Representation space in a network of coupled chaotic oscillators with modulated synchronization

David DeMaris

Dept. of Electrical and Computer Engineering University of Texas at Austin, Austin TX , 78712, USA

Abstract Statistics over an ensemble of chaotic oscillators are used to perform recognition among a set of 3-D paperclip objects. The homogeneous local connections resemble cortical column structure, while dynamic aspects such as temporal coding, fluctuating correlations, and population coding accord with recent physiological observations. The coupled map formalism targets mesoscopic dynamics, of units consisting of mixed excitatory-inhibitory populations. This paper focuses on aspects of the representation space. Multidimensional scaling reveals the clustering of objects in the representation space, and distortions of the clustering resulting from small shifts in one of the dynamical parameters of each recognizer are examined.

Keywords: Object recognition, Population code, Neuronal transient, Synchronized chaos, Dynamical recognizer

Introduction

A regular array of discrete-time nonlinear oscillators with recurrent self-connectivity and local diffusive coupling, known as coupled map lattice or CML, was used to perform pattern recognition, extending the concept of a dynamical recognizer [13] to high dimensional recurrent networks. The related problems of similarity and stimulus equivalence are addressed via cooperative processing and partial synchronization after rapid expansion from a constrained subspace. A representation space is formed whose dimensions are partition cells (intervals of the state space) of the each units' state space. Spatio-temporal patterns formed by an expanding wavefront on the object boundaries are co-determined by intrinsic time-varying dynamics of the oscillators and the forms presented as initial conditions. The evolving state of the entire lattice is sampled at a particular readout time, resulting in a population code over all units. The current system is a hybrid, with algorithmic processes for learning and statistical recognition, but with the construction of the reprentation space employing a regular lattice of chaotic oscillating units with a local mean field coupling as a dynamical recognizer. An ensemble of such CML recognizers run in parallel during recognition, comparing the resulting statistics of each recognizer against the previously recorded statistics of its preferred object, can classify known or unknown objects, presented in familiar or novel views.

Aspects of this system were reported previously, focusing on the nature of the population code, dynamics, and correlations with physiological observations [2]. The discussion is expanded here to reflect improved understanding of the network functioning, and to situate the work in the context of other work on neural computation with transients. In addition, preliminary studies of clustering in the representation space and the effect of parameter deviations on recognition rates and other measures of the representation space are presented.

Simulation methods and results

The CML formalism targets medium scale (mesoscopic) dynamics between locally synchronized, mixed excitatory-inhibitory populations in column-like units *without explicitly modeling the sub-populations*. The logistic map iterated at each site in a 75x75 square lattice was originally introduced as a model of population dynamics; the specific form used here produces values ranging from -1 to +1, supporting a neuronal population interpretation with a baseline activity (value 0) and suppressed (- range) or enhanced (+ range) activity. A long modeling tradition has stressed the possibility that collective behavior in column like units may arise from such population dynamics [15], [6]; micro-circuit level parameters such as the ratio of excitatory and inhibitory connections and threshold serve effectively as a

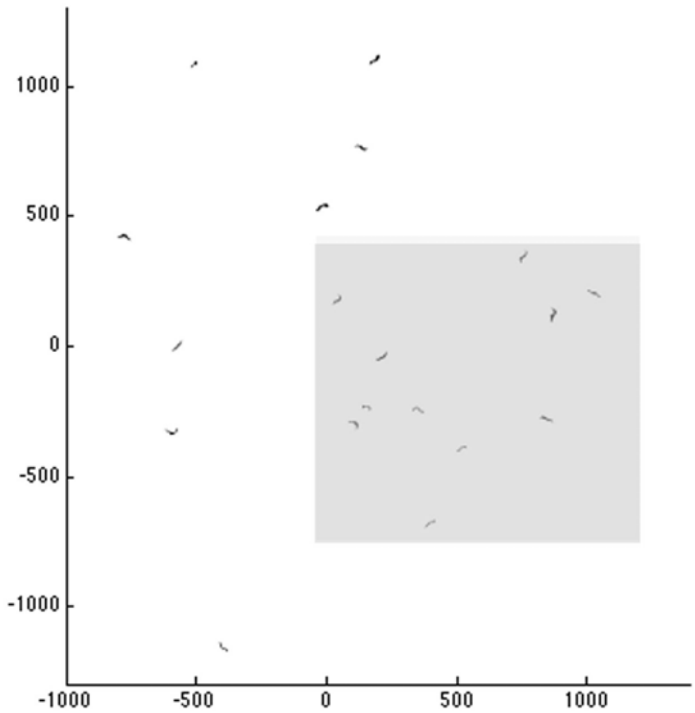
bifurcation or nonlinearity parameter. Such behavior is taken here as the fundamental operating unit, with local mean field operations between such units providing the substrate for invariance computations. The equations for each iteration of the coupled maps, and the large scale temporal structure of two parameter epochs are given as an appendix.

For each object, a dynamical recognizer is trained which attempts to transform all the presented views to an identical set of statistics measured over 64 bins in the activity space (range -1 to +1). During recognition, this statistical view of the network state is measured over the entire lattice, so it is the spatio-temporal interactions between such mesoscopic scale units, rather than the constituent neurons of each unit, which are the primary computational substrate. Any particular column-like unit will not produce the same trajectory when evolving from a different view of the target object, but the states of all units at some designated readout time must reach the same distribution. It is interesting to note that certain lattice sites may, in the course of the network dynamics, reach and remain in the enhanced activity regime, while others fluctuate. An observer with only a single unit probe, seeking confirmation for grandmother cell or sparse distributed representations, would naturally interpret such local high activity patterns as confirming evidence.

Evolutionary computation is used to train the CML constituting each recognizer, but without explicit testing on the object recognition task until training of all recognizers is complete. Instead, an objective function is formulated based solely on desirable characteristics of the distribution sampled after the candidate control parameters (i.e. bifurcation, coupling, and number of iterations) transform the target image. This objective function must balance two goals: 1) tight clustering of different views of an object and 2) separation of classes corresponding to various views of other objects. A view normalization strategy (similar to that used in a radial basis function (RBF) ensemble [4] approach to object recognition) serves the tight clustering goal, with a cross-entropy term comparing current output with previously stored representations enforcing separation of objects in the representation space. Details of this objective function are presented in [3].

Another noteworthy aspect of the training is that classifiers are formed sequentially, based on the known signature distributions of previously learned objects. In other systems with comparable or higher recognition rates, the statistics of the entire "world" of objects must be known *a priori* to make optimal choices for either the center views in an RBF formulation [4], or for feature weights in a rich-feature-space statistical matching system [11]. While relearning or refinement of representations is possible and has biological support, the sequential learning of objects with immediate feedback on success in discrimination is more ecologically realistic than the global world knowledge used in the cited computer vision frameworks. There is evidence from single unit studies in inferotemporal cortex [12] that the representations formed, as measured during the delay (priming) period of a match to sample task, correlate more strongly with temporal proximity during learning than with features of the object. This observation lends support to the view based theories, and closely fits the representation strategy described here. The task chosen for study is the recognition of paperclip objects enhanced with added geons. Paperclip objects have been widely studied in psychophysics and neurophysiology. This class of data is challenging to human subjects, who exhibit pair matching error rates ranging from 30% to near 100% depending on training circumstances [14]; thus the problem is non-trivial. With training on all seven available views separated by 30 degrees, the present system achieved a recognition rate of 88% for a set of twenty paperclip objects, degraded to 60% correct for forty objects. Note that the system here is required to identify which object is presented, not simply whether two successive presentations are the same or different.

A 2-D visualization of the clustering of objects (obtained via multidimensional scaling over average distributions produced by each recognizer) suggests that a reasonable clustering in this representation space emerges from the constraints of this invariance problem and the requirement to spread the representations for individual, sequentially presented objects over the space (fig 1a, 1b). Similar results regarding the emergence of reasonable categorization and clustering from the constraint to map different views to the same output have been shown for RBF ensemble networks [4].



a.

b.

Fig 1 Multidimensional scaling of the representation space. a) Each of 20 objects (in the 0 degree view) is mapped to its location after projecting from 64 to two dimensions. The point in 64 space is computed by the fraction of oscillating lattice sites occupying each of 64 equally sized bins at the end of the second parameter epoch (e.g. time t1+t2). The units are dimensionless quantities. b) The gray area of a) is enlarged. The clustering in this representation space strikes a compromise between aspects of shape such as the number and location of extra geons and the set or orientations present in the various views under z axis rotation.

In connectionist systems with sigmoidal activation units, performance degradations due to network “lesions” have been studied for many years now [7]. For recurrent networks with nonlinear unit transfer functions corresponding to larger scale distributed networks, the dynamical disease framework [1] serves to assess cognitive or perceptual disturbances. In this case, the dynamical parameters (bifurcation and coupling) may be shifted from baseline or modulated values, with corresponding degradation in network performance. These shifts would correspond to biological syndromes like lesions in remote connected areas responsible for slow wave interactions related to the modulated synchronization parameters, or to disturbances in microcircuit parameters regulating baseline values.

The distortion of the representation space for one such shift is quantified in Table 1. For a given shift in a single parameter (c1, the inter-unit coupling of the first parameter epoch) several measures of distortion in the representation space are tabulated below, along with the corresponding reduction in recognition rates. Further study would be required to ascertain whether there are systematic trends in certain parameter distortions and the representation space, and whether such trends could be helpful in validating the theory of object representation. It might be possible to improve recognition robustness by examining such distortion measures during learning, and favoring solutions with greater robustness to such parameter distortion. While not critical for digital implementations for pattern recognition, this robustness would be important for analog implementations of this recognition principle, which would have greater sensitivity to process variations and noise. It is also of interest in understanding the limits of these operating principles in biological systems. Treating the control parameter as a signals, these simulations indicates that noise levels of 0.2-2% in the "memory" signal which controls recognition in the network produce serious disruptions. This sensitivity is perhaps the strongest theoretical argument against the plausibility of this operating principle.

Table 1: Measures of Representation Space Distortion

parameter distortion	recognition rate	sum of MDS position shifts	nearest / mean similarity	sum of inter-object distances
c1=c1 + 0 (mean .525)	87%	0	56.3	73.3
c1=c1+.001	66%	0.058	56.6	72.5
c1=c1-.001	78%	0.030	55.3	73.1
c1=c1+.005	49%	0.089	54.3	72.0
c1=c1-.005	59%	0.090	54.1	72.5
c1=c1+.01	36%	0.108	52.5	71.8
c1=c1-.01	38%	0.128	54.2	72.8

Discussion

While building on past theory [16],[9] and in accord with contemporary proposals [5] on high dimensional coupled chaotic systems, the present network makes substantive breaks with other systems. The most important is to introduce two parameter epochs, with a sharp change in the parameters during the recognition cycles, rather than stationary or smoothly changing dynamics. Over many evolutionary learning trials, these recognizer ensembles were found to achieve shorter recognition times and higher peak recognition rates using two parameter epochs, rather than stationary chaotic dynamics [3]. However, stationary networks can also exceed 80% recognition rates. In the early reports on this network, the operating principle motivating the two epochs was described as "synchronization opponents", as it was anticipated that solutions found by genetic search would follow a certain pattern. While the synchronization properties of the network are indeed modulated, the original intuition that the stages would be monotonically desynchronized (via high chaos level and low coupling), followed by partial synchronization (lowered chaos and/or coupling) is less complex than the actual behavior exploited by the genetic algorithm. Phase synchronization of the oscillators, as reflected by strong peaks in the instantaneous statistics, and the entropy of the distribution at any time step, is seen to fluctuate within both epochs. The increased performance with tow epochs appears to arise from increasing the reachable states during the transient.

It also breaks with typical recurrent networks by concentrating on the *transient* regime of dynamics rather than on attractors. The first stage distributes the states out of a confined subspace (the high contrast image used as an initial condition), to cover the state space of the system, in only a few steps. While the second stage ultimately concentrates the states in a subspace to facilitate pattern recognition, parameters (high coupling and parameters leading to stable periodic dynamics) which would induce rapid synchronization must be avoided. Otherwise, all initial conditions will map near that particular recognizer characteristic distribution.

Pattern recognition [8] or more general computation [10] with transients and readout units rather than stable states (attractors) is an emerging trend in computational neuroscience. The present work is distinguished by an assumption that memory of a useful computation, such as approximate view invariance, is not achieved strictly through local microcircuit parameter adjustments. Rather, the memory for the recognition process may be achieved by the interaction of slow bursts of long range excitatory connections from limbic memory areas projecting into cortical regions performing transformations to invariant population codes.

In the perceptual models of Freeman and colleagues, recognition is identified with dimensionality reduction, evidenced by the constraint of the system dynamics to a subspace. The computational approach here differs by beginning with a constraint assumed to be imposed from an externally feedforward coupled network performing segmentation. This highlights the fit of the present model and theory with the traditional role ascribed to the intermediate (V2) visual areas, the extraction of boundaries to form a primal sketch. The primal sketch formation is *required* to constrain states in a restricted subspace during a perceptual frame, while in RBF theory it plays no role; V1 like receptor field inputs are directly fed to the ensemble. Thus, this model incorporates elements of feedforward processing - but with feedforward links between recurrent processes, with long range inter-regional reactions from memory integral in their processing. This type of structure fits the known anatomy of the ventral pathway more closely than the other feedforward network models cited.

Future work in the family of computations is expected to take several forms. One is to move beyond the separate recognizer framework to a unified network model more closely resembling likely operations; the present system stands chiefly as a proof of concept that transient interactions in coupled chaotic units can operate on difficult pattern recognition tasks. Another is to seek methods which reduce the sensitivity of recognition on exact parameter values. The incorporation of noise into the learning and recognition process is one possible avenue to investigate.

Acknowledgements

Paperclip images courtesy of Michael Tarr, Dept. of Psychology, Brown University. Multidimensional Scaling Matlab functions courtesy of Malcom Slaney and Interval Research. Thanks to Gerhard Werner, Walter Freeman, and Paul So for useful discussions.

Appendix I. Equations used in the dynamical evolution of patterns

The short range or local mean field coupled logistic map is used here; the computation is divided into two steps; first a diffusion step is applied:

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

where S_d is a temporary array holding the state of the diffusion computation, t is the current time step, x,y are the spatial indices of the pixel array S at the center of the diffusion neighborhood, S is the state variable at each pixel of the array restricted to the range [-1.0 to 1.0], and c is the coupling constant restricted to the range [0.0 to 1.0]. The scaling term 1-c is a squashing function, preventing the sum of the surround from exceeding the stable range of the map; in microcircuit parameters it may be interpreted as a gain control limiting the instantaneous influence of a highly active neighborhood.

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

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

where *S, t, x,* and *y* are as above and where *b* is the bifurcation parameter, restricted to the range [0.0 < b < 2.0]. Initial and intermediate states S are restricted to the range [-1.0 < S < 1.0].

The basic iterations described above are applied at every lattice site, and are imbedded in an "outer loop" computation with the parameters b and c changing as a step function at time t1 and t2. This is interpreted in neural terms as a regularly changing modulation of the local and short range synchronization dynamics in a cortical region, induced by bursting behavior from long range excitatory connections from memory areas.

Each parameter epoch consists of a triple {*b*, *c*, *t*}, where *t* is number of iterations in the stage. Two such frames are applied in turn. During evolutionary search for the parameter sets, each stage is constrained to 7 iterations. The total iterations required to create the representation space ranges from 6 to 14 iterations, with a mean of 11.3 in the first learning trial on the 39 objects in the Tarr paperclip+ geon set.

References

[1] J. Belair, L. Glass, U. an der Heiden, and J. Milton, Dynamical disease: Identification, temporal aspects, and treatment strategies of human illness, Chaos 5 (1995) 1-7.

[2] D. DeMaris, Distributions of Sampled Attractor Basin Transients: a population code for object representation, Neurocomputing (2001) 547-554.

[3] D. DeMaris and B.F. Womack, Recognition of Objects Rotated in Depth Using Partial Synchronization of Chaotic Units, in International Joint Conference on Neural Networks '01 (IEEE / INNS, Washington, D.C., 2001).

[4] S. Duvdevani-Bar and S. Edelman, Visual recognition and categorization on the basis of similarities to multiple class prototypes, Intl. J. of Computer Vision 33 (1999) 201-228.

[5] W.J. Freeman, R. Kozma, and P.J. Werbos, Biocomplexity: adaptive behavior in complex stochastic dynamical systems, Biosystems 59 (2001) 109-123.

[6] E.M. Harth, T.J. Csermely, B. Beek, and R.D. Lindsay, Brain functions and neural dynamics, Journal of Theoretical Biology 26 (1970) 93-120.

[7] G.E. Hinton and T. Shallice, Lesioning an attractor neural network: investigation of aquired dyslexia., Psychological Review 98 (1991) 74-95.

[8] J. Hopfield and C. Brody, What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration, Proceedings of the National Academy of Sciences 98 (2001) 1282-1287.

[9] K. Kaneko, Spatiotemporal chaos in one and two dimensional coupled map lattices, Physica D 37 (1989) 1-47.

[10] W. Maas, T. Natschlager, and H. Markram, Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations, (preprint) (2001) .

[11] B.W. Mel, SEEMORE: Combining Color, Shape, and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition, Neural Computation 9 (1997) 777-804.

[12] Y. Miyashita, Neuronal correlate of visual associative long-term memory in the primate cortex, Nature 335 (1988) 817-820.

[13] J.B. Pollack, The Induction of Dynamical Recognizers, Machine Learning 7 (1991) 227-252.

[14] M.J. Tarr, H.H. Bulthoff, M. Zabinski, and V. Blanz, To what extent do unique parts influence recognition across changes in viewpoint?, Psychological Science 8 (1997) 282-289.

[15] T. Wennekers and F. Pasemann, Synchronous Chaos in High Dimensional Modular Neural Network, International Journal of Bifurcation and Chaos 6 (1996) 2055-2067.

[16] Y. Yao and W.J. Freeman, Model of Biological Pattern Recognition with Spatially Chaotic Dynamics, Neural Networks 3 (1990) 153-170.