Modeling large cortical networks with growing self-organizing maps

James A. Bednar, Amol Kelkar, and Risto Miikkulainen

Department of Computer Sciences, The University of Texas at Austin, Austin, TX 78712 {jbednar,amol,risto}@cs.utexas.edu

Abstract

Self-organizing computational models with specific intracortical connections can explain many features of visual cortex. However, due to their computation and memory requirements, it is difficult to use such detailed models to study large-scale object segmentation and recognition. This paper describes GLISSOM, a method for scaling a small RF-LISSOM model network into a larger one during self-organization, dramatically reducing time and memory needs while obtaining equivalent results. With GLISSOM it should be possible to simulate all of human V1 at the single-column level using existing workstations. The scaling equations GLISSOM uses also allow comparison of biological maps and parameters between individuals and species with different brain region sizes.

Key words: Self-organization, Cortical modeling, Vision, Orientation maps, Growing networks

1 Introduction

Self-organizing computational models have shown that input-driven development can explain much of the topographic organization of the visual cortex, such as retinotopy and orientation preference, as well as many functional properties, such as short-range contour segmentation and binding (see [12] for review). However, other important phenomena have remained out of reach because they require too much computation time and memory to simulate. These phenomena, such as longrange visual contour and object segmentation and integration, are thought to arise out of the specific lateral interactions between large numbers of neurons over a wide spatial area [5]. Simulating such behavior requires an enormous number of specific, modifiable connections, which is computationally infeasible using existing models. In this paper we present a general scaling approach called GLISSOM

Supported in part by the National Science Foundation under grants IRI-9309273 and IIS-9811478. We are grateful to Lisa Kaczmarczyk for comments. Published in *Neurocomputing* 44–46:315–321, 2002.



Fig. 1. Architecture of the RF-LISSOM network. A small RF-LISSOM network and retina are shown, along with connections to a single neuron (shown as a large circle). The input is an oriented Gaussian activity pattern on the retinal ganglion cells (shown by grayscale coding); the LGN is bypassed for simplicity. The afferent connections form a local anatomical receptive field (RF) on the simulated retina. Neighboring neurons have different but highly overlapping RFs. Each neuron computes an initial response as a scalar (dot) product of its receptive field and its afferent weight vector, i.e. a sum of the product of each weight with its associated receptor. The responses then repeatedly propagate within the cortex through the lateral connections and evolve into activity "bubbles". After the activity stabilizes, weights of the active neurons are adapted using a normalized Hebbian rule.

that allows much larger networks to be simulated faithfully in a given time and amount of memory, thereby making larger-scale simulations possible.

2 Growing LISSOM (GLISSOM)

GLISSOM is based on the RF-LISSOM computational model (Receptive-Field Laterally Interconnected Synergetically Self-Organizing Map; figure 1) [10, 11]. RF-LISSOM focuses on the two-dimensional topographic organization of the cortex, modeling a cortical area as an $N \times N$ sheet of neurons and the retina as an $R \times R$ sheet of ganglion cells. Neurons receive afferent connections from broad patches of radius r_A on the retina, and receive lateral excitatory and inhibitory connections from nearby and more distant neurons in the cortex (radii r_E and r_I , respectively). The connection weights are initially random or isotropic, and are subsequently organized through an unsupervised Hebbian learning process using visual input. Weak connections are eliminated periodically, resulting in patchy lateral connectivity similar to that observed in the visual cortex. Due to the large number of connections involved in a realistic map, this self-organization process is very computation and memory intensive.

GLISSOM's computation and memory savings are achieved by starting simulations with a small network and smoothly scaling up the network as it self-organizes. This approach is effective without sacrificing accuracy because: (1) self-organizing models tend to have peak computational and memory requirements at the beginning of training [7], and (2) self-organization tends to proceed in a global-to-local fashion, with large-scale order established first, followed by more detailed local self-organization [4]. Thus small maps, which are much quicker to simulate and

take less memory, can be used to establish global order, with larger maps used only to achieve more detailed structure.

GLISSOM relies on a set of parameter scaling equations developed for RF-LISSOM networks, and on an interpolation procedure developed to convert an existing network into a larger one [7]. The scaling equations and interpolation procedure were derived by treating a cortical network as a finite approximation to a continuous map composed of an infinite number of units [1, 9]. Under such an assumption, networks of different sizes represent coarser or denser approximations to the continuous map, and any given approximation can be transformed into another by (conceptually) reconstructing the continuous map and then resampling it. Such resampling is equivalent to the smooth bitmap scaling done by computer graphics programs, and extends a similar procedure proposed by Rodriques and Almeida [8] for the more abstract Self-Organizing Map architecture.

For each GLISSOM scaling step from a cortex of size N_o into one of size N, we first scale the cortical architecture parameters:

$$r_{\rm E} = \frac{N}{N_o} r_{\rm E_o}, \, \alpha_{\rm E} = \frac{r_{\rm E_o}^2}{r_{\rm E}^2} \alpha_{\rm E_o}, \, r_{\rm I} = \frac{N}{N_o} r_{\rm I_o}, \, \alpha_{\rm I} = \frac{r_{\rm I_o}^2}{r_{\rm I}^2} \alpha_{\rm I_o}, \, D_{\rm I} = D_{\rm I_o} \frac{r_{\rm I_o}^2}{r_{\rm I}^2}, \tag{1}$$

where $r_{\rm E}$ and $r_{\rm I}$ are the lateral excitatory and inhibitory radii, $\alpha_{\rm E}$ and $\alpha_{\rm I}$ are the lateral connection learning rates, and $D_{\rm I}$ is the connection pruning cutoff level at a given iteration. These equations specify how scaling should be done to preserve large-scale structural and functional properties. If used to construct a separate RF-LISSOM simulation, the scaled parameters would result in a final map similar to but larger than the original map [7]. The above equations scale only the cortex size, but in RF-LISSOM the retina size and total area can be scaled similarly [7].

Once the new architecture parameters have been determined, we generate the connection strengths in the new network from the strengths in the original, partiallyorganized network. Each connection weight in the new network is calculated as a normalized proximity-weighted linear combination of the connection strengths from the corresponding region of the original network (for details see [7]). This smooth scaling allows neuron density to be increased while keeping the large-scale structural and functional properties constant, such as the organization of the orientation preference map. In essence, the large network is grown in place, thereby minimizing the computational resources required for simulation.

3 Results

The two requirements for GLISSOM's dynamic scaling are (1) the maps it produces should be equivalent to those produced by RF-LISSOM, and (2) the computation time and memory should be significantly reduced. Figures 2 and 3 show that both of these requirements are met. For example, for the simulations in figure 2 on a 600MHz Pentium III workstation, GLISSOM takes 1.9 hours and 60MB, while RF-LISSOM requires 5.1 hours and 317MB. Thus GLISSOM is faster by a factor



Fig. 2. GLISSOM produces final orientation maps that match those of RF-LISSOM. Plots (a) and (b) show orientation maps measured for a 144×144 RF-LISSOM network near the beginning and at the end of self-organization. Plots (c) and (d) show the corresponding maps for a GLISSOM network that gradually scaled a 54×54 cortex up to 144×144 . Each neuron in each plot is grayscale-coded according to the orientation of its preferred visual stimulus, with shading varying from black (horizontal) to light gray (vertical). An example neuron is marked with a white square in each plot; the lateral inhibitory connections of this neuron are outlined in white around it. Most neurons in the early maps have random, weak orientation preferences, although small patches of orientation-selective neurons are emerging at similar locations in (a) and (c). The selective patches are driven by the stream of input patterns (which was identical for both networks), and are insensitive to the different random initial weight values in each map [7]. After self-organization, the final maps (b) and (d) are nearly identical, as are the pruned lateral connection patterns within them. For both final maps, nearby neurons have similar orientation preferences, short-range lateral connections target neurons of all orientations, and long range connections preferentially target neurons with similar orientations (as found in experimental animals [2]).

of two while using only one-fifth of the memory. At the same time, the resulting orientation maps are both qualitatively and quantitatively similar (as measured by e.g. the average difference between corresponding weight values [7].) Importantly, the relative speedup and the memory savings increases with larger network sizes (figure 3), which means that GLISSOM makes much larger networks practical.

Based on the connectivity patterns observed in these simulations, the scaling equations allow us to obtain a rough estimate of how large a simulation would be needed to approximate the density, area, and connectivity of the biological cortex. The result of these calculations is that GLISSOM should be able to represent all of human primary visual cortex (V1) at the single-column level using approximately 450 megabytes of RAM, which is within the reach of current desktop workstations [7]. Existing algorithms with specific connections, such as RF-LISSOM, would require more than 135 gigabytes of RAM, necessitating the use of large supercomputers.

4 Discussion

The simulation results show that growing realistic cortical maps during self-organization can reduce simulation time and memory usage, making the study of largescale phenomena tractable. Similar maps are developed in each case, as long as



Fig. 3. For large networks, GLISSOM uses significantly less memory and time. Summary of results from fourteen simulations on the same 600MHz Pentium III workstation, each running either RF-LISSOM (with the fixed size N indicated in the x-axis) or GLIS-SOM (with a starting size of N = 36 and the final N indicated in the x-axis). Each point on each curve represents one run; variance between runs was negligible. (*a*) RF-LISSOM keeps the peak number low throughout training. (*b*) The total simulation times for RF-LISSOM also increase dramatically for larger networks, but because GLISSOM has fewer connections to process, its computation time increases only modestly for the same range of N. Thus, given fixed computational resources, GLISSOM can simulate a much larger network.

the GLISSOM starting size is sufficiently large [7]. Scaling equations similar to those in section 2 can also be developed for other similar models, and then smooth interpolation can be used to scale them up during self-organization. Essentially, the GLISSOM method allows a fixed model to be adapted into one that grows in place, by adding scaling equations and an interpolation algorithm.

The scaling equations GLISSOM uses also give insight into how the corresponding quantities may differ between different individuals, between different species, and during development (cf. [6]). The equations specify how the biophysical correlates of the model parameters should differ between any two otherwise-similar brains (or areas thereof) that differ in size. The discrepancy between the actual parameter values and those predicted by the scaling equations can give insight into the difference in function and performance of different individuals and species.

Future work can utilize GLISSOM to study larger-scale and more complex phenomena at a detailed level. For instance, the visual tilt illusion is thought to occur through orientation-specific lateral interactions between spatially-separated stimuli [3]. Memory constraints limited the lateral connection length in previous RF-LISSOM simulations, and thus such long-range interactions could not be studied. With GLISSOM it should be possible to integrate both the spatial and orientation aspects of the tilt illusion by self-organizing a network with a large area and the full range of lateral connectivity found in the cortex.

5 Conclusion

The GLISSOM method should allow detailed laterally-connected cortical models like RF-LISSOM to be applied to much more complex, large-scale phenomena. The scaling method also provides insight into the cortical mechanisms at work in organisms with brains of widely different sizes. Thus GLISSOM can help explain brain scaling in nature as well as helping scale up brain simulations.

References

- S.-I. Amari, Topographic organization of nerve fields, *Bulletin of Mathemati*cal Biology 42 (1980) 339–364.
- [2] W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick, Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex, *Journal of Neuroscience* 17 (1997) 2112–2127.
- [3] R. H. S. Carpenter and C. Blakemore, Interactions between orientations in human vision, *Experimental Brain Research* 18 (1973) 287–303.
- [4] B. Chapman, M. P. Stryker, and T. Bonhoeffer, Development of orientation preference maps in ferret primary visual cortex, *Journal of Neuroscience* 16 (1996) 6443–6453.
- [5] C. D. Gilbert, A. Das, M. Ito, M. Kapadia, and G. Westheimer, Spatial integration and cortical dynamics, *Proceedings of the National Academy of Sciences*, USA 93 (1996) 615–622.
- [6] J. H. Kaas, Why is brain size so important: Design problems and solutions as neocortex gets bigger or smaller, *Brain and Mind* 1 (2000) 7–23.
- [7] A. Kelkar, J. A. Bednar, and R. Miikkulainen, Modeling large cortical networks with growing self-organizing maps, Technical report, Department of Computer Sciences, The University of Texas at Austin, (2000), Technical Report AI-2000-285.
- [8] J. S. Rodriques and L. B. Almeida, Improving the learning speed in topological maps of patterns, in: *Proceedings of the International Neural Networks Conference* (Paris, France) (Kluwer, Dordrecht; Boston, 1990) 813–816.
- [9] A. C. Roque Da Silva Filho, *Investigation of a Generalized Version of Amari's Continuous Model for Neural Networks*, Ph.D. thesis, University of Sussex at Brighton, Brighton, UK, (1992).
- [10] J. Sirosh and R. Miikkulainen, Cooperative self-organization of afferent and lateral connections in cortical maps, *Biological Cybernetics* 71 (1994) 66–78.
- [11] J. Sirosh, R. Miikkulainen, and J. A. Bednar, Self-organization of orientation maps, lateral connections, and dynamic receptive fields in the primary visual cortex, in: J. Sirosh, R. Miikkulainen, and Y. Choe, eds., *Lateral Interactions in the Cortex: Structure and Function* (The UTCS Neural Networks Research Group, Austin, TX, 1996) Electronic book, ISBN 0-9647060-0-8, http://www.cs.utexas.edu/users/nn/web-pubs/htmlbook96.
- [12] N. V. Swindale, The development of topography in the visual cortex: A review of models, *Network Computation in Neural Systems* 7 (1996) 161–247.