ARTICLE Communicated by Peter Dayan

Cortical Map Reorganization as a Competitive Process

Granger G. Sutton III
James A. Reggia
Steven L. Armentrout
C. Lynne D'Autrechy
Department of Computer Science, A. V. Williams Bldg., University of Maryland, College Park, MD 20742 USA

Past models of somatosensory cortex have successfully demonstrated map formation and subsequent map reorganization following localized repetitive stimuli or deafferentation. They provide an impressive demonstration that fairly simple assumptions about cortical connectivity and synaptic plasticity can account for several observations concerning cortical maps. However, past models have not successfully demonstrated spontaneous map reorganization following cortical lesions. Recently, an assumption universally used in these and other cortex models, that peristimulus inhibition is due solely to horizontal intracortical inhibitory connections, has been questioned and an additional mechanism, the competitive distribution of activity, has been proposed. We implemented a computational model of somatosensory cortex based on competitive distribution of activity. This model exhibits spontaneous map reorganization in response to a cortical lesion, going through a two-phase reorganization process. These results make a testable prediction that can be used to experimentally support or refute part of the competitive distribution hypothesis, and may lead to practically useful computational models of recovery following stroke.

1 Introduction

Feature maps in primary sensory cortex are highly plastic in adult animals: they undergo reorganization in response to deafferentation (Merzenich *et al.* 1983; Kaas 1991), de-efferentation (Sanes *et al.* 1988), localized repetitive stimuli (Jenkins *et al.* 1990), and focal cortical lesions (Jenkins and Merzenich 1987). During the past few years there have been several efforts to develop computational models of such cortical map self-organization and map refinement (Obermeyer *et al.* 1990; Pearson *et al.* 1987; Grajski and Merzenich 1990; Sklar 1990; von der Malsburg 1973; Ritter *et al.* 1989). For example, models of the hand region of primary somatosensory cortex (SI) have demonstrated map refinement, map reorganization after localized repetitive stimulation and deafferenta-

tion (Pearson *et al.* 1987), and the inverse magnification rule (Grajski and Merzenich 1990).

These computational studies show that fairly simple assumptions about network architecture and synaptic modifiability can qualitatively account for several fundamental facts about cortical map self-organization. However, it is known from limited animal experiments that focal cortical lesions also produce spontaneous map reorganization (Jenkins and Merzenich 1987). Developing a model of map reorganization following cortical lesions is important not only for the insights it may provide into basic cortical physiology, but also because it could serve as a model of nervous system plasticity following stroke (Reggia *et al.* 1993). In the only previous computational model we know of that simulated a focal cortical lesion, map reorganization would not occur unless it was preceded by complete rerandomization of weights (Grajski and Merzenich 1990). Map reorganization following a cortical lesion is fundamentally different from that involving deafferentation or focal repetitive stimulation. In the latter situations there is a change in the probability distribution of input patterns seen by the cortex, and such a change has long been recognized to result in map alterations (Kohonen 1989). In contrast, a focal cortical lesion does not affect the probability distribution of input patterns, so other factors must be responsible for map reorganization.

Past computational models of cortical map self-organization and plasticity have all assumed that the sole mechanism of intracortical inhibition is horizontal (lateral) inhibitory connections. Recently, the validity of this assumption has been called into question and an additional mechanism, the *competitive distribution of activity*, has been proposed for some intracortical inhibitory phenomena (Reggia *et al.* 1992). We recently implemented a model of cerebral cortex and thalamocortical interactions based on the hypothesis that competitive distribution is an important factor in controlling the spread of activation at the level of the thalamus and cortex (Reggia *et al.* 1992).

Because of the flexible nature of the competitive cortex model, we hypothesized that a version of this model augumented by making thalamocortical synapses plastic would not only demonstrate map formation and reorganization as with previous models, but would also demonstrate spontaneous map reorganization following cortical lesions. In the following we describe simulations that show that this is correct. Our computational model of somatosensory (SI) cortex uses competitive distribution of activity as a means of producing cortical inhibitory effects (Sutton 1992). This model, augmented with an unsupervised learning rule, successfully produces cortical map refinement, expansion of cortical representation in response to focal repetitive stimulation, and map reorganization in response to focal deafferentation. More importantly, our model exhibits something that previous models have not yet produced: map reorganization following focal cortical damage without the need to rerandomize weights. This reorganization is a two-phase process, and

results in a testable prediction that can be examined experimentally. We conclude that competitive distribution of activity can explain some features of cortical map plasticity better than the traditional view of cortical inhibition.

2 Methods

We augmented the original competitive distribution cortical model with an unsupervised learning rule that modifies synaptic strengths over time. The intent was to examine how attributing peristimulus inhibition in cortex to competitive distribution of activity rather than to lateral inhibitory connections affected map formation and reorganization. We refer to our augmented model as the *competitive SI model* because it is a crude representation of a portion of the thalamus and primary somatosensory cortex (area 3b of SI), specifically portions of those structures receiving sensory input from the hand. This area was chosen because of its topographic organization, the availability of interesting experimental data (Jenkins and Merzenich 1987; Merzenich *et al.* 1983), and to allow comparison with some previous models of SI which make more traditional assumptions about intracortical inhibition (Pearson *et al.* 1987; Grajski and Merzenich 1990).

The competitive SI model is constructed from two separate hexagonally tesselated layers of 32×32 volume elements representing the thalamus and the cortex. Each element represents a small set of spatially adjacent and functionally related neurons. To avoid edge effects, opposite edges of the cortical sheet are connected to form a torus. All connections are excitatory and competitive. Each thalamic element connects to its corresponding cortical element and the 60 surrounding cortical elements within a radius of four. With probability 0.5 a thalamocortical connection is initially assigned the minimum weight value 0.00001; otherwise, the weight is chosen uniformly randomly between this minimum and 1.0. Each cortical element connects to its six cortical neighbors; all corticocortical weights are equal (their magnitude then has no effect due to the activation rule).

Each element's functionality is governed by a competitive activation rule (Reggia *et al.* 1992). The activation $a_j(t)$ of cortical element j, representing the mean firing rate of the neurons contained in element j, is governed by

$$\frac{d}{dt}a_j(t) = c_s a_j(t) + [M - a_j(t)]in_j(t) \tag{2.1}$$

where $in_j(t) = \sum_i out_{ij}(t)$ with i ranging over all thalamic and cortical elements sending connections to cortical element j. Activation of thalamic element j is also determined by equation 2.1, but its $in_j(t)$ term represents only input from sensory receptors. Equation 2.1 bounds $a_j(t)$ between

zero and a constant M. An output dispersal rule provides for competitive distribution of activation:

$$out_{ji}(t) = c_p \left\{ \frac{w_{ji}(t)[a_j(t) + q]}{\sum_k w_{ki}(t)[a_k(t) + q]} \right\} a_i(t) \tag{2.2}$$

for both thalamic and cortical elements. The small predefined, nonnegative constant q serves two purposes: it dampens competitiveness and it prevents division by zero.

The competitive learning rule for changing weight w_{ji} on the connection to cortical element j from thalamic element i is $\Delta w_{ji}(t) = \epsilon[a_i(t) - w_{ji}(t)]a_j(t)$. To maintain normalized incoming weight vectors, an explicit weight renormalization step is needed after the weight update takes place. Simulations were run both with and without this explicit weight renormalization step, and only minor differences were observed. The results reported here are for the normalized model to be consistent with Grajski and Merzenich (1990). Random, uniformly distributed hexagonal patches were used as input stimuli because of their simplicity, intuitive appeal, and similarity to stimuli used with some past models of SI cortex.

To evaluate topographic map formation, we defined the measures

$$\text{Total response:} \quad r_j = \sum_i a_{ji} \tag{2.3}$$

$$\text{Center:} \quad \bar{x}_j = \left(\sum_i x_i a_{ji} \right) / r_j. \quad \bar{y}_j = \left(\sum_i y_i a_{ji} \right) / r_j. \quad \text{and} \tag{2.4}$$

$$\text{Moments:} \quad wx_j = \sqrt{\left[\sum_i (x_i - \bar{x}_j)^2 a_{ji} \right] / r_j}.$$

$$wy_j = \sqrt{\left[\sum_i (y_i - \bar{y}_j)^2 a_{ji} \right] / r_j}. \tag{2.5}$$

Here a_{ji} is the activation level of cortical element j when a point stimulus (input of 1.0 to a single thalamic element) is applied at thalamic element i, r_j is the total response of cortical element j summed over the thalamic point stimuli, x_i and y_i are the x and y coordinates of thalamic element i, $\bar{x}_j$ and $\bar{y}_j$ are the x and y coordinates for the center of cortical element j's receptive field, and wx_j and wy_j are the x and y moments of cortical element j's receptive field. The x and y moments of the cortical receptive field do not indicate the entire extent of the receptive field, but rather are measures of its width analogous to the standard deviation.

3 Results

Figure 1 shows the coarse topographic map that existed before training due to initially random weights and the topographic projection of thalamocortical connections. The topographic map is plotted by placing

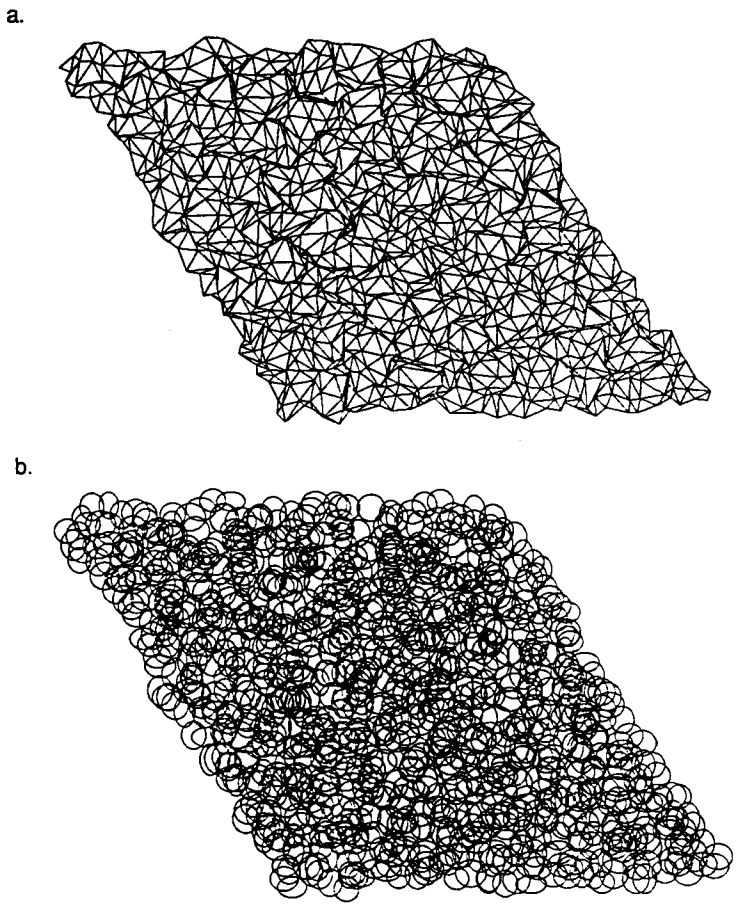

Figure 1: Cortical receptive field plots for the model before learning: (a) receptive field centers with nearest neighbors connected, and (b) receptive field centers and moments.

points at the computed centers of the cortical elements' receptive fields and connecting any two points which represent cortical elements that are nearest neighbors (Fig. 1a). The x and y cortical receptive field moments are plotted by drawing ellipses in sensory/thalamic space centered on the computed centers (Fig. 1b). The lengths of the x and y axes of each ellipse represent the x and y moments (not the full receptive fields). Anal-

ogous plots can be defined for thalamic response fields (each thalamic point stimulus generates a response across the cortical layer called its response field), cortical incoming weight vectors, and thalamic outgoing weight vectors. Figure 2 shows that with training a finely tuned, uniform topographic map appeared, and receptive field moments became uniform and decreased in size. Incoming weight vectors of cortical elements also became very uniform compared to their initial random state; after training with hexagonal patches of radius two, the incoming weight vectors became roughly bell-shaped when the weights are plotted as a surface.

When the fingertips of attending monkeys are stimulated much more frequently than the rest of the hand, the region of SI cortex which represents the fingertips increases in size (Jenkins *et al.* 1990). This increase in size is mostly at the expense of neighboring regions, but also at the expense of more distant regions. The receptive fields in the expanded fingertips region of cortex also show a decrease in size. For our model, these results were simulated by first performing topographic map formation as above. After the map formed (Fig. 2), the input scheme was changed so that an 8×16 region of the thalamic layer designated as the repetitively stimulated finger (second finger region from the left), was now seven times more likely to be stimulated than other regions.

After the topographic map reorganized due to this change in the input scheme, a number of effects observed in animal studies occur in the competitive SI model (Fig. 3). The number of cortical elements whose receptive field centers were in the repetitively stimulated finger representation increased dramatically, more than doubling. Thus, there was a substantial increase in the cortical magnification for the repetitively stimulated finger, as is observed experimentally. The neighboring finger representations decreased in size and shifted, and even more distant finger representations were reduced in size. Following repetitive stimulation, for the enlarged representation of the second finger shown in Figure 3b the mean receptive field size did not decrease. However, receptive field size did decrease for a large number of the cortical elements whose receptive field centers lie toward the edges of the repetitively stimulated finger representation, consistent with the inverse magnification rule in these regions.

To simulate an afferent lesion, our model was trained and then a contiguous portion of the thalamic layer corresponding to a single finger was deprived of sensory input. Input patterns within the deafferented finger region ceased to occur. With continued training, some cortical receptive fields which were in the deafferented finger region shifted outside of this region, forcing a resultant shift of surrounding cortical receptive fields. Some cortical elements near the center of the deafferented region received insufficient activation to reorganize and remained essentially unresponsive to all input patterns. Much reorganization in biological cortex involves replacement of deafferented glabrous representation by

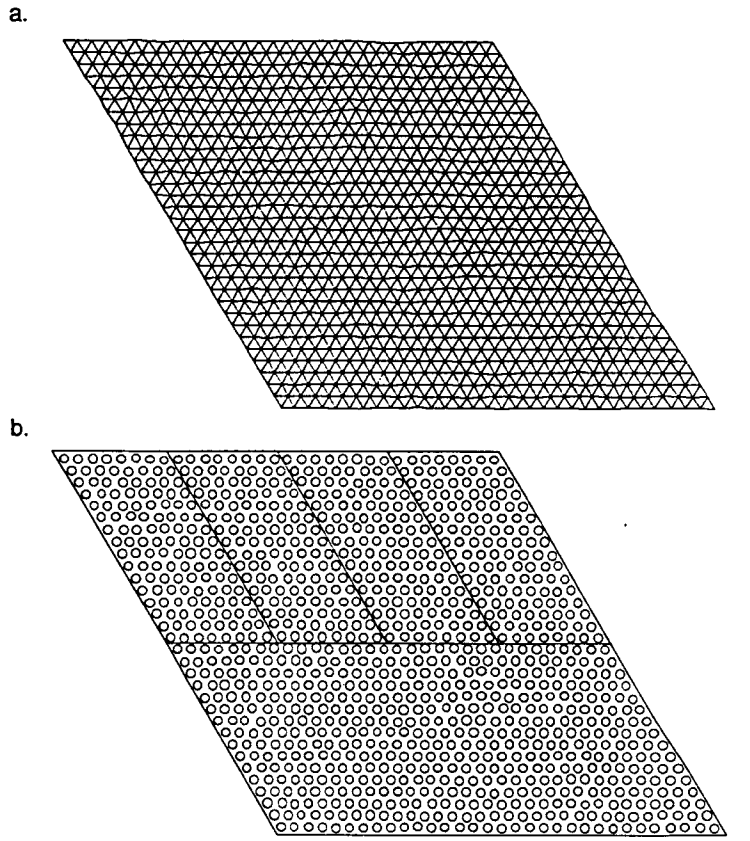

Figure 2: Cortical receptive field plots for the model after training with randomly positioned hexagonal patches of radius two as input stimuli: (a) receptive field centers with nearest neighbors connected and (b) receptive field centers and moments. Contrast with Figure 1. The refined topographic map (b) has small, roughly equal-size receptive fields [actual receptive fields extend beyond the ovals drawn in (b) and overlap]. As a reference for discussion, boundaries for four "fingers" and a "palm" are added. In all the simulations described in this paper $c_s = -2.0$ (self-inhibition), $c_p = 0.6$ (excitatory gain), $M = 3.0$, $q = 0.0001$ in equations 2.1 and 2.2, and a time step of 0.5 is used, for cortical elements. The same values are used for thalamic elements except $c_p = 1.0$. A learning rate of $\epsilon = 0.01$ is used. Distance between neighboring receptive fields here and in subsequent figures is 1.0.

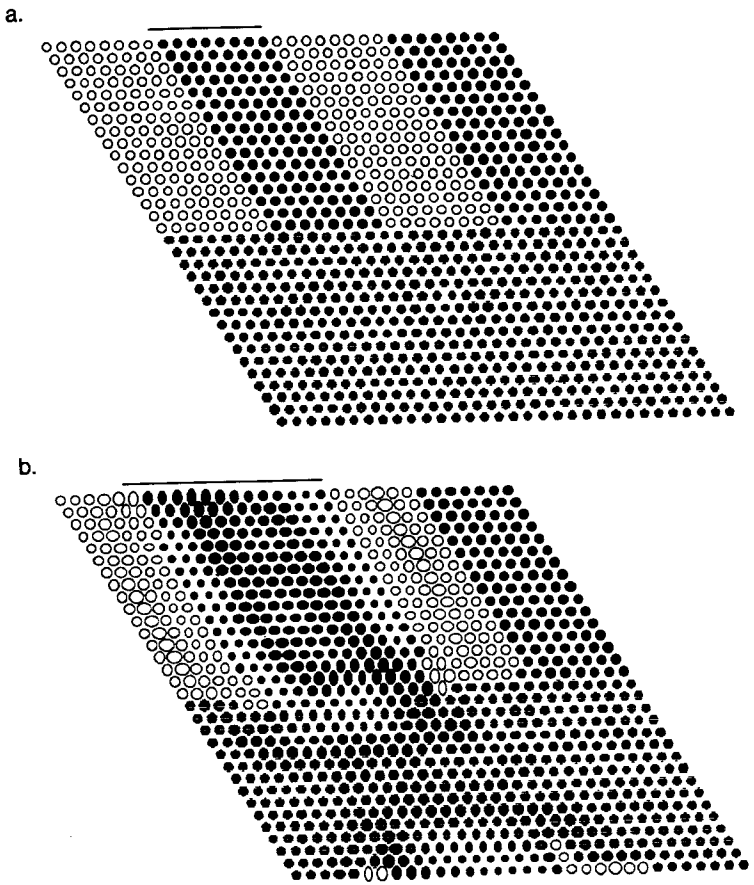

Figure 3: Cortical receptive field moments plotted in cortical space and filled to show regional location of receptive field centers (a) after training with uniformly distributed stimuli, and (b) after subsequent repeated finger stimulation. The horizontal line indicates the breadth of the cortical representation of the selectively stimulated finger before (a) and after (b) repetitive stimulation. The x and y moments of the cortical receptive fields are represented by ellipses centered at the physical location of the cortical element in the cortex rather than the location of the receptive field center in sensory/thalamic space. The general location of receptive field centers is shown by filling ellipses with different gray scale patterns to indicate the different finger and palm regions.

new/extended representation of dorsal hand surfaces (Kaas 1990); this could not happen in our model as no sensory input corresponding to the dorsum of the hand was present.

To simulate a focal cortical lesion in our model, a contiguous portion of the trained cortical layer (elements representing the second finger from the left) was deactivated after training (i.e., activation levels clamped at 0.0). After lesioning, the topographic map showed a two-phase reorganization process. Immediately after lesioning and before any retraining, the receptive fields of cortical elements adjoining the lesioned area shifted towards the second thalamic finger and increased in size (Fig. 4a). This immediate shift was due to the competitive redistribution of thalamic output from lesioned to unlesioned cortical elements. The second phase of map reorganization occurred more slowly with continued training and was due to synaptic weight changes (Fig. 4b).[1] Cortical representation of the "lesioned finger" was reduced in size (reduced magnification). The mean receptive field x moment (y moment) prior to the lesion was 0.626 (0.627) for the entire cortex. Following the cortical lesion and subsequent map reorganization, the mean receptive field x moment (y moment) increased to 0.811 (0.854) for elements within a distance of two of the lesion site (mostly shaded black in Fig. 4b), consistent with an inverse magnification rule.

4 Discussion

It has recently been proposed that competitive distribution of activity may underlie some inhibitory effects observed in neocortex (Reggia *et al.* 1992). The present study supports this hypothesis in two ways. First, we have shown that a computational model of cerebral cortex based on the competitive distribution hypothesis, starting from a coarse topographic map, can simulate the development of a highly-refined topographic map with focused, bell-shaped receptive fields. Once such a map was formed, changing the probability distribution of input stimuli resulted in substantial map reorganization. With repetitive stimuli to a localized region of the sensory surface, the cortical representation of that region increased dramatically and changes occurred that were in part consistent with an inverse magnification rule as has been observed experimentally (Jenkins *et al.* 1990). With deafferentation of a localized region many of the cortical elements originally representing the deafferented region came to represent surrounding sensory surface regions, as has been described experimentally (Kaas 1991). Of course, neither our model nor previous ones, being tremendous simplifications of real cortex (e.g., having very limited radii of interconnectivity and modeling only a small cortical

[1] Results reported here are for unchanged uniformly random stimuli. If stimuli frequency in the region originally represented in the lesioned cortex was increased, reorganization was even more pronounced.

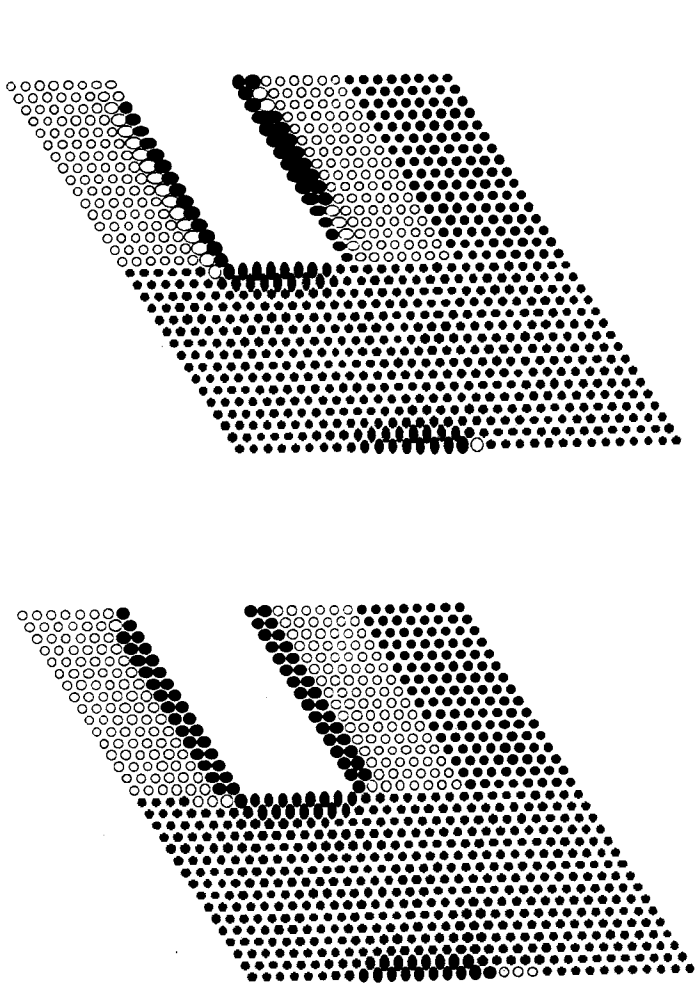

Figure 4: Cortical receptive field moments and centers plotted in cortical space (a) immediately following a cortical lesion, and (b) following a cortical lesion and further training with uniformly distributed stimuli. Finger regions originally represented in the now-lesioned cortical area (indicated by solid black ovals) have shifted substantially into the surrounding intact cortex.

region), can account for all experimental data related to these phenomena. For example, none of these models, including ours, adequately accounts for the almost immediate changes in receptive fields observed following deafferentation (Calford and Tweedle 1991; Kaas *et al.* 1990; Chino *et al.* 1992; Gilbert and Wiesel 1992) nor for long-term "massive" reorganization (Pons *et al.* 1991).

Second, our competitive SI model exhibited dramatic map reorganization in response to a focal cortical lesion. Reorganization following a cortical lesion is fundamentally different from repetitive stimulation and deafferentation as there is no change in the probability distribution of input stimuli. To our knowledge, the only previous cortical model which has tried to simulate reorganization following a focal cortical lesion is the three-layer model of Grajski and Merzenich (1990). Following a focal cortical lesion, map reorganization did not occur with this earlier model unless the synaptic strengths of all intracortical and cortical afferent connections that remained intact were randomized, that is, unless the model reverted to its completely untrained state (it was also necessary to enhance cortical excitation or reduce cortical inhibition). In our SI model no special procedure such as weight randomization was required: sensory regions originally represented in the lesioned cortex spontaneously reappeared in cortex outside the lesion area. This spontaneous map reorganization is consistent with that seen experimentally following small cortical lesions (Jenkins and Merzenich 1987). Further, the model's receptive fields increased in size in perilesion cortex as has also been described experimentally (Jenkins and Merzenich 1987). Demonstration that the competitive SI model reorganizes after a cortical lesion provides a potential computational model of stroke.[2]

Map reorganization following a cortical lesion to the competitive SI model involved a two-phase process where each phase, rapid and slow, is due to a different mechanism. Immediately after a cortical lesion, competitive distribution of activation caused some finger regions orginally represented by the lesioned area of cortex to "shift outward" and appear in adjacent regions of intact cortex. This result provides a specific testable prediction for half of the competitive distribution hypothesis: if competitive distribution of activity is present from thalamus to cerebral cortex, then significant shifts of sensory representation out of a lesioned cortical area should be observed right after a cortical lesion. The second, slower phase of additional map reorganization is due to synaptic plasticity, and is apparently triggered by the first phase. It is not yet clear whether a model based on more traditional intracortical inhibitory connections can produce spontaneous reorganization following a cortical lesion. Further computational studies should determine whether the difficulties encountered in obtaining such reorganization are a general property of cortical models using inhibitory connections or whether they

[2]We have recently extended this work to a model of proprioceptive cortex based on length and tension input from muscles in a model arm (Cho *et al.* 1993).

reflect specific details of the one computational model studied so far (Grajski and Merzenich 1990).

Acknowledgments

Supported by NINDS Awards NS 29414 and NS 16332. The authors are also with the Department of Neurology and the Institute for Advanced Computer Studies at the University of Maryland.

References

Calford, M., and Tweedle, R. 1991. Acute changes in cutaneous receptive fields in primary somatosensory cortex after digit denervation in adult flying fox. *J. Neurophys.* **65**, 178–187.

Chino, Y., Kaas, J., and Smith, E., *et al.* 1992. Rapid reorganization of cortical maps in adult cats following restricted deafferentation in retina. *Vision Res.* **32**, 789–796.

Cho, S., Reggia, J., and D'Autrechy, C. 1993. *Modelling map formation in proprioceptive cortex.* Tech. Rep. CS-TR-3026, Dept. of Computer Science, Univ. of Maryland.

Gilbert, C., and Wiesel, T. 1992. Receptive field dynamics in adult primary visual cortex. *Nature (London)* **356**, 150–152.

Grajski, K., and Merzenich, M. 1990. Hebb-type dynamics is sufficient to account for the inverse magnification rule in cortical somatotopy. *Neural Comp.* **2**, 71–84.

Jenkins, W., and Merzenich, M. 1987. Reorganization of neocortical representations after brain injury: A neurophysiological model of the bases of recovery from stroke. In *Progress in Brain Research*, Vol. 71, F. Seil, E. Herbert, and B. Carlson, eds., pp. 249–266. Elsevier, Amsterdam.

Jenkins, W., Merzenich, M., Ochs, M., Allard, T., and Guic-Robles, E. 1990. Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J. Neurophys.* **63**, 82–104.

Kaas, J. 1991. Plasticity of sensory and motor maps in adult mammals. *Ann. Rev. Neurosci.* **14**, 137.

Kaas, J., Krubitzer, L., Chino, Y., *et al.* 1990. Reorganization of retinotopic cortical maps in adult mammals after lesions of the retina. *Science* **248**, 229–231.

Kohonen, T. 1989. *Self-Organization and Associative Memory.* Springer-Verlag, Berlin.

Merzenich, M., Kaas, J., *et al.* 1983. Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience* **8**, 33–55.

Obermayer, K., Ritter, H., and Schulten, K. 1990. A neural network model for the formation of topographic maps in the CNS: Development of receptive fields. *Proceedings of International Joint Conference on Neural Networks*, Vol. II, pp. 423–429. San Diego, CA.

Pearson, J., Finkel, L., and Edelman, G. 1987. Plasticity in the organization of adult cerebral cortical maps: A computer simulation. *J. Neurosci.* **7**, 4209–4223.

Pons, T., Garraghty, P., Ommaya, A., *et al.* 1991. Massive cortical reorganization after sensory deafferentation in adult macaques. *Science* **252**, 1857–1860.

Reggia, J., D'Autrechy, C., Sutton, G., and Weinrich, M. 1992. A competitive distribution theory of neocortical dynamics. *Neural Comp.* **4**, 287–317.

Reggia, J., Berndt, R., and D'Autrechy, C. 1993. Connectionist models in neuropsychology, *Handbook of Neuropsychology*, Vol. 9, Elsevier, in press.

Ritter, H., Martinetz, T., and Schulten, K. 1989. Topology-conserving maps for learning visuo-motor-coordination. *Neural Networks* **2**, 159–168.

Sanes, J., Suner, S., Lando, J., and Donoghue, J. 1988. Rapid reorganization of adult rat motor cortex somatic representation patterns after motor nerve injury. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2003.

Sklar, E. 1990. A simulation of cortical map plasticity. *Proc. IJCNN* **III**, 727–732.

Sutton, G. 1992. Map formation in neural networks using competitive activation mechanisms. Ph.D. Dissertation, Department of Computer Science, CS-TR-2932, Univ. of Maryland.

von der Malsburg, C. 1973. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetic* **14**, 85–100.

Received January 4, 1993; accepted June 1, 1993.