

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 (30 μm 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 structures¹¹. 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 junctions¹² 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 often 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.

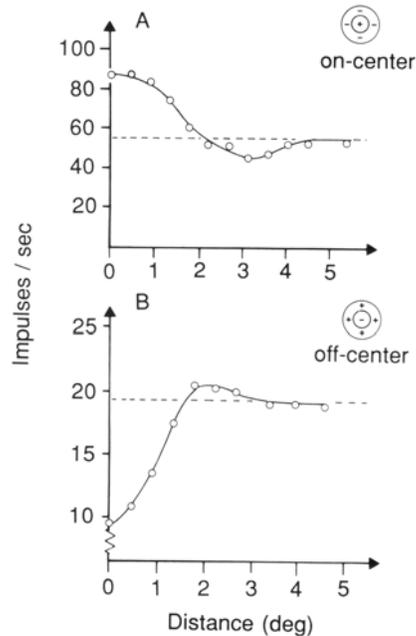


Fig. 3. The response of (A) on-center off-surround receptive fields and (B) off-center on surround in the retina as a function of the distance of a bar stimulus from the field center. From Maffei, L. (1968). "Inhibitory and facilitatory spatial interactions in retinal receptive fields." *Vision Research*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

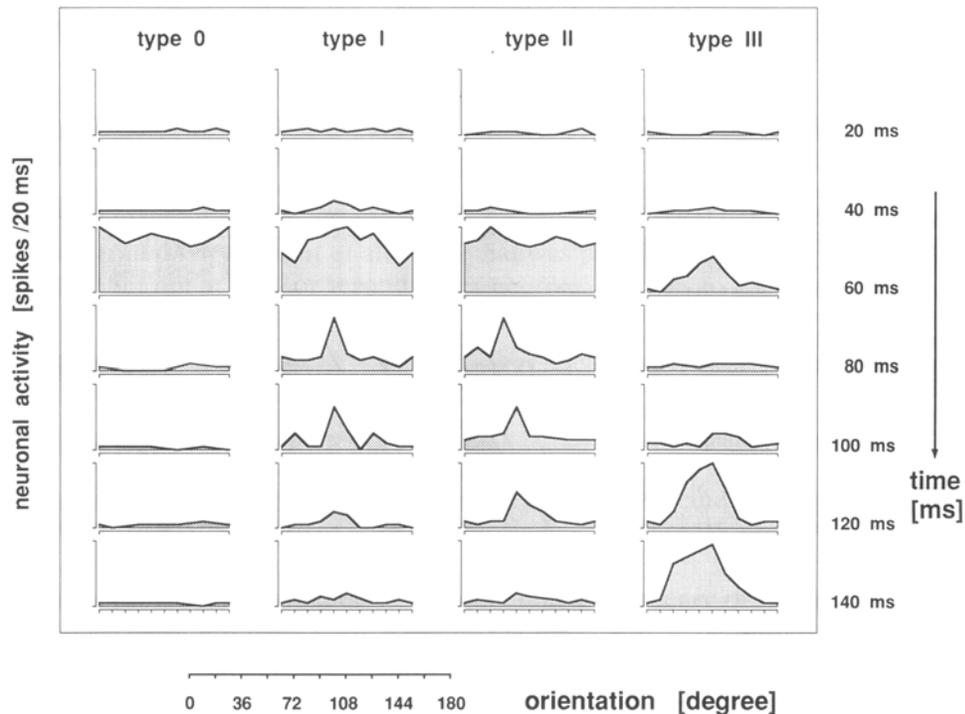


Fig. 4. Time varying receptive field structures. See text above for description From Dinse, H. R., K. Kruger, et al. (1991). Temporal Structure of Cortical Information Processing: Cortical Architecture, Oscillations, and Non-Separability of Spatio-Temporal Receptive Field Organization, in Neuronal Cooperativity.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

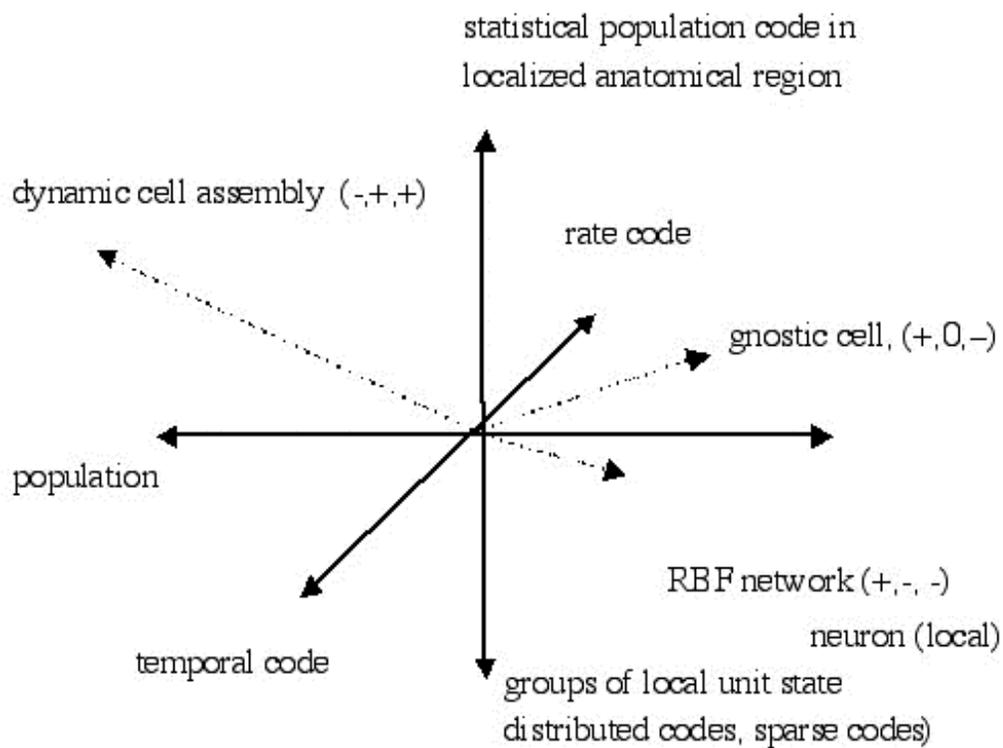


Fig. 5. A space of neural codes, with the dotted lines pointing to prototypical functional units. The x axis indicates whether coding occurs at single neurons or via groups. The y axis indicates whether, for group or population codes, whether the code is local or statistical. For the former, the code 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:

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

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).

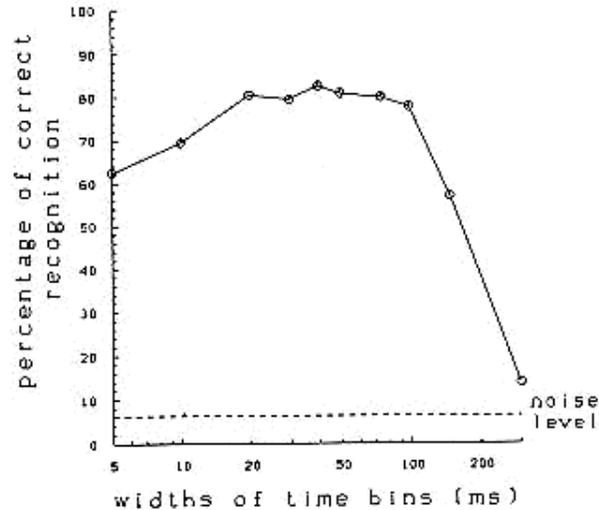


Fig. 6. The percentage of corrected responses as a function of the temporal resolution for averaging spike rates of a 30 channel spatio-temporal response vector for various time bins. Using a rate code (i.e. averaging over the entire 300 ms response interval) approaches a chance prediction level. From 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.

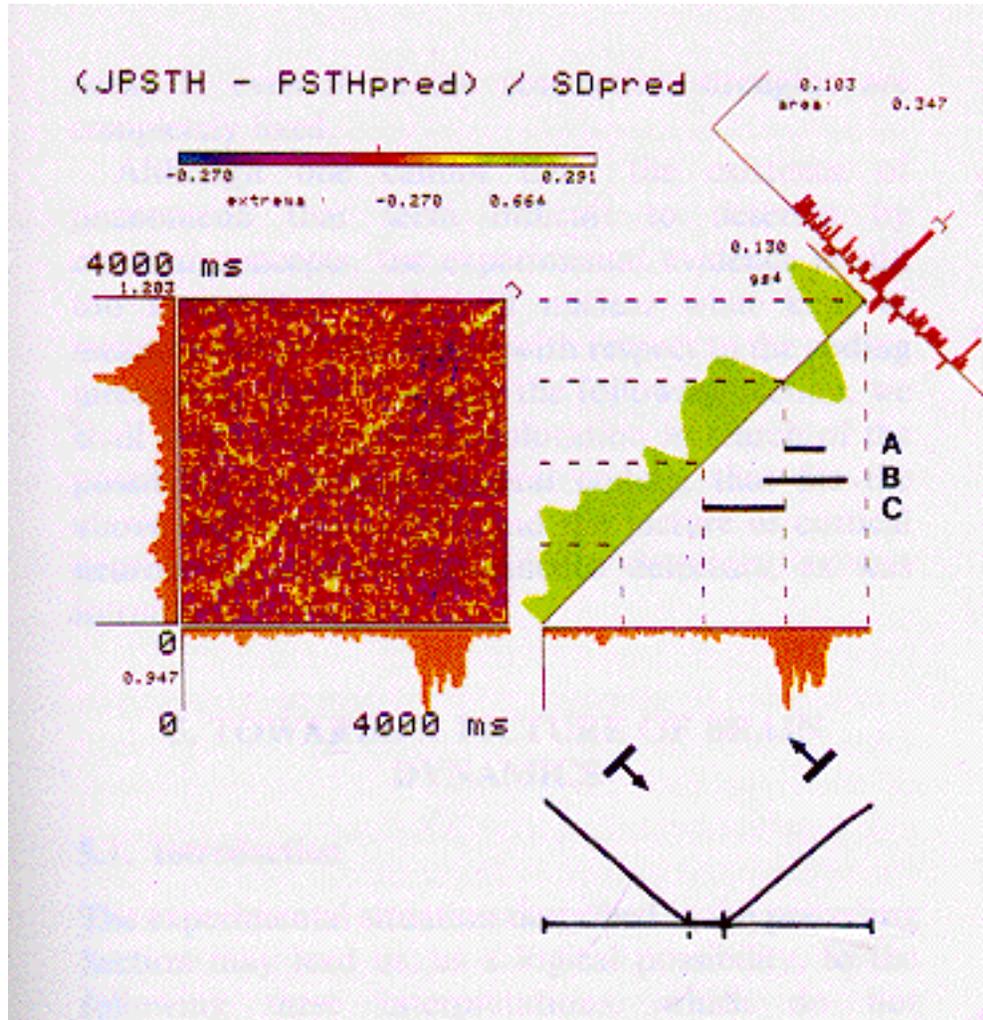


Fig. 7. Temporal changes in correlation structure for non-stationary signals. See text above for explanation. From Aertsen, A. M. H. J. and G. Gerstein (1991). Dynamic aspects of neuronal cooperativity: fast stimulus-locked modulations of effective connectivity. *Neuronal Cooperativity*. J. Kruger. Berlin, 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 systems¹⁴. 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.

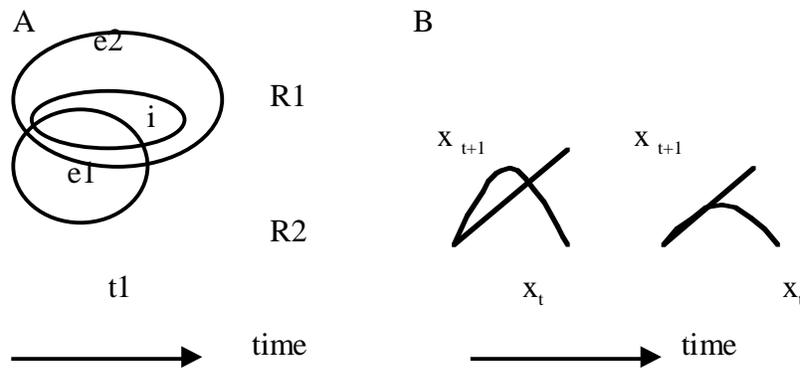


Fig. 8. Illustration of assemblies in two regions R1 and R2 cooperating to change dynamical parameters. A) The ovals represent coupled pools of randomly connected excitatory and inhibitory neurons. At time t_1 one of the pools decreases its activity. B) Each curve on the right indicate spike firing probabilities 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 t_1 .

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.

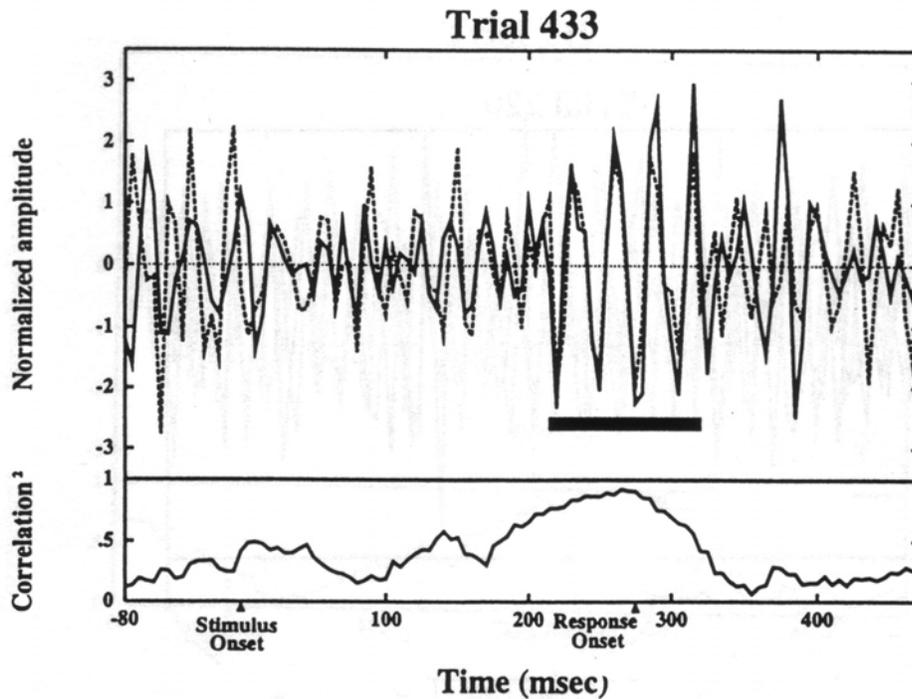


Fig. 9. Local field potential, filtered from 30-80 Hz, for a single trial, showing modulations in coherence of IT and striate cortex. In the framework of the Soca model described here, these modulations might correspond to transfer of phase synchronized contour border shapes from V1/V2 to IT (stimulus 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 mm² 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 terminating¹⁵ 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.

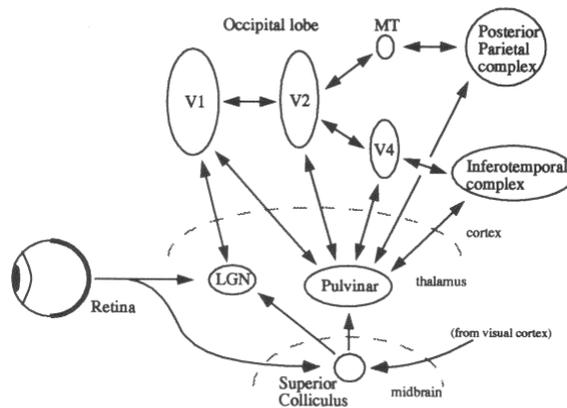


Fig. 10. Connections in the ventral pathway. From Van Essen, D. C., C. H. Anderson, et al. (1994). *Dynamic Routing Strategies in Sensory, Motor, and Cognitive Processing. Large-Scale Neuronal Theories of the Brain.* C. Koch and J. Davis. Cambridge, MA, MIT Press: 271-299. Reproduced with permission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

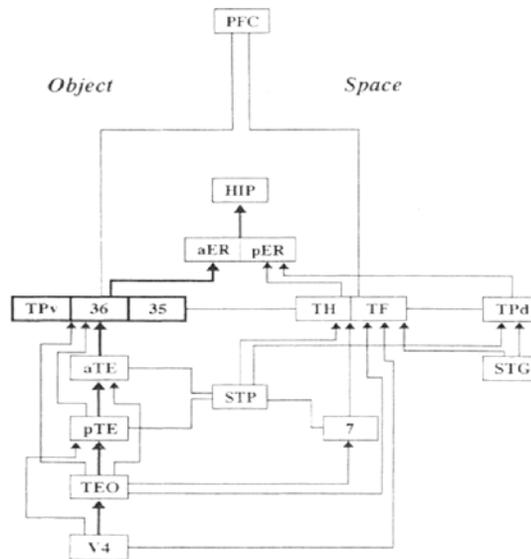


Fig. 11. Flow of information in subregions of IT cortex and connected cortical areas. pTE and TEO are implicated in processing of features, while aTE (anterior) is implicated in visual memory (ER = entorhinal, HIP = hippocampus, PFC. = prefrontal cortex). Connections to subcortical areas are not shown. 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.

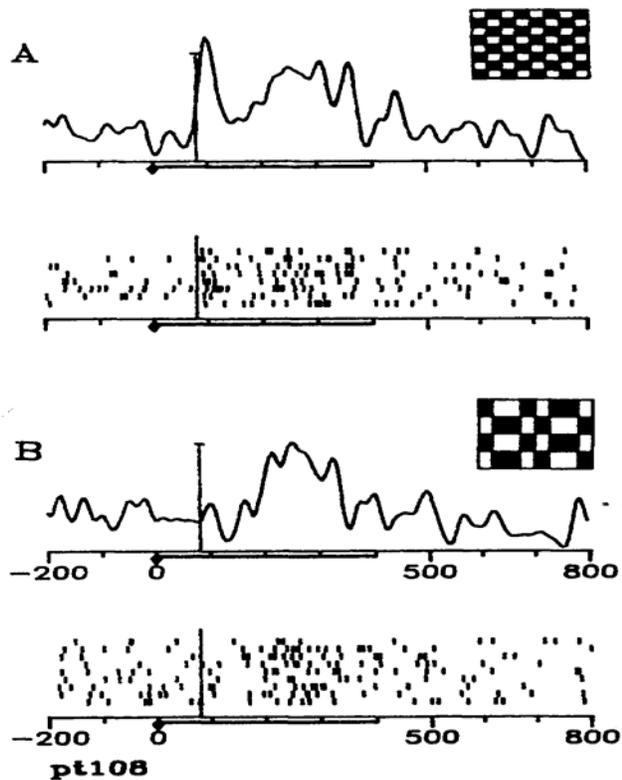


Fig. 12. The mean response of one neuron to two different Walsh patterns A and B, with rasters from individual trials plotted. Both excitation and inhibition are seen, in agreement with data of Gochin and coworkers described below. The bar height represents 50 spikes/ sec. The bar beneath shows the stimulus 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.

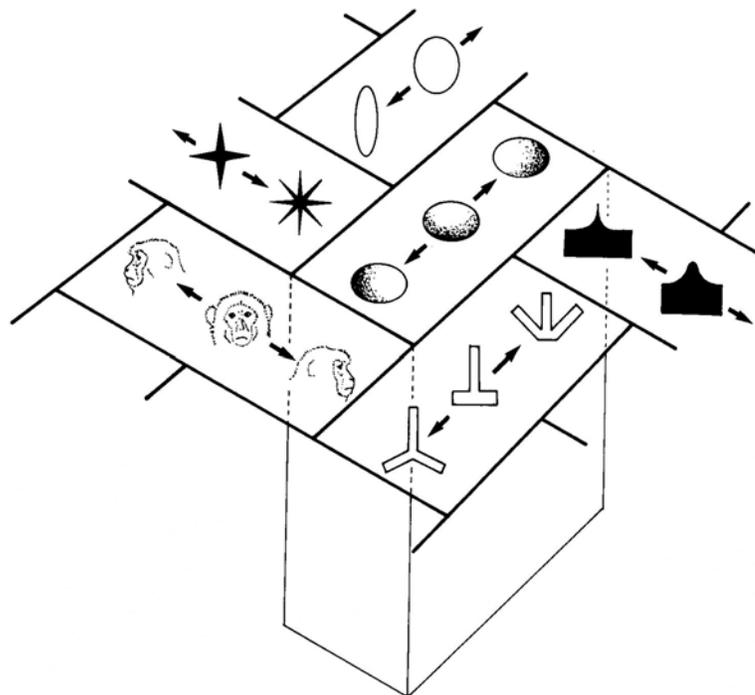


Fig. 13. A schematic view of overlapping critical feature columns hypothesized by Tanaka based on single neuron and optical recordings. From Tanaka, K. (1996). "Representation of visual features of objects in the inferotemporal cortex." *Neural Networks* 9(8): 1459-1475. Reproduced with permission of 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, Tsodyks 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 tasks¹⁶.

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-monotonic 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 monotonic *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 scale¹⁷ 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 monotonic 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 monotonic 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 (Natschlagler 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 KO 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 classically 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.