removed.

Using larger coupling kernels has potential to improve the performance, based on some preliminary results in an early experiment seeking a universal classifier (i.e. one which did not develop a classifier based on multiple presentations of an object). While the overall results were poor for scaled and objects rotated in the plane, better results were obtained with larger coupling matrices with specific orientation kernels.

It would be possible to make the coupling kernels part of an evolutionary learning strategy. I have mentioned the possibility of multiple scales of analysis of the image previously; it is likely that particular objects would require different weighting of statistics at particular scales, or coupling *between* scales, to handle situations when local or global information is more important. For example, with non-rigid objects we might wish to weight the response of local information more heavily, or allow it to have more influence on a common, cross-scale determined distribution by asymmetric coupling.

Use of All or Portions of the Ensemble Time Series

The strategy of using lattice state values sampled at a single instant was chiefly chosen for simplicity, and a low computation and storage burden. It is possible that improved performance could be achieved by examining the entire N channel time series or some measure of this dataset. Additional measures might serve as witnesses casting votes for one or another object when the standard distance measure fails to produce a clear match. For example, if the mapping of an object is close to two or more signatures, a vector of rate averages might discriminate. These additional measures would have to be saved as part of each object (or prototype's) signature, of course.

In fact, since the learning procedure currently does not perform actual comparisons, there is really no proof that the optimal discrimination performance for the given set is obtained at time step (t1 + t2). It would be interesting to examine recognition rates at earlier times. It might prove useful to consider vectors at the last two steps as predictors, and to adjust the objective function to measure and score these.

Improving Search Performance with Indexing

If adequate performance on a large and diverse object set were obtained, it would be important to address efficiency in the *search process* as well. While dynamical methods might exist, their nature remains a topic for future research. In the short term, an *indexing* strategy could be used to speed up the search process.

For example, I have shown that the application of multi-dimensional scaling to the database of mean distributions over the partition cells results in a clustering of cells.

The following steps would improve recognition performance, with speedup depending on the number of clusters obtained. The basic idea is to evaluate (i.e. execute the CML dynamics) only those prototypes near to the best performing clusters:

```
*Perform clustering via Multidimensional scaling or some other procedure.*
*Choose a classifier with a signature nearest the center of each cluster.*
*Evaluate the phenotype network for each cluster center.*
*For the classifier with minimum distance from Soca(target) to signature*
  *Evaluate all cluster members.*
```

If the occupation of the representation space is not distinctly clustered, another technique might allow screening of distance computations after the CML evaluation. A low resolution version of the signature histogram (the partition cell occupancies) can be computed and stored. For example, the current implementation used 64 bins; construct a 32 bit string by choosing odd bins and setting the corresponding bits 1 if the occupancy exceeds half the value, 0 otherwise. (Signatures with bins near the threshold could be scheduled for full evaluation to avoid false negatives). A simple bit comparison will exclude objects with rather different signatures. More sophisticated variants of this approach have been developed for comparison of similarity in chemical structures (Willet, Barnard et al. 1998); (Flower 1998).

Such a technique could be of even greater importance if all or portions of the lattice time series were used as a signature, rather than an instantaneous sampling of the lattice as was performed in this thesis.

Noise and Occlusion

As mentioned earlier, scale space methods typically avoid *creating new spatial structure*, while the procedure described here is completely dependent on fine structure and cooperative interactions derived from those structures. This suggests that the process would be noise sensitive, and to date I have not performed any experiments by degrading the initial binary images with noise. One factor likely to mitigate the effects of both noise and occlusion is the collective measurement process.

Learning from Single Views

It is clear that humans (and monkeys) can learn to recognize exact views and some limited transformations after a single exposure. How can the current framework account for this?

One possibility is that ensemble networks serve as a back end allowing interpolation in the space spanned by a limited number of existing prototypes. This is essentially the approach taken by Chorus. Single shot learning gets a free ride on prior learning of objects seen from multiple views.

Another possibility would be to *generate multiple views from a single view* by sampling different subregions of the image, applying different morphological or disortion operators which approximate viewing transformations, or other procedures.

SOCA NETWORKS AS A BIOLOGICAL THEORY

The work here has adopted the view based recognition paradigm for recognition of objects, but differs in assumptions of the underlying neural mechanics. Admittedly, the existing framework is completely hybrid in character and has a long way to go to be a complete biological theory. Yet there are intriguing correspondences between the nature of the computation here and aspects of biological computation. In summary, the major distinguishing aspects of the Soca network approach from classical connectionist models are as follows:

1. The neuronal group level is taken as the basic functional unit, with macrostate variables representing the ongoing ensemble state. Computations and representations are formed by collective measurements (distributions) over spatially uniform ensembles of these units.
2. The typical connectionist unit transfer functions (sigmoidal activation and threshold output) is replaced with a nonmonotonic, highly nonlinear (chaotic) function. There is no threshold, since the state variable represents some collective measure such as ensemble frequency or phase distributions. Connections between groups are not excitatory or inhibitory, but serve to construct intrinsic state flows in the space of the recurrent network dynamics; these can be tailored to the statistics of input configurations. In Marr's hierarchy of processing levels, the algorithmic level involves spatial cooperative processes in these state flows.
3. The code for family of input patterns (i.e. outline shapes) is a sample of the states in the entire network at a specific point in time, rather than activation of an optimal unit (place codes), a connectionist sparse or distributed output layer, or a recurrent network in a static (fixed point) attractor. The code is computed via synchronization processes operating on stimulus-linked aperiodic oscillations with local cooperative interactions. Synchronization or clustering may arise in distant units with similar local input configurations, with no communication between the areas resulting in the synchronization. Rather, a common history and local cooperative dynamics results in occupying the same region of a dynamical phase space.
4. Synchronization may be considered a form of competitive processing, but differing from the "inhibition of rivals" scenarios usually seen in rate code models. Instead, through cooperative processes, particular oscillation modes occupy sites in the network. This is effectively competitive since there is selection from broadband oscillations to a sharply peaked distribution. Whether this "cooperative competition" offers advantages over the usual type requires further study, but it seems to map more readily onto a view of neurons as spike coincidence detectors and participants in spatio-temporal processing assemblies exhibiting local coherence (Bullock, Achimowicz et al. 1995).

5. This spatial cooperative processing and partial synchronization effectively performs a low-dimensional transformation of high dimensional objects (i.e. image arrays) after first projecting them into higher dimensional spaces (by *relaxing* subspace synchronization constraints created by earlier levels of the visual system). In this higher dimensional space, the spatial diffusion process generates information about the initial spatial configuration for increasing orders of local statistics, but conditioned by lower order statistics.
6. The network is time varying. These changes in the dynamics are interpreted as corresponding with major signal bands and modulations as observed in local field potential EEG and multi-unit neuron correlation studies. For the most part, connectionist accounts of cognitive processing make little contact with medium and large scale electrophysiology; rather, they focus on the level of neurons, stressing that these may stand for groups. In contrast, the present work investigates computational processes involving temporal changes in coupling (modulating synchronization) and in bifurcation parameters (modulating synchronization and the aperiodic or periodic tendencies of oscillations), motivated by observed changes in coherence, spectral sharpening at various space and time scales of biological neural systems.

While these attributes are all unusual in neural modeling which attempts to address the perceptual or psychological level, there are many grounds for criticism in terms of biological realism.

The arbitrary starting, stopping and injecting an image as an initial condition into the network is problematic. Some dynamical networks use an input stage to inducing a perturbation of an ongoing state vector, perhaps by increasing coupling parameters in a processing layer coupled to a learning and recognition layer.

A general criticism that might be leveled is the rather loose commitment to the meaning of the state variables at the lattice sites. I have generally followed the approach that they refer to ensemble average frequency (or spike density), which should correlate with the amplitude of the local field potential. This is the approach taken by Freeman, and I offered the justification for this in terms of random firing models producing a single humped curve.

On the other hand, I have held open the door that these principles could also apply to more complex spike coincidence and phase modulation networks; one of the strengths of nonlinear dynamics is the universality of the principles, so that many different underlying physical systems may exhibit similar dynamical trends.

I have claimed support from observations interpreted as temporal codes, but in fact the code used here is an *instantaneous state vector*. The distinction between code and computation producing the code has perhaps been blurred. I have implictly suggested that the observed temporal modulations may be a computation “steering” the system toward a partially synchronized state. It would be interesting to perform “probe” experiments on the units, to see whether in fact stimuli can be predicted from the time series of units to verify this.

In spite of these limitations, I argue that the general principles here are a viable algorithmic and mechanism-level theory (or proto-theory, perhaps) of object representation and that experiments related to many of these computational principles may serve to distinguish between the competing algorithmic theories within a view based computational framework. I now examine several such possibilities.

Slow Wave Phenomena

It is possible that the idea that slow wave phenomena in neural masses acts as a bifurcation or coupling parameter control, as I have suggested, is overstated or oversimplified. No detailed models based on integrate and fire dynamics have been developed to date to support this hypothesis; it is possible that spiking dynamics alone result in the slow wave phenomena, without anything as regular as a clock. Studies of the time course of synchronization phenomena show considerable variation from trial to trial, rather than a predictable time course (Gray, Engel et al. 1992). In all cases, episodes of oscillation are transient, lasting from 100-200 ms, followed by epochs of aperiodic activity. Whether this indicates intrinsic rhythms, the intermittent dynamics described by Tsuda as a binding strategy, or some unsuspected phenomena is not clear. If more regularity in synchronization were seen it would better support the theory of coding – however, the studies above are in primary visual cortex of anaesthetized cats, where one would really like to see similar studies in areas (V4, TE) more implicated in object level coding and recognition.

It may be that no model at the abstraction level of coupled maps could reproduce synchronization phenomena in EEG measures, though this has been difficult to achieve in more detailed models such as the Freeman group's KIII model; finding the invariant response to a stimulus remains a challenge in animal studies (Kay, Shimoide et al. 1995). Dynamics at the level of bifurcation and coupling parameters may play an interpretive role, but if the time course of their modulations are irregular and emerge directly from spike interactions, simplification may not be possible. I think it is too early to make this negative conclusion, though, as the line of investigation pursued here is still novel and untested, but tracks observations at many scales of neural dynamics which remain largely uninterpreted in computational or algorithmic terms.

The phenomenon of alpha blocking arguably is evidence against the Soca scheme. Alpha is maximal with eyes closed, especially in occipital (primary visual) and parietal areas, but is greatly reduced when a visual stimulus is presented (Basar 1998). This might be taken to indicate that the alpha band is an emergent cortical process governed by spike dynamics and graded potentials alone, rather than being a kind of dynamical control function as hypothesized here.

However, this reduction may be in part an artifact of coarse measurement in scalp recorded EEG. As reviewed in chapter 3, slow wave phenomena and changes in coherence are evident in local field potential in primary visual cortex. Modulatory dynamics and functional role of slow waves are still poorly understood. Thus, I argue the present demonstration of a computational role for modulations in parameters corresponding to slow wave phenomena is important and deserves further study.

There is some evidence that alpha frequencies are implicated in memory performance. Klimesch investigated relationships of short and long term memory with alpha band among age matched subjects; the good memory performers had 1.5 Hz higher alpha peaks compared to poorer performers. Alpha desynchronization was pronounced for subjects with poor memory (Klimesch 1996). The Soca network style described here, while primarily perceptual, can be described in part in terms of using slow wave control to turn image-like structures into a usable distributed representation suitable for comparison by hypothetical synchronization operators. This general principle may be effective in other areas, such as memory. The failure of slow waves to synchronize adequately might lead to deficits in perceptual and memory processes.

These deficits could be studied in the present conceptual framework, by examining temporal dispersion around the precise transitions to new dynamical parameters currently used. Parameters would be inhomogeneous across the array, but still drawn from the same value sets.

In the more complex biological network, such sampling of a transient process can not be the end result, but might be a starting point for further encoding in long term memory, and for matching against some version in working memory during search tasks. Changes in inter-regional coherence corresponding to readout or further processing of such an encoding should be expected, and these are precisely what the studies of Bressler, Basar and colleagues cited above demonstrate.

If Soca-like computation and coding occurs in IT cortical complex (or related form processing areas like V4) it is certain to work in conjunction with memory formation networks and comparison networks. Soca-like networks might play roles in view interpolation, invariant recognition, and similarity mapping. Memory formation might be handled by recurrent attractor networks, as advocated by Rolls and Amit; however, due to its superior handling of the compositionality issue, I favor the “exotic attractor” or itinerancy coding model, which could play a role in both (long term) memory formation and short term memory.

Stimulus Related Spatial Localization of Activity in Inferotemporal Cortex

Since the population code of Soca does not lead to asymptotic spatially localized activity in certain units, some alternative explanation is required to explain the observations of Tanaka and coworkers on spatial localizations, i.e. a small set of activated regions, with some small displacements of the spots by changing stimuli.

Several possible explanations are envisioned. In a biological system, the state vectors established by learning could be provided as input to the itinerancy coding network of Tsuda (Tsuda 1992) for permanent storage. Comparison would be implemented by some synchronization operations of the Soca network output with such an encoded memory, with success or frustration of synchronization driving behavioral response.

In this view, synchronization opponents, itinerancy coding, and competitive synchronization comparison processes are envisioned as the basic operations of object representation and recognition in biological systems. The work presented here emphasizes only the first aspect of this view of biological form processing, while other investigators have focused on memory (Tsuda 1992) and competitive synchronization processes.

This is a rather complex formulation which I must consign to the future work category; I also think it is more appropriate to a longer time scale encoding and recognition strategy involving eye movement scanning strategies. Some proposals more suited to preattentive rapid recognition are outlined here, which envision memory as more directly related to the oscillation structure and control parameters in the Soca network.

Selectionist Responses for Ensemble Response Character

One possibility is that Soca-like classifier networks are arrayed across anterior IT cortex; as they respond to input, some selection process occurs in which ensemble responses with undesirable characteristics (too synchronized during a pre-readout epoch) are suppressed by auxiliary networks sensitive to these ensemble responses characteristics.

Possible support for this theory comes from the optical studies of Tanaka and colleagues demonstrating a small spatial shift of maximum activity given different viewpoints of a stimulus. One possibility for the developmental structure and function of the anterior TE is that regular spatial gradients in some microcircuit parameter lead to a spatial array of columns analogous to the control space of the Soca networks. The small variations in response surface in the {bifurcation, coupling} control plane might lead to changes in maximum response to different views under such a selection criteria, with abrupt changes at critical transitions.

Anterior IT as Temporal Pattern Recognizer

An alternative proposal for explaining the results would re-envision a role for locally coded, combination code column “units” as proposed by Tanaka. Rather than being selective for (distal) feature combinations, such units might instead be responding to specific *ensemble frequency combinations* in the *transformed* signal, which might be computed upstream in the TEO, pIT or V4 regions. This detection could be on the basis of recognizing specific temporal patterns, such as a preferred set ensemble frequencies occurring in a short time window, or more distributed response where a local column responds to the ensemble frequency change in the last two fast cycles at the optimal readout time. The latter columns might then be linked by learning procedures.

Memory Based Search by Localized Slow Wave Interactions

Another proposal would suggest that a remote region cooperates with certain regions of IT to establish a Soca-like computation by injecting spatially localized slow wave activity into certain regions. This extra activity functions to change bifurcation or coupling parameters and resulting “stimulus tuned” ensemble frequency or phase distributions; however, it is detectable as an activity level when the incoming stimulus matches the expectation. This would suggest that the amplitude might be modulated and that if multi-channel local field potential studies were conducted in an area known to be activated for a particular stimulus through optical recording, the level of slow wave activity might be increased relative to non-active regions.

The findings of Nakamura of slow (4-6 Hz) stimulus related oscillations at the neuron level in the temporal pole are intriguing, suggesting to me that this may be the most plausible theory of matching. Correlation studies between these neurons and the temporal structure of upstream areas (TEO, V4, and pTE) could be performed to investigate whether these are deterministically linked and whether a causal flow can be determined.

A slight variation on this scheme envisions more intrinsic dynamics in the local regions. Local columns might have characteristic oscillatory structures and distributions which attempt to match, by synchronization or resonance phenomena, stimulus-created oscillations like those of the Soca network. Activation of parallel areas could be an artifact of best possible matches, while those areas with intrinsic dynamics failing to match the incoming structure have their activity suppressed.

In a primed search scenario, the slower oscillations observed by Nakamura in the temporal pole could interact with the characteristic local oscillatory structures to modulate the responses, allowing a match only for the local structures with a compatible slow wave dynamics. Correlation studies between the TE areas and the anterior temporal pole, using data such as Tanaka feature combinations, could explore possible interactions. Searching for one object A while measuring the response to putatively optimal complex stimulus B for a local region could show changes in the correlation structures of A and B, ultimately leading to suppression of activity. Searching for A and local region with optimal stimulus A should show a different correlation structure; observations of these dynamics could provide clues on how comparison is performed in a more complex dynamics scenario.

Transient Synchronization

The synchronized periodic oscillations seen in primary visual and prefrontal cortex have been chiefly interpreted as signatures of binding of features for segmentation and other Gestalt phenomena. I have suggested that such transient synchronization could also play a role in establishing the correct initial conditions for a Soca-style cycle of spatial cooperative processes under a modulated synchronization dynamics.

If some technique could be devised to interfere selectively with synchronization processes observed early in the recognition cycle in IT cortex, the effects of this on recognition performance could be examined. Theories of object recognition predicated on rate codes should predict little effect of synchronization; the present theory predicts substantial effects. This could be modeled by dispersing initial conditions about their current baseline values.

Aperiodic Oscillations

Aperiodic oscillations with synchrony and modulations of synchrony have been observed at various scales; certainly nonlinear dynamics are a potential source of such structure. At issue, though, is can we determine whether temporal codes are relevant, and whether they are produced by chaotic micro-circuit dynamics or reflect other types of computation (perhaps modulation of inputs with primed memories at the neural microcircuit level, as suggested by Eskandar et al (Eskandar, Optican et al. 1992).

The work here, stressing population codes, suggests that population measures through local field potential arrays or optical methods are the most promising techniques to address the question. If coding is distributed over populations and not localized to specific units, optimal stimulus predictions should be obtained from population measures. The differences in performance might not be evident unless complex stimuli are used.

Epochs of Desycnchronization and Synchronization

The modulations in effective connectivity (interpreted from correlations) can be explained in this way. As rhythmic volleys come into local recurrent networks from thalamic or other cortical regions, they change effective bifurcation and coupling parameters in the target local structure. For a given stimulus, the time course of synchronization is co-determined by inputs and by the dynamical flows inherent at a point in the bifurcation-coupling parameter plane. As demonstrated in the present work, such a dynamical flow may serve computational roles, functioning as a dynamical recognizer.

However, it is possible that modulations in effective connectivity arise naturally in the processing of coordinated arrays of chaotic oscillators, during representation formation or recognition, even in a single epoch. This should naturally occur as oscillators approach one another in phase space, whether the macrostate variable is interpreted as spike phase or ensemble average frequency.

Studies of modulations in correlation and effective connectivity as pioneered by Aertsen et al. (Aertsen, Gerstein et al. 1989), if performed in IT cortex along with response histogram procedures of Gochin et al (Gochin, Colombo et al. 1994), could reveal whether the best stimulus predictions are obtained in intervals of increased correlation for many pairs, as the present theory of computation and coding for object recognition would predict. RBF networks would predict no relationship between the highly correlated epochs and best histogram prediction.

The issue is confused by doubts over whether sub-regions of IT represent the locus of a feed-forward representation network, or an area of comparison of incoming dynamics with memory representation by unknown processes (possibly involving synchronization dynamics). In tasks with a visual memory component (the monkey must release a bar when a match to the target is shown after several distractors) it has been shown that neuronal responses are *decreased* substantially from their optimal stimulus when a stimulus matching the target is shown (Miller, Li et al. 1991).

Based on the various reviews cited earlier, we can exclude the anterior section (aTE) as being the likely site of comparison and priming dynamics. The best candidate regions to explore for Soca-like dynamics to be generated are intermediate ventral stages, primarily V4 and TEO.

Spatio-Temporal Patterns

One of the results presented here might be investigated in multi-channel local field potential or microelectrode array studies. In the study of parametric curves, the representations (sampled distributions) tended to have higher co-occurrence entropies as learning progresses. I am unaware of any multi-channel data analyzed by this criteria. One might examine whether this measure is predictive of differences in performance on recognition tasks.

Psychophysical Approaches to Theory Resolution

When specific views of 3D objects are used for training, error rates are shown to increase with rotation angle. This is one argument in favor of the RBF ensemble approach over the Soca model. However, this is presumably due to the fact that I did not supply or enforce conforming to any particular view in the learning process; another *network based* learning process using the same basic normalization principle but with a bias towards conforming to a view meeting some abstract (but temporally localized) principle might exhibit this effect. When only two views were used for training, peformance was highest on the trained views, but the fallout with view distance was irregular. This may also in part be due to the large angular separation between views provided to the network.

In order to discriminate between normalization in the Soca and RBF approaches, one might perform psychophysical investigations into preferred or canonical views. In the Soca approach, one might expect the preferred view to be that which is closest to the mean classifier output over all views, given the optimization criteria found to be successful here. If some class-separation strategy such as the maximum cross-entropy strategy is employed, the representation chosen, and corresponding best-map-to-mean view, would be expected to vary based on presentation order.

Recent view-based psychological theories have focused on the role of diagnostic features. A finding that presentation order dominates diagnostic features in determining preferred views would be evidence in favor of the maximum cross-entropy principle in determining representation. Presentation order has been observed to affect the neuron level activity during the delay (memory) period of a matching task, with the most similar activity arising from successively presented objects, even though there is little obvious relationship between the successively presented stimuli.

As I have repeatedly pointed out, however, multiple neuropsychological strategies and subsystems may be active, exploiting different regularities of the environment. Diagnostic features may be exploited by an alternative subsystem which is biased towards dominance by the nature of the task. Tarr et al. state that match to sample tasks, where a single target is identified, may bias subjects to rely on critical diagnostic features which may not scale well. In contrast sequential matching and naming, with continual presentation of different objects, elicits the formation or activation of a new representation at every trial.

The Encoding of Stimulus Meaning

On the basis of extensive studies in olfaction, W. Freeman has repeatedly stressed that neural dynamics encode the stimulus *meaning*, rather than a direct transformation of the feature space (Freeman and Barrie 1994). This is evidenced by changes in the stimulus evoked dynamical patterns for the same stimulus depending on the reinforcement paradigm last used with that stimulus. While I have not stressed this important point, which arguably should apply to other sensory modalities, I believe that this paradigm is compatible with the Soca classifier approach, as follows. We might consider that restricted areas of the network control space correspond to different motivation and arousal states related to emotional control of behavior. When a stimulus is learned, the dynamical encoding is restricted to those producible by control parameter sets in this subregion of the control space.

This assumes that possible solutions to a task like forming normalization solutions to multiple views are rather dense in the control space. The only evidence I have for this is the relatively successful search process with relatively few (3000) parameter sets explored. Further research on the density of such solutions would be of general interest, and in support of this theory of behaviorally relevant encodings.

SUMMARY: MAIN CONTRIBUTIONS

The main contributions of the thesis, from the perspectives of several disciplines, are summarized now.

In terms of **pure dynamics of discrete maps**, exploratory computational studies examined the short time (transient) behavior in two stage coupled logistic networks, where the stages form *synchronization opponents* adapted to a family of input forms. By opponents, I mean that the first stage has dynamics resulting in *expansion of the phase space volume*, while the second stage is, at least transiently, *contracting in phase space volume*. Initial conditions were binary distributions, which considered from the standpoint of synchronization dynamics represent an artificially synchronized starting point for the evolution of the network.

Some simple preliminary experiments examined transients in two map systems and lattices, with a range of parametric variation in both single stage and opponent stage scenarios. Parameters studied included synchronization over the bifurcation and coupling control plane, the effect of distance between the initial distribution components, and the instantaneous time distributions over a range of parameters with geometrically structured and spatially randomized binary distributions as initial conditions. In accordance with the more detailed spatially tuned chaotic neural models (Hansel and Sompolinsky 1992), it was shown that random spatial distributions produce relatively smooth distributions as the bifurcation parameter is varied while oriented initial conditions produce more irregular distributions. Coupled map lattices produce similar phenomena with relatively low computational cost. Smoothness in the time evolution and synchronization of a two map systems after small iteration counts suggests that some parameter perturbations will be tolerated.

From the perspective of **pattern recognition**, a method involving locally coupled, synchronization modulated nonlinear oscillator arrays was demonstrated to form a *population metric space* representation of objects, with partition cells in the dynamcal phase space taken as the dimensions of the space. In one family of experiments a metric space is explicitly constructed; in another, the objects do not occupy an obvious metric space, but clustering of objects with similar structure in a representation space emerges from learning to normalize views to a common distribution.

This spatiotemporal style of computing a representation contributes to resolving the classic dilemmas for representing local features, feature conjunction, and binding, while affording rapid recognition. The method was demonstrated to reach recognition rates as high as 85% in a nearest neighbor match scenario with 20 objects. The formalism of dynamical recognizers was extended from low dimensional dynamics to high dimensional dynamics

At the intersection of **dynamics, computation and pattern recognition**, preliminary evidence is presented that, for the given tasks and view normalization approach taken here, the *non-stationary* dynamics outperform a single preferred dynamical regime, such as the *transition to chaos* dynamics. At the outset of this work, it was hoped that a single dynamical stage might be found which solved the similarity problem for many objects. In a family of experiments recognizing objects rotated in depth, it was shown that a single dynamical stage is effective at rates far above chance, but results in 4% higher error rates than the non-stationary Soca strategy. There was also a slight trend toward lower average classification times with the nonstationary system. Given the non-deterministic nature of evolutionary learning, additional trials are needed to establish statistical significance of this result. This work is in progress, along with efforts to explore the effects of varying mutation rates from the baseline values used throughout the present study.

This small difference between single and two stage dynamics was unexpected; examination of sampled lattice time series reveals that during coupled chaotic evolution there are *fluctuations* of the occupied phase space volume at each step, so that characterization of a *stage* as subspace-desynchronizing or subspace-synchronizing relies on time averaged behavior, rather than an obvious monotonic increase or decrease in volume. If the recognition rate and time advantages hold with more trials, a deeper explanation of these advantages in terms of Markov chain dynamics would still be desirable. At this time, it seems that the original intuition motivating the Soca cyle is too simple an explanation, failing to capture the performance of the single stage system. The conceptual framework of probabilistic finite state automata and dynamical recognizers may be sufficient to explain the latter, but why should non-stationarity offer any improvement?

However, it is intriguing that an evolutionary scenario of improving performance with subtle improvements in timing control is suggested. The present task of recognizing the equivalence of views is emblematic of task-specific, constructive memory systems, where the memory consists of programming a general network topology with an implicit network of state flows between dynamical partition cells. Coordination dynamics between memory and recurrent computational modules may involve such programming of both task related computational systems and possibly also the programming of readout windows, sensitizing "receiver assemblies" to perform their tasks at specific times in concert with Markov-chain like computations.

A few preliminary results are presented regarding **evolutionary learning** of a computation with coupled maps. Unsurprisingly, evolutionary learning outperforms a random search strategy (15% error for the best evolved solution vs 29% for nearest neighbor classification with random search), with both evaluating 3000 parameter sets (genotypes). One notable aspect of the learning is that it occurs on the basis of an *abstract objective function* rather on the ultimate task for which the network is designed. Since the task itself involves parallel evaluation of all classifiers, this approach provides substantial speedup.

The stimulus equivalence experiments showed that the usual practice of Gaussian mutation of real valued genes, performed worse (also 29% vs.15% error) than random selection of new values from the entire range of possible values. This was conjectured to be due to the nonlinear mapping of the gene to the output space. These preliminary results concerning learning are based on very few learning trials; computing mean error rates over additional learning trials should be performed before drawing firm conclusions on the relative merits of mutation rates and single stages vs. multiple stage dynamics. . The possibility of improved performance through timing control seems helpful in understanding biological evolution, which presumably might have more difficulty searching a network space where effective topologies and network configurations are sparse islands in a sea of ineffective parameters. Experiments to assess the ease of learning tasks through network evolution in CML systems and more conventional connectionist architectures are called for.

At the intersection of **high level computer vision**, **visual psychology**, and **neuroscience of the ventral pathway**, it is shown that cooperative interactions of nonlinear oscillator transient trajectories can perform view normalization and view based recognition with recognition rates comparable to a recent feed-forward model, a statistical model with a rich input feature space. Unfortunately no comparison has yet been performed on a common set of data, but the data used here are challenging for several reasons:

1. Larger angular separation between views.
2. Inclusion of "extreme" end views in the recognition task.
3. Lower raw information due to use of silhouettes rather than color or grey scale images.

Another important distinction is that the Soca classifier set was trained solely on information theoretic principles, in sequential presentation. In contrast, the Chorus and SEEMORE systems used statistical overviews of the entire object world to choose network centers and set weights on feature spaces. I would not claim that there is no possible justification for such a procedure on biological grounds[34] .

---

[34] Freeman and Barrie have argued that one of the advantages of encoding all memory in a chaotic attractor is precisely that it contains the entire history of experience, which can be used to influence the formation of new memories (Freeman and Barrie 1994).

View interpolation with an ensemble of radial basis function networks has been recently proposed as a biologically motivated object recognition, and claimed to fit a variety of psychological and experimental neuroscience data; I explored several questions about the relative merits of the two approaches and suggested a multi-channel experiment to discriminate between the approaches.

The visual world used for training and testing was limited to isolated objects to focus on the problem of recognizing objects rotated in depth. Related paperclip objects have resulted in substantial error rates for humans and monkeys in learning and recognition tasks; the choice of data was motivated by continuity with such literature, and by the possibility of making predictions which can be tested with the same stimulus set. Modeling of primed search and multiple objects in a scene would be logical target tasks for next generation of this research; this might proceed in parallel with extensions to strictly network based learning and recognition.

The model of recognition is still a hybrid network - computational model rather than a pure network; in the present form, it cannot be considered as a complete biological model. Further work is required to synthesize a completely network based approach, but the principles discovered over the course of modeling and embodied in the objective functions suggest the effective form of network learning dynamics. In addition to network based learning,, a network based recognition strategy is required to more closely match biological observations. Two possible network architectures – synchronization of itinerency codes, and a back-end of ensemble frequency combination coding units were proposed as possible extended frameworks to increase the biological realism.

Finally, in the field of **computational neuroscience**, it was shown that modulations in the synchrony of large scale neural ensembles could play a role in learning and recognition. These modulations are affected in terms of bifurcation and coupling parameters of spatially distributed nonlinear systems. I offer a new interpretation for such modulations in the context of the *synchronous opponent cooperative activity* dynamics (Soca networks).

Previously, such modulations in synchrony, observed in multi-unit studies in pre-frontal and primary visual areas, have been interpreted as signatures of feature binding over space. I have proposed an understanding of such modulations as the unfolding of computational and coding processes in what I termed “coherence assemblies”.

The concept of information processing by symbolic dynamics, proposed long ago in the context of low dimensional systems, (Nicolis 1986) is extended to a higher dimensional system, resulting in a hypothesis of population coding via partition cell metric spaces.

A novel interpretation for transient synchronization episodes observed in primary visual cortex is proposed. My claim is that synchronization may be required in earlier stages of the visual pipeline - concurrent with or after segmentation processes - in order to prepare for the phase space expansion (desynchronization) stage of the two opponent stages performing a view normalization computation in IT or closely related form processing areas. This preparation consists of the creation of a high contrast primal sketch, which serves as an effective *initial subspace* for the spatiotemporal computation and population coding.

If epochs of increased synchronization and correlations are observed in association areas (as they have been in primary and prefrontal cortex) these may correspond to the state space contraction (synchronization) stage. These epochs should be correlated with increased stimulus prediction performance relative to epochs of desynchronization obtained from multi-channel vector measurements.

The intrinsic difficulties of reconstructing the behavior of non-stationary, spatially distributed oscillating systems in the brain suggest an important role for modeling. Models and simulations allow a constructive approach (Kaneko and Tsuda 1994) to supplant or enhance attempts at reverse-engineering the behavior of neural systems from correlations of neuron level components with stimulus, behavior, and other low level components. Constructive, in this context, refers both to the role of the network designer (human, evolution, or evolutionary computation) and to the rich capabilities of dynamical state flows and synchronization operators to rapidly construct implicit networks.

This is especially important if neural systems function by intrinsic or stimulus linked non-stationary behavior, since few methods exist for characterizing non-stationary dynamics and deducing the system function. Given clues about system function derived from evoked potential and spatial and temporal trends in synchronization, along with classical information theoretic principles from pattern recognition, models may be built for specific cognitive phenomena. Once a model exists, the same methods used for analysis of multi-channel biological signals (cross-correlation, coherence, and higher order spectral measures) can be applied to these artificial systems, leading to a fruitful interaction between theory and experiment.

In the simulation work presented, such signal analysis measures have not yet been applied to the resulting time series; rather, the observed trends in medium scale dynamics are taken as motivation to explore base nonstationarities in the service of some task, presently this exploration proceeds by evolutionary computing methods. Thus the work is in the spirit of constructive modeling, as defined above. By demonstrating a functional role for such nonstationarities in a classical perceptual task, I hope to contribute to expanding the dialog on neural computing mechanisms to include such relationships between modulations in synchrony, bifurcations as control for such modulations, and population computation and coding.

While speculative discussions of these topics are common among experimentalists and theorists in the last decade or so, concrete *engineering* models which perform well on a challenging task are still rare, as are attempts to systematically relate chaotic synchronization to any underlying neuroscience. I have attempted to realize both goals here in a fashion which is accessible to workers coming from psychology, experimental neuroscience, and engineering, providing enough background to justify what otherwise might seem very abstract forms of computation. If, by this effort, the coupled map style of computation gains credibility as a technique in computational neuroscience, these research communities may all benefit from an important new tool for understanding the programming style hidden in the rhythms of the brain and mind.

Appendix 1: Formal Language Definitions

Portions of this material are adapted from (Gurari 1989).

A Type 0 grammar ***G*** is defined as a quadruple <***A*** , Σ , ***P*** , ***S*** >, where

A is an alphabet, whose elements are called non-terminal symbols. ***A**** denotes the set of possible words formed from concatenating these symbols.

∑ is an alphabet disjoint from N, whose elements are called terminal symbols.

P is a relation of finite cardinality on (***A***)*, whose elements are called production rules. Moreover, each production rule (α,β) in P, denoted α→β, must have at least one non-terminal symbol in ***A*** . In each such production rule, α is said to be the left-hand side of the production rule, and β is said to be the right-hand side of the production rule.

S is a symbol in ***A*** called the start , or sentence , symbol.

A grammar ***G*** = <***N, Σ, P, S*** > is said to be a right-linear grammar if each of its production rules is either of the form α→ *x*β or of the form α → x, where α and β are non-terminal symbols in ***A*** and *x* is a string of terminal symbols in Σ*.

The grammar is said to be a left-linear grammar if each of its production rules is either of the form α→β*x* or of the form A→ *x*, where α and β are nonterminal symbols in ***A*** and *x* is a string of terminal symbols in Σ*.

The grammar is said to be a regular grammar if it is either a right-linear grammar or a left-linear grammar. A language is a **regular language** if it is generated by a regular grammar.

Such a Type 0 grammar G = <***A, Σ, P, S***> is said to be context-free if each of its production rules has exactly one non-terminal symbol on its left hand side, that is, if each of its production rules is of the form A->alpha .

The concept of a grammar can be extended to a *stochastic* grammar. The set ***P*** of productions can be associated one to one with an set π of probabilities on each production. The resulting system <***A, Σ, P,*** *π,* ***S***> is a stochastic grammar. The set of productions is a stochastic language. A state transition graph consisting of nodes for each element in **A** can be defined, with labeled arcs giving the probabilities of emitting a new token; this graph specifies a Markov process which generates strings in the stochastic language ***L.***

Appendix 2: Discrete Markov Chains

In this appendix basic definitions related to Markov chains are presented following the treatment in the Sochastic Processes text of Lawler (Lawler 1995). I occasionally interject comments drawing connections to dynamical systems. While Markov chains share certain concepts with dynamics systems such as periodicities, transient states, the basic dichotomy between deterministic and probabilistic systems remains. By considering coarse grained partitions of the dynamical phase space(as is also performed for symbolic dynamics) as the states in a Markov chain, one can convert from the deterministic to the stochastic mode of analysis.

STOCHASTIC AND MARKOV PROCESSES

A **stochastic process** is a collection of random variables X_t indexed by time. When time is a subset of the nonnegative integers {0,1,2 ...} the process is called discrete time. The random variables take values in a state space; this may be discrete (a finite or countably infinite set) or continuous. A **Markov process** is a stochastic process with the restriction that the change at time t is determined by the value of the process (i.e. the value of the state space scalar or vector in R^d) at time t, and not by values at times before t.

MARKOV CHAINS

A time homogeneous Markov chain is a Markov process described by a initial probability distribution and a transition probability matrix P, where the elements in the matrix P_{ij} are independent of time. The matrix P must be a stochastic matrix, satisfying the conditions:

$$0 \le P_{ij} \le 1, 1 \le i, j \le N$$

$$\sum_{j=1}^{N} P_{ij} = 1, 1 \le i \le N$$

The n-step transition probabilities $p_n(i,j)$ are given by P^n.

An absorbing state is a state which leads to itself with probability one. This is equivalent to a fixed point attractor in a dynamical system.

A Markov chain is irreducible if all states communicate, i.e. there is a path between the two states in the transition matrix. Otherwise, the state space is partitioned into disjoint sets called communication classes. These may be transient or recurrent, inheriting properties from their constituent states. A transient state will leave the state with probability 1 (when the system is captured by an absorbing state).

The partitioning of states into transients and absorbing states corresponds to the partitioning of contracting dynamical systems into attractors and (basin) transients. For expanding (chaotic) dynamical systems, there may be forbidden regions of phase space on the attractor, which nevertheless could be given as an initial condition or reached by a perturbation.

A recurrent Markov chain is one for which each state is visited infinitely many times. In contrast, a transient chain is one for which each state is visited a finite number of times.

Appendix 3: Signal Analysis

Autocorrelation

Autocorrelation emphasizes periodic components of a time series by comparing values separated by a regular time interval (lag). Each sample of the time series is multiplied by the value shifted in time by a fixed lag; the sum of these products is the autocorrelation function for a particular lag time. The autocorrelogram is generated by

1. Removing the mean from the signal
2. Normalizing by signal power
3. Plotting lag time (x axis) vs. the normalized correlation coefficient.

A similar process, cross-correlation, is used to compare two time series. Autocorrelation can, of course, be viewed as a special case of cross correlation.

$$R_{xy}\left[k\right] = 1/N \sum_{n=0}^{N-1} x[n]y[n+k]$$

where x[n] and y[n] are values of time series x and y at time n, k is the lag of y[n] with respect to x[n], and N is the number of samples in series or window in which the function is computed.

Let N be the number of sampled points in two signals x and y.

Power Spectrum

$$S_{xx}(f) = \frac{FFT(x) * FFT^{*}(x)}{N}$$

where *FFT** is the complex conjugate of *FFT(x)* (e.g. the negated imaginary part of *FFT(x))*.

Cross Power Spectrum

$$S_{xy}(f) = \frac{FFT(y) * FFT^{*}(x)}{N^{2}}$$

Coherence function

$$C_{xy} = \frac{[|(Averaged\ S_{\mathrm{xy}}(f))|]^{2}}{Averaged\ \mathrm{S}_{\mathrm{xx}}(f) * Averaged\ \mathrm{S}_{\mathrm{yy}}(f)}$$

The coherence function requires an average of two or more measurements of the signals under analysis. For a single measurement, it would register unity at all frequencies. To average a complex quantity such as the cross power spectrum $S_{xy}(f)$, sum it in the complex form, divide by the number of averaging trials, then convert to magnitude and phase with rectangular to polar conversion.. The auto power spectra, $S_{AA}(f)$ and $S_{BB}(f)$ are real quantities.

Bibliography

Abbot, L. F. and Sejnowski, T. J. (1999). Introduction to Neural Codes and Distributed Representations. Neural Codes and Distributed Representations: Foundations of Neural Computation. L. F. Abbot and T. J. Sejnowski. Cambridge, MIT Press**:** vii-xxiii.

Abraham, R. H., Corliss, J. and Dorband, J. E. (1991). “Order and Chaos in the Toral Logistic Lattice.” Journal of Bifurcations and Chaos **1**(1): 227-234.

Aertsen, A. M. H. J. and Gerstein, G. (1991). Dynamic aspects of neuronal cooperativity: fast stimulus-locked modulations of effective connectivity. Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag.

Aertsen, A. M. H. J., Gerstein, G., Habib, M. K. and Palm, G. (1989). “Dynamics of neural firing correlation: Modulation of effective connectivity.” Journal of Neurophysiology **61**(5): 900-917.

Allman, J., Miezen, F. and McGuiness, E. (1985). “Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons.” Perception **14**: 105-126.

Allman, J., Miezin, F. and McGuiness, E. L. (1985). “Direction- andvelocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT).” Perception **14**: 105-126.

Amari, S.-I. (1974). “A method of statistical neurodynamics.” Kybernetic **14**: 201.

Amit, D. J. (1995). “The Hebbian Paradigm Reintegrated: Local Reverberations as internal representations.” Behavioral and Brain Sciences **18**: 617-657.

Anninos, P. A., Beek, B., Csermely, T. J., Harth, E. M. and Pertile, G. (1970). “Dynamics of Neural Structures.” Journal of Theoretical Biology **26**: 121-148.

Balakrishnan, K. and Honar, V. (1995). Evolutionary Design of Neural Architectures. Ames, Iowa State University Dept. of Computer Science.

Barreto, E., Kostelich, E., Gregobi, C., Ott, E. and Yorke, J. A. (1995). "Efficient switching between controlled unstable periodic orbits in higher dimensional chaotic systems." Physical Review E **51**: 4169-4172.

Basar, E. (1998). Brain Function and Oscillations Vol. I. Berlin, Springer -Verlag.

Basar, E. (1998). Brain Function and Oscillations Vol. II. Berlin, Springer -Verlag.

Biederman, I. and Gerhardstein, P. C. (1993). "Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance." Journal of Experimental Psychology: Human Perception and Performance **19**(1162-1182).

Biedermann, I. (1987). "Recognition-by-components: A theory of human image understanding." Psychological Review **94**: 115-147.

Bovik, A. C., Clark, M. and Geisler, W. S. (1989). "Multichannel texture analysis using localized spatial filters." IEEE Trans. Pattern Anal. Machine Intell. **11**: 674-693.

Bower, G. H. and Clapper, J. P. (1989). Experimental Methods in Cognitive Science. Foundations of Cognitive Science. M. I. Posner. Cambridge MA, MIT Press**:** 245-300.

Bressler, S. L. (1995). "Large-scale cortical networks and cognition." Brain Research Reviews **20**: 288-304.

Bressler, S. L. and Nakamura, R. (1993). "Episodic multi-regional cortical coherence at multiple frequencies during visual task performance." Nature **366**: 153-156.

Brigger, P. (1995). Morphological shape representation using the skeleton decomposition: application to image coding. Dissertation (Computer Science),. Lausanne, Ecole Polytechnique Federale de Lausanne.

Brown, V. R., Levine, D. S. and Shirey, T. (2000). Oscillations in Neural Systems, Lawrence Erlbaum Associates.

Bullock, T. H. (1993). How Do Brains Work? Papers of a comparative neurophysiologist. Boston, Birkhauser.

Bullock, T. H., Achimowicz, J. Z., Duckrow, R. B., Spencer, S. S. and Iragui-Madoz, V. J. (1995). "Bicoherence of intracranial EEG in sleep, wakefulness, and seizures." EEEG Clinical Neurophysiology **103**: 661-678.

Bulthoff, H. H. and Edelman, S. (1992). "Psychophysical support for a two-dimensional view interpolation theory of object recognition." Proc. Natl. Acad. Sci. USA **89**: 60-64.

Calvin, W. H. (1995). Cortical Columns, Modules, and Hebbian Cell Assemblies. The Handbook of Brain Theory and Neural Networks. M. Arbib. Cambridge, MIT Press**:** 269-272.

Campbell, S. and Wang, D. (1996). "Synchronization and Desynchronization in a Network of Locally Coupled (Wilson-Cowan) Oscillators." IEEE Transactions on Neural Networks **7**(3): 541-554.

Carvalho, R., R., V. l. M. and Seixas, J. (1999). "Feigenbaum Networks." Physica D **126**: 27-37.

Chapeau-Blondeau, F. and Chauvet, G. (1992). "Stable, Oscillatory and Chaotic Regimes in the Dynamics of Small Neural Networks With Delay." Neural Networks **5**: 735-743.

Charniak, E. (1993). Statistical Language Learning. Cambridge, MA, MIT Press.

Chate, H. and Manneville, P. (1989). "Coupled map lattices as cellular automata." Journal of Statistical Physics **56**: 357-370.

Chen, C.-h. (1973). Statistical Pattern Recognition. Rochelle Park NJ, Spartan Books.

Chua, L. O. and Yang, L. (1988). "Cellular Neural Networks: Theory." IEEE Transactions on Circuits and Systems **35**(10): 1257.

Cowan, J. D. (1974). Models of Large Scale Nervous Activity. Some Mathematical Questions in Biology V. J. D. Cowan. Providence, American Mathematical Society**:** 99-133.

Del Bimbo, A. (1999). Visual Information Retrieval. San Francisco, Morgan Kaufmann.

DeMaris, D. (1995). Computing shape similarity with chaotic reaction diffusion spectra. World Congress on Neural Networks, Washington D.C.

DeMaris, D. (1995). Spatially Extended Chaos and the Perception of Form. Electrical and Computer Engineering. Austin, University of Texas.

Dinse, H. R., Kruger, K., Mallot, H. A. and Best, J. (1991). Temporal Structure of Cortical Information Processing: Cortical Architecture, Oscillations, and Non-Separability of Spatio-Temporal Receptive Field Organization. Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag**:** 68-104.

Donath, W. and Hoffman (1972). "Algorithms for partitioning of graphs and computer logic based on eigenvectors of connection matrices." IBM Technical Disclosure Bulletin **15**: 938-944.

Driebe, D. (1999). Fully chaotic maps and broken time symmetry. Dordrecht, Kluwer Academic.

Duda, R. O. and Hart, P. (1973). Pattern classification and scene analysis. New York, Wiley.

Duvdevani-Bar, S. and Edelman, S. (1999). "Visual recognition and categorization on the basis of similarities to multiple class prototypes." Intl. J. of Computer Vision **33**: 201-228.

Eckhorn, R. (2000). Cortical Processing by Fast Synchronization: High Frequency Rhytmic and Non-rhythmic Signals in the Visual Cortex Point to General Principles of Spatiotemporal Coding. Time and the Brain. R. Miller. Lausanne, Gordon & Breach.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Munk and Reitbock, H. J. (1988). "Coherent Oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat." Biological Cybernetics **60**: 121-130.

Edelman, S. (1995). "Class similarity and viewpoint invariance in the recognition of 3D objects." Biological Cybernetics **72**: 207-220.

Edelman, S. (1999). Representation and Recognition in Vision. Cambridge MA, MIT Press.

Edelman, S. and Weinshall, D. (1991). “A self-organizing multiple-view representation of 3D objects.” Biological Cybernetics **64**: 209-219.

Elbert, T., Ray, W. J., Wowalik, Z. J., Skinner, J. E., Graf, K. E. and Birbaumer, N. (1994). “Chaos and physiology: deterministic chaos in excitable cell assemblies.” Physiological Reviews **74**(1): 1-40.

Eskandar, E. N., Optican, L. M. and Richmond, B. J. (1992). “Role of anterior temporal neurons in visual memory 1: Temporal encoding of information about visual images, recalled images, and behavioral context.” Journal of Neurophysiology **68**(4): 1277-1295.

Eskandar, E. N., Optican, L. M. and Richmond, B. J. (1992). “Role of anterior temporal neurons in visual memory 2: Multiplying temporal wave forms related to vision and memory.” Journal of neurophysiology **68**(4): 1296--13.

Farhat, N. H. and del Moral Hernandez, E. (1996). Recurrent networks with recursive processing elements: paradigm for dynamical computing. Adaptive Computing: Mathematical and Physical Methods for Complex Environments, Bellingham, Wash., SPIE.

Flower, D. R. (1998). “On the Properties of Bit String-Based Measures of Chemical Similarity.” Journal of Chemical Information and Computer Science **38**: 378-386.

Freeman, H. (1985). Image Processing and Pattern Recognition. Proc. of Advances in Image Processing and Pattern Recognition. V. Capellini and R. Marconi. Amsterdam, Elsevier Science.

Freeman, W. J. (1992). Predictions on neocortical dynamics derived from studies in paleocortex. Induced Rhythms in the Brain. E. Basar and T. H. Bullock. Boston, Birkhauser.

Freeman, W. J. (1999). “Noise-induced first-order phase transitions in chaotic brain activity.” International Journal of Bifurcations and Chaos **9**(11): 2215-2218.

Freeman, W. J. (2000). Neurodynamics: an exploration in mesoscopic brain dynamics. London, Springer.

Freeman, W. J. and Barrie, J. M. (1994). Chaotic oscillations and the genesis of meaning in the cortex. Temporal Coding in the Brain. G. Buzsaki, R. Llinas, W. Singer, A. Berthoz and Y. Christen. Berlin, Springer-Verlag**:** 13-37.

Fuji, H., Ito, J., K., A., N., I. and Tsukada M.. Neural Networks, 1303-1350, 1996 (1996). “Dynamical Cell Assembly Hypothesis: Theoretical Possibility of Spatio-temporal Coding in the Cortex.” Neural Networks **9**: 1303-1350.

Gallant, J. L., Braun, J. and van Essen, D. C. (1993). “Selectivity for polar, hyperbolic and cartesian gratings in macaque visual cortex.” Science **259**: 100-103.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P. and Gore, J. C. (1997). “Levels of categorization in visual object studied with functional MRI.” Current Biology.

Gauthier, I. and Tarr, M. J. (1997). “Becoming a Greeble Expert: Exploring the face recognition mechanism.” Vision Research **37**: 1673-11682.

Gersch, W. (1987). Non-Stationary Multichannel Time Series Analysis. Methods of Analysis of Brain Electrical and Magnetic Signals. A. S. Gevins and A. Remond. Amsterdam, Elsevier.

Gerstein, G. L. (1988). Information Flow and State in Cortical Networks: Interpreting Multi-neuron Experiments. Organization of Neural Networks. W. v. Seelen, G. Shaw and U. M. Leinhos. Weinheim, VCH**:** 53-75.

Giblin, P. J. and Kimia, B. B. (1999). On the intrinsic reconstruction of shape from its symmetries. 1999 IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, USA, IEEE Computer Society.

Giblin, P. J. and Kimia, B. B. (1999). On the local form and transitions of symmetry sets, medial axes, and shocks. Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, IEEE Computer Society.

Gochin, P. M., Colombo, M., Dorfman, G. A., Gerstein, G. L. and Gross, C. G. (1994). “Neural ensemble coding in inferior temporal cortex.” Journal of Neurophysiology **71**: 2325-2337.

Goldmeier, E. (1972). Similarity in Visually Perceived Form. New York, International Universities Press.

Gomez, F. and Miikkulainen, R. (1997). "Incremental Evolution of Complex General Behavior." Adaptive Behavior **5**: 317:342.

Grassberger, P. (1988). "On symbolic dynamics of one-humped maps of the interval." Zeitschrift fur Naturforschung **43A**: 671-680.

Grassberger, P. (1991). Information and Complexity Measures in Dynamical Systems. Information Dynamics. H. Atmanspacher and H. Scheingraber. New York, Plenum Press. **256**.

Gray, C. M., Engel, A. K., Konig, P. and Singer, W. (1992). "Synchronization of oscillatory neuronal responses in cat striate complex: Temporal properties." Visual Neurosciences **8**(337-347).

Gray, C. M., P., K., Engel, A. K. and W., S. (1989). "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties." Nature **338**: 334-337.

Gray, C. M. and Singer, W. (1989). "Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex." Proc. National Academy of Science USA **86**: 1698-1702.

Gregson, R. A. M. (1988). Nonlinear Psychophysical Dynamics. Hillsdale, NJ, Lawrence Erlbaum Associates Inc.

Gregson, R. A. M. (1995). Cascades and Fields in Perceptual Psychophysics. Singapore, World Scientific.

Griniasty, M., Tsodys, M. V. and Amit, D. J. (1993). "Conversion of Temporal Correlations between Stimuli to Spatial Correlations Between Attractors." Neural Computation **5**: 1-17.

Grossberg, S. (1980). "How does the brain build a cognitive code?" Psychological Review **87**: 1-51.

Gurari, E. (1989). Introduction to the Theory of Computation. New York, Computer Science Press.

Hansel, D. and Sompolinsky, H. (1992). “Synchronization and Computation in a Chaotic Neural Network.” Physical Review Letters **68**(5): 718-721.

Hansel, D. and Sompolinsky, H. (1996). “Chaos and Synchrony in a Model of a Hypercolumn in Visual Cortex.” Journal of Computational Neuroscience **3**: 7-34.

Hansen, L. and Salamon, P. (1990). “Neural network ensembles.” IEEE Trans. Pattern Analysis and Machine Intell. **12**: 993-1001.

Hari, R. (1997). “Human Cortical Oscillations.” Trends In Neuroscience **20**: 44-49.

Harth, E. and Tzanakou, E. (1974). “Alopex: a stochastic method for determining visual receptive fields.” Vision Research **14**: 1475-1482.

Harth, E. M., Csermely, T. J., Beek, B. and Lindsay, R. D. (1970). “Brain functions and neural dynamics.” Journal of Theoretical Biology **26**: 93-120.

Hayward, W. G. (1998). “Effects of Outline Shape in Object Recognition.” Journal of Experimental Psychology: Human Perception and Performance **24**(2): 1-14.

Hayward, W. G. and Tarr, M. J. (1997). “Testing Conditions for Viewpoint Invariance.” Journal of Experimental Psychology Human Perception and Performance **23**(5): 1511-1521.

Holden, A. J., Tucker , J. V. and Thompson, B. C. (1991). Excitable Media as Computational Systems. Nonlinear Wave Processes in Excitable Media. New York, Plenum Press.

Hordijk, W., Crutchfield, J. P. and Mitchell, M. (1998). Mechanisms of Emergent Computation in Cellular Automata. Parallel Problem Solving from Nature V, Springer-Verlag.

Hubel, D. H. and Wiesel, T. N. (1962). “Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex.” J. Physiol. Lond. **160**: 106-154.

Ingraham, R. L. (1991). A Survey of Nonlinear Dynamics "chaos theory". Singapore, World Scientific.

Ito and Kaneko, K. (2000). “Self organized hierarchical structure in a plastic network of chaotic units.” Neural Networks **13**(3): 275-281.

Kaminski, M. J. and Blinowska, K. J. (1991). “A new method of the description of the information flow in brain structures.” Biological Cybernetics **65**: 203-210.

Kaneko, K. (1986). Collapse of Tori and Genesis of Chaos in Dissipative Systems. Singapore, World Scientific.

Kaneko, K. (1989). “Spatiotemporal chaos in one and two dimensional coupled map lattices.” Physica D **37**: 1-47.

Kaneko, K. (1990). “Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements.” Physica D **41**: 137-142.

Kaneko, K. (1993). “Overview of coupled map lattices.” Chaos **2**(3): 279-282.

Kaneko, K. (1993). Theory and applications of coupled map lattices. Chichester, John Wiley & Sons.

Kaneko, K. and Tsuda, I. (1994). “Constructive Complexity and Artificial Reality.” Physica D **75**: 1-10.

Karr, C. (1991). Air-injected hydrocyclone optimization via genetic algorithm. Handbook of genetic algorithms. L. Davis. New York, Van Nostrand Reinhold**:** 223-236.

Kay, L., Shimoide, K. and Freeman, W. J. (1995). “Comparison of EEG time series from rat olfactory system with model composed of nonlinear coupled oscillators.” International Journal of Bifurcations and Chaos **5**(3): 849-858.

Kelso, J. A. S., Case, P., Holroyd, T., Horvath, E., Raczaszek J, B., T. and M., D. (1995). Multistability and Metastability in Perceptual and Brain Dynamics. Ambiguity in Mind and Nature. K. P. and S. M. Berlin, Springer Verlag.

Kimia, B. B. and Siddiqi, K. (1994). Geometric heat equation and nonlinear diffusion of shapes and images. Computer Vision and Pattern Recognition, Seattle, IEEE Computer Society.

Klimesch, W. (1996). "Memory processes, brain oscillations and EEG synchronization." Int. J. Psychophysiology **24**: 61-100.

Korn, F., Sidiropouls, N., Faloultsos, C., Siegel, E. and Protopapas, Z. (1996). Fast and Effective Similarity Search in Medical Tumor Databases Using Morphology. Multimedia Storage and Archiving Systems.

Kovacs, I. and Julesz, B. (1994). "Perceptual sensitivity maps within globally defined visual shapes." Nature **370**: 644-646.

Kowalski, J. M., Albert, G. L., Rhoades, B. K. and Gross, G. W. (1992). "Neuronal Networks with Spontaneous, Correlated Bursting Activity: Theory and Simulations." Neural Networks **5**: 805-822.

Kozma, R. (2000). Personal Communication.

Kruger, J. (1991). Spike train correlations on slow scales in monkey visual cortex. Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag.

Kruger, J. and Becker, J. D. (1991). "Recognizing the visual stimulus from neuronal discharges." Trends in Neuroscience **14**: 282-285.

Krumhansl, C. L. (1978). "Concerning the applicability of geometric models to similarity data: the interrelationship between similarity and spatial density." Psychological Review **85**: 445-463.

Lashley, K. S. (1942). The problem of cerebral organization in vision, Jaques Cattell Press.

Lawler, G. F. (1995). Introduction to Stochastic Processes. Boca Raton, Chapman & Hall/CRC.

Lee, T. S., D., M., Romero, R. and V.A.F, L. (1998). "The role of the primary visual cortex in higher level vision." Vision Research **38**: 2429-2454.

Leedham, G. (1991). Pattern Recognition. Image Processing. D. Pearson. London, McGraw-Hill.

Lind, D. and Marcus, B. (1995). An Introduction to Symbolic Dynamics and Coding. Cambridge, Cambridge University Press.

Logothetis, N. K., Pauls, J., Bulthoff, H. H. and Poggio, T. (1994). “View-dependent object recognition by monkeys.” Current Biology **4**: 401-414.

Logothetis, N. K., Pauls, J. and Poggio, T. (1995). “Shape representation in the inferior temporal cortex of monkeys.” Current Biology **5**(5): 552-563.

Logothetis, N. K. and Sheinberg, D. L. (1996). Recognition and Representation of Visual Objects in Primates: Psychophysics and Physiology. The Mind-Brain Continuum: sensory processes. F. Llinas and P. S. Churchland. Cambridge, MIT Press**:** 147-172.

Maistrenko, Y., Popovych, O. and M., H. (2000). “On strong and weak chaotic partial synchronization.” Intl. Journal of Chaos and Bifurcations **10**(1): 179-203.

Mannion, C. L. T. and Taylor, J. G. (1992). Information Processing by Oscillating Neurons. Coupled Oscillating Neurons. J. G. Taylor and C. L. T. Mannion. London, Springer-Verlag.

Maragos, P. (1988). “Pattern spectrum and multiscale shape representation.” IEEE Trans. Pattern Analysis and Machine Intelligence **11**(7): 701-716.

Marr, D. and Nishihara, H. K. (1978). “Representation and recognition of the spatial organization of three dimensional structures.” Proceedings of the Royal Society of London B **204**: 301:328.

Marr, D. C. (1982). Vision. San Francisco, W.H. Freeman and Co.

McIntosh, A. R. and Gonzalez-Lima, F. (1994). “Structural equation modeling and its application to network analysis is functional brain imaging.” Human Brain Mapping **2**: 2-22.

Mel, B. W. (1997). “SEEMORE: Combining Color, Shape, and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition.” Neural Computation **9**: 777-804.

Miller, E. K., Li, L. and Desimone, R. (1991). “A neural mechanism for working and recognition memory in inferior temporal cortex.” Science **254**: 1377-1379.

Milner, A. D. (1999). Neuropsychological studies of perception and visuomotor control. Attention, Space, and Action. G. W. Humphreys, J. Duncan and A. Treisman. Oxford, Oxford University Press**:** 217-231.

Mitchell, M., Crutchfield, J. P. and Das, R. (1996). Evolving Cellular Automata with Genetic Algorithms: A Review of Recent Work. First International Conference on Evolutionary Computation and Its Applications, Moscow, Russia, Russian Academy of Sciences.

Mitchell, M., Hraber, P. and Crutchfield, J. (1993). "Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations." Complex Systems **7**: 89-130.

Miyashita, Y. (1988). "Neuronal correlate of visual associative long-term memory in the primate cortex." Nature **335**: 817-820.

Miyashita, Y. and Chang, H. S. (1988). "Neuronal correlate of pictorial short-term memory in the primate temporal cortex." Nature **331**: 68-70.

Moore, C. (1998). "Dynamical Recognizers: Real-time Language Recognition by Analog Computers." Theoretical Computer Science **201**(99-136).

Moriarty, D. E. and Miikkulainen, R. (1996). "Efficient reinforcement learning through symbiotic learning." Machine Learning **22**: 11-32.

Mountcastle, V. B. (1978). The Mindful Brain: Cortical Organization and the Group-Selective Theory of Higher Brain Function. Cambridge, MIT Press.

Mumford, D. (1989). Analysis and Synthesis of Human and Avian Categorization of Fifteen Simple Polygons, Harvard University.

Mumford, D. (1994). Neuronal Architectures for Pattern Theoretic Problems. Large-Scale Neuronal Theories of the Brain. C. Koch and J. Davis. Cambridge, MA, MIT Press**:** 125-152.

Nakamura, K. and Kubota, K. (1996). "The primate temporal pole: its putative role in object recognition and memory." Behavioral Brain Research **77**: 53-77.

Nakamura, K., Mikami, A. and Kubota, K. (1991). "Unique oscillatory activity related to visual processing in the temporal pole of monkey." Neuroscience Research **12**: 293-299.

Natschlager, T. and Ruf, B. (1998). “Spatial and temporal pattern analysis via spiking neurons.” Network: Computation in Neural Systems **9**(3): 319-332.

Nicolis, J. S. (1986). “Chaotic dynamics applied to information processing.” Reports on Progress in Physics **49**(10): 1109-1196.

O'Brian, G. L. (1981). “The road coloring problem.” Israel Journal of Mathematics **39**: 145-154.

Palm, G. (1982). Neural Assemblies. Berlin, Springer-Verlag.

Pardey, J., Roberts, S. and Tarassenko, L. (1996). “A Review of Parametric Modeling Techniques for EEG Analysis.” Medical Engineering Physics **18**(1): 2-11.

Pavlidis, T. (1977). Structural Pattern Recognition. New York, Springer.

Pearlmutter, B. A. (1990). Dynamic recurrent neural networks. Pittsburgh, School of Computer Science, Carnegie Mellon University.

Perez, J. C. (1988). De Nouvelle Voies Vers L'Intelligence Artificielle: Pluri-Disciplnarity, Auto-organizatin, Reseau Neronaux. Paris, Masson.

Perez, J. C. and Massotte, P. (1987). Chaos Fractal Attractor, IBM Corp.

Poggio, T. and Edelman, S. (1990). “A network that learns to recognize three-dimensional objects.” Nature **343**: 263-266.

Poggio, T. and Girosi, F. (1990). “Regularization algorithms for learning that are equivalent to multilayer networks.” Science **247**: 978-982.

Pollack, J. B. (1990). Recursive distributed representations. Connectionist Symbol Processing. B. Hington. Cambridge, MIT press**:** 77-105.

Pollack, J. B. (1991). “The Induction of Dynamical Recognizers.” Machine Learning **7**: 227-252.

Potter, M. A. and De Jong, K. A. (1994). A Cooperative Coevolutionary Approach to Function Optimization. Third Parallel Problem Solving from Nature, Jerusalem, Israel, Springer-Verlag.

Price, C. B., Wambacq, P. and Oosterlinck, A. (1993). “The plastic coupled map lattice: a novel image processing paradigm.” Chaos **2**(3): 351-363.

Purves, D., Riddle, D. R. and LaMantia, A.-S. (1992). “Iterated patterns of brain circuitry (or how the cortex gets its spots).” Trends in the Neurosciences **15**: 362-368.

Rabiner, L. R. and Juang, B. H. (1986). An introduction to hidden Markov models. IEEE ASSP Magazine**:** 4-15.

Ratcliff, R., Van Zandt, T. and McKoon, G. (1999). “Connectionist and Diffusion Models of Reaction Time.” Psychological Review **106**: 261-300.

Ratliff, F. (1965). Mach Bands: quantitative studies on neural networks in the retina. San Francisco, Holden Day.

Rentschler, I., Hubner, M. and Caelli, T. (1988). “On the discrimination of compound Gabor signals and textures.” Vision Research **28**: 279-291.

Richmond, B., Optican, L., Podell, M. and Spitzer, H. (1987). “Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex.” Journal of Neurophysiology **57**: 132-146.

Riehle, A., Grun, S., Diesmann, M. and Aertsen, A. (1997). “Spike Synchronization and Rate Modulation Differentially Involved in Motor Cortical Function.” Science **278**: 1950-1953.

Rieke, F., Bodnar, D. A. and Bialek, W. (1995). “Naturalistic Stimuli Increase the Rate and Efficiency of Information Transmission by Primary Auditory Afferents.” Proceedings of the Royal Society of London B **262**: 259-265.

Rolls, E. T. (1992). “Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical areas.” Philosophical Transactions of the Royal Society, London [B] **335**: 11-21.

Rolls, E. T. and Baylis, G. C. (1986). “Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey.” Experimental Brain Research **65**: 38-48.

Rolls, E. T., Baylis, G. C., Hasselmo, M. E. and Naiwa, V. (1989). “The effect of learning on the face-selective responses of neurons in the cortex in the superior temporal sulcus of the monkey.” Experimental Brain Research **76**: 153-164.

Rolls, E. T. and Treves, A. (1998). Neural Networks and Brain Function. Oxford, Oxford University Press.

Rosch, E. (1975). “Cognitive representation of semantic categories.” Journal of Experimental Psychology: General **104**: 192-233.

Rosenfeld, A. (1979). Picture Languages: Formal Models for Pattern Recognition, Academic Press.

Scasseleti, B., Alexopoulos, S. and Flickner, M. (1994). Retrieving Images by 2D shape: a comparison of computation methods with human perceptual judgements. Conference on Storage and Retrieval for Image and Video Databases, SPIE.

Sergent, J., Ohta, A. and MacDonald, B. (1992). “Functional neuroanatomy of face and object processing: A positron emission tomography study.” Brain **115**: 15-36.

Simoncelli, E. P., Freeman, W. T., Adelson, E. H. and Heeger, D. J. (1992). “Shiftable Multi-Scale Transforms or, "What's Wrong with Orthonormal Wavelets".” IEEE Trans. Information Theory **38**(2): 587-607.

Singer, W. (1996). Neuronal Synchronization: A solution to the binding problem? The Mind-Brain Continuum Sensory Processes. R. R. Llinas and P. S. Churchland. Cambridge, MIT Press.

Skarda, C. and Freeman, W. (1987). “How brains make chaos in order to make sense of the world.” Behavioral and Brain Sciences **10**: 161-195.

Softky, W. P. and Koch, C. (1994). “The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs.” Journal of Neuroscience **7**: 177-191.

Stamford, J. A. (1990). “Fast Cyclic Voltammetry: Measuring Transmitter Release in Real Time.” Journal of Neuroscience Methods **34**: 67-72.

Swindale, N. V. (1990). “Is the cerebral cortex modular?” Trends in the Neurosciences **12**: 487-492.

Tanaka, K. (1993). "Neuronal mechanisms of object recognition." Science **261**: 685-688.

Tanaka, K. (1996). "Representation of visual features of objects in the inferotemporal cortex." Neural Networks **9**(8): 1459-1475.

Tanaka, K., Saito, H., Fukuda, Y. and Moriya, M. (1991). "Coding visual images of objects in the inferotemporal cortex of the macaque monkeys." Journal of Neurophysiology **66**: 170-189.

Tarr, M. J. (2000). Visual Pattern Recognition. Encyclopedia of Psychology. A. E. Kazdin. Washington, D.C., American Psychological Association.

Tarr, M. J., Bulthoff, H. H., Zabinski, M. and Blanz, V. (1997). "To what extent do unique parts influence recognition across changes in viewpoint?" Psychological Science **8**(4): 282-289.

Tarr, M. J. and Pinker, S. (1989). "Mental rotation and orientation-dependence in shape recognition." Cognitive Psychology **21**: 233-282.

Taylor, J. G., Krause, B., Shah, N. J., Horwitz, B. and Mueller-Gaertner, H.-W. (2000). "On the Relation Between Brain Images and Brain Neural Networks." Human Brain Mapping **9**: 165-182.

Taylor, J. G. and Mannion, C. L. T. (1992). Coupled Oscillating Neurons. London, Springer-Verlag.

Tek, H. and Kimia, B. B. (1999). Symmetry maps of free-form curve segments via wave proagation. Seventh IEEE Conference on Computer Vision, Los Alamitos, IEEE Computer Society.

Tovee, M. J. and Rolls, E. T. (1995). "Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex." Visual Cognition **2**: 35-38.

Tovee, M. J., Rolls, E. T., Treves, A. and Bellis, R. P. (1993). "Information encoding and the responses of single neurons in the primary visual cortex." Journal of Neurophysiology **72**: 1049-1060.

Traub, R. D., Whittington, M. A. and Jefferys, J. G. (1997). "Gamma oscillation model predicts intensity coding by phase rather than frequency." Neural Computation **9**(6): 1251-1264.

Tsuda, I. (1992). “Dynamic Link of Memory.” Neural Networks **5**: 313-326.

Tsuda, I. (1992). “Dynamic Link of Memory: Chaotic Memory Map in Nonequilibrium Neural Networks.” Neural Networks **5**(2): 313-326.

Tversky, A. (1977). “Features of Similarity.” Psychological Review **84**: 327-352.

Tversky, A. and Hutchinson, J. W. (1986). “Nearest neighbor analysis of psychological space.” Psychological Review **93**: 3-22.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J. and Aston-Jones, G. (1999). “The Role of Locus Coeruleus in the Regulation of Cognitive Performance.” Science **283**: 549-.

Uttal, W. R. (1988). On Seeing Forms. Hillsdale NJ, Lawrence Erbaum Associates.

Vaadia, E., Ahissar, E., Bergman, H. and Lavner, Y. (1991). Correlated Activity of Neurons: A Neural Code for Higher Brain Functions? Neuronal Cooperativity. J. Kruger. Berlin, Springer-Verlag.

Van Essen, D. C., Anderson, C. H. and Olshausen, B. A. (1994). Dynamic Routing Strategies in Sensory, Motor, and Cognitive Processing. Large-Scale Neuronal Theories of the Brain. C. Koch and J. Davis. Cambridge, MA, MIT Press: 271-299.

Van Gelder, T. (1990). “Compositionality: a connectionist variation on a classical theme.” Cognition **14**(355-384).

van Leeuwen, C., Styvers, M. and Nooter, M. (1997). “Stability and intermittence in large-scale coupled oscillator models for perceptual segmentation.” Journal of Mathematical Psychology **41**: 319-344.

Verfaillie, K. and L., B. (1995). “A corpus of 714 full-color images of depth-rotated objects.” Perception and Psychophysics **57**: 925-961.

Wallis, G. (1994). Neural Mechanisms Underlying Processing in the Visual Areas of the Occipital and Temporal Lobes. Experimental Psychology, Oxford University: 226.

Wang, G., Tanaka, K. and Tanifuji, M. (1996). "Optical imaging of functional organization in the monkey inferotemporal cortex." Science **272**(1665-1668).

Wennekers, T. and Pasemann, F. (1996). "Synchronous Chaos in High Dimensional Modular Neural Network." International Journal of Bifurcation and Chaos **6**: 2055-2067.

Wiener, N. (1985). Time and Organization. Collected Works. P. Masani. **IV**.

Willet, P., Barnard, J. M. and Downs, G. M. (1998). "Chemical Similarity Searching." Journal of Chemical Information and Computer Science **38**: 983-996.

Wilson, H. R. and Cowan, J. D. (1972). "Excitatory and Inhibitory interactions in localized populations of model neurons." Biophysics Journal **12**: 1-24.

Wilson, R. and Knutsson, H. (1988). "Uncertainty and inference in the visual system." IEEE Trans. System Man and Cybern. **13**(305-312).

Wolfram, S. (1986). "Approaches to Complexity Engineering." Physica D **22**: 385-399.

Wolfram, S. (1986). Theory and Applications of Cellular Automata. Singapore, World Scientific.

Wolfson, H. J. and Yehezkel, L. (1992). Transformation Invariant Indexing. Geometric Invariance in Computer Vision. J. L. Mundy and A. Zisserman. Cambridge, MIT Press.

Wright, A. H. (1991). Genetic Algorithms for Real Parameter Optimization. Foundations of Genetic Algorithms. G. E. Rawlins, Morgan Kaufman. **3:** 205-218.

Wu, C. W. (1998). Global Synchronization in Coupled Map Lattices. IEEE Intl. Symposium of Circuits and Systems, Monterey CA, IEEE.

Wu, C. W. (1999). Synchronization in Arrays of Coupled Chaotic Circuits and Systems: Theory and Applications. Controlling Chaos and Bifurcations in Engineering Systems. G. Chen. Boca Raton, CRC Press**:** 1-27.

Wuensche, A. (1996). The Emergence of Memory: Categorization Far From Equilibrium. Toward a Science of Consciousness: The First Tucson Discussion and Debates. S. Hameroff, A. Kazniak and A. Scott. Cambridge, MIT Press.

Yao, Y. and Freeman, W. J. (1990). "Model of Biological Pattern Recognition with Spatially Chaotic Dynamics." Neural Networks **3**: 153-170.

Yoshizawa, S., Morita, M. and Amari, S. (1993). "Capacity of associative memory using a non-monotonic neuron model." Neural Networks **6**: 167-176.

Yuille, A. L. and Ullman, S. (1990). Computational Theories of Low-Level Vision. Visual Cognition and Action: An Invitation to Cognitive Science. D. N. Osherson, S. M. Kosslyn and J. M. Hollerbach. Cambridge, MIT Press. **2**.

Zhao, I., Macau, E. E. N. and Omar, N. (2000). "Scene segmentation of the chaotic oscillator network." International Journal of Bifurcations and Chaos **10**(7): 1697-1708.

Vita

David DeMaris was born and attended public schools in Salem, Illinois. He received a Bachelor of Science in Electrical Engineering from the University of Illinois in 1982 and a Master of Science in Engineering from the University of Texas at Austin in 1995. Since 1982 he has worked in VLSI design and design automation at IBM Burlington and Austin laboratories, where he is currently an employee of the Microelectronics division. In graduate studies at the University of Texas, his work in perceptual processing with coupled map lattices has attracted international attention, evidenced in speaking invitations in Europe and Japan and collaboration with mathematical psychologists in Australia. In fall 2000, he developed a biomedical engineering graduate course entitled Neurodynamics with Gerhard Werner, emphasizing nonlinear dynamics and synchronization processes.

His research interests include nonlinear dynamics, perception and cognition, pattern recognition; scientific visualization, computer architecture, VLSI design automation, and digital arts. His technical studies have been complemented by training in music composition and algorithmic visual art. He developed the compositional strategy dynamical symbolism, in which the interior forms of perception are made visible or audible. He has performed interactive computer visual art, live electronic music and ensemble settings, and composes and plays piano and trumpet in a jazz ensemble.

Permanent address: 1514 W. 9th Austin TX 78703

This dissertation was typed by the author.

The poet's eye, in a fine frenzy rolling,
Doth glance from heaven to earth, earth to heaven;
And, as imagination bodies forth
The forms of things, unknown, the poet's pen
Turns them to shapes, and gives to airy nothing
A local habitation and a name.

Shakespeare, A Midsummer Night's Dream, Act V, sc. 1

Solve et coagula, painting / assemblage by D. DeMaris, 2000