Formation of Dimension-Reducing Somatotopic Maps

E. Erwin, K. Obermayer and K. Schulten
Beckman Institute and Department of Physics
University of Illinois at Urbana-Champaign, Urbana IL 61801
email: erwin@lisboa.ks.uiuc.edu

The somatosensory cortex of humans and higher animals contains a map of the body surface with a largely topographically ordered connectivity between tactile skin receptors and cortical neurons. Merzenich, et. al., have extensively studied the hand region of the somatotopic map in areas 3b and 1 of the cortex in a number of adult owl and squirrel monkeys [10]. They found that the maps are highly individual to a particular animal and change even in adult animals in response to the frequency of sensory stimuli received, although some general map characteristics do seem to be constant between individuals. In an earlier model [14] based on the self-organizing feature map algorithm [5, 6], we showed that some aspects of the formation of the somatotopic cortical maps may be explained by an input-driven, self-organizing process. However, our earlier model treated the hand surface as a two-dimensional object and did not take into account the interaction between the representations of the dorsal and glabrous hand surfaces. Here we report on an extended model which considers the hand as a true three-dimensional object whose surface must be mapped onto the two-dimensional cortical array. Our model cortical maps reproduce many features of the experimental maps well. However, in order for the model to produce maps with the relative locations of map features showing the same tendency to be correlated as in the experimentally observed maps, it is necessary that the initial pattern of connectivity contain a bias toward the representation of the dorsal surface in particular regions of the largely topographic map of the glabrous surface. This finding then suggests how specific the biological pattern of connectivity between skin receptors and the cortex might need to be before an activity-based refinement of the map structure begins. Due to the high computational requirements, our model was implemented on the Connection Machine CM-2.

1   Introduction

Feature maps are one of the major architectural features of the cortex. They appear in their simplest form in the primary sensory and motor areas, but some higher cortical areas also make use of similar structures. In these cortical regions, the tuning properties of a subset of the neurons are spatially segregated, forming a map which often preserves much of the topological relationships of the feature space being represented. Recently, high-resolution images of the somatosensory [10] and visual [2] feature maps in primate cortex have become available, making it possible to study the structure of these cortical maps in

detail. Predictions from models of the development of these maps may now be compared with observed map structures.

Cortical areas 1 and 3b in primates are the primary cortical destination of efferents from tactile sensory receptor cells on the skin surface. Inputs to these areas are spatially segregated such that distorted, two-dimensional maps of the skin surface are formed, where neurons in different regions of the maps process input from different locations on the skin surface. Merzenich, et. al., have mapped the representation of somatosensory input from the hand in these cortical areas in a number of owl and squirrel monkeys [10]. Their data clearly shows that there does not exist one single cortical "hand map" which is identical for all individuals of a species, but rather that the maps are unique to each single monkey, and only some rather general characteristics are shared. Even in adult animals the hand maps are not constant in their arrangement. The amount of area of the map which responds to a particular location on the skin surface may increase or decrease depending on the relative amount of stimulation received by that area [10]. The maps may even show large-scale rearrangement to adapt to extreme changes in the patterns of nerve stimulation caused by, e. g., amputation of a digit [11] or other pathological changes [9]. Yet there is a limit to the amount of rearrangement which can take place, with the *plasticity* of the maps far greater during an early, developmental stage of the animal's life.

These and other observations have led to the generally accepted view that brain maps are not genetically predetermined, or "prewired", but rather are only roughly specified genetically and are refined by an activity-based process during a developmental period. (See, e.g., [1].) The initial connections which grow in the developing brain could perhaps be guided by cell-adhesion molecules [4, 19] or by other means to form a rough topographic order for the somatosensory maps. A second phase of map formation follows which refines the initially crude map by selective strengthening, weakening or elimination of connections between the receptors on the skin (or interneurons) and the neurons in the cortical map. This later phase of map refinement, which in primates may last long after birth, appears to be driven by temporal correlations in the patterns of sensory input. The hand maps seen in adult animals are thus the result of both genetic factors and the "pattern of use" of the hand by the individual [10].

Our previous efforts in modeling the formation of brain maps in the visual [13, 15, 16], auditory [8], and somatosensory [14] cortical areas have helped us to understand how a simple algorithm for modifying synaptic efficacies, based on a Hebbian learning rule, can give rise to the types of feature maps observed experimentally. Our model visual maps have compared well qualitatively and even quantitatively with experimentally observed maps [13]. To achieve such detailed agreement, it was necessary to assume that the initial connections between the cortex and retina, which form prior to the start of the self-organizing process, be roughly topographic. When it becomes possible to compare model predictions and experimental observations in more detail, additional mechanisms may also be needed to explain the formation of the detailed structure of the maps. Genetically prespecified patterns of cell growth and subsequent refinement of connections by an activity-driven process may only be two elements in a modular and partially redundant set of processes [12].

Since a simple model could well describe the formation of the visual maps, we wondered whether the same simple model would be equally capable of explaining the development of

the detailed structure of the somatosensory maps, and if so how specific the connectivity patterns set up before the self-organizing process of map refinement begins would need to be. To address this question, it was necessary to extend our earlier model [14] of formation of cortical hand maps to include a three-dimensional, rather than two-dimensional representation of the hand surface.

The problems which arise in mapping the three-dimensional body surface onto the essentially two-dimensional cortical array are quite different from those arising in the mapping of the high-dimensional space describing features in visual input onto the cortex. In the visual cortex, a map of a high-dimensional, *continuous* feature space, comprising stimulus location, orientation, ocularity, etc., must be formed on a two-dimensional area. Such mappings from continuous spaces onto a regular lattice of neurons have been successfully treated by analytical methods [15, 16, 17, 18] as well as by simulations [13, 16]. In the somatosensory cortex, a map of an irregularly-shaped manifold, the skin surface, embedded in a higher-dimensional space must be formed on the two-dimensional cortical sheet. This more difficult type of dimension-reducing mapping may currently be treated only by simulations.

2 Experimentally Observed Maps

Cortical region 3b, which contains one of the primary cortical representations of the skin surface, has been extensively studied by Merzenich, et. al. They have published detailed maps of the region of this brain area processing tactile input from the hands for nine individual adult owl monkeys and five squirrel monkeys, and have also studied the similar hand maps in cortical area 1 [10]. A few of their maps from area 3b of squirrel monkeys are reproduced in Figure 1. Although clearly there does not exist one single "hand map" which is shared by all members of the species, some features of the maps are constant between individuals. For instance, the position of the hand map in the cortex is constant relative to the maps of other skin areas. All parts of the glabrous (front) surface of the hand are usually represented in the map and the representation is roughly topographic, with the thumb represented in the most medial region and the fingers represented, in order, along the lateral direction. The tips of the fingers appear in the rostral area of the maps and the palm is represented on the opposite (caudal) side.

Studying the structure of the maps in detail, Merzenich, et. al., were able to make several additional general observations: (i) At the finest scale, strict internal topography is maintained; (ii) The overall size of the representations of individual finger surfaces often varies severalfold as does the overall size of the complete hand map; (iii) The dorsal (hairy) surface of the fingers is only partially represented in disconnected, but internally continuous "islands", and the hairy surface of the hand itself is never represented in area 3b; (iv) The largest areas of representation of the hairy finger surfaces tend to occur at the lateral and medial margins of the map. Although the representation of the hairy surface in area 3b is never complete, it is more fully represented in area 3b of owl monkeys than in squirrel monkeys.

Although not emphasized by Merzenich, the "islands" where the hairy surface of the fingers are represented seem to occur more often at the rostral edge of the map, where the
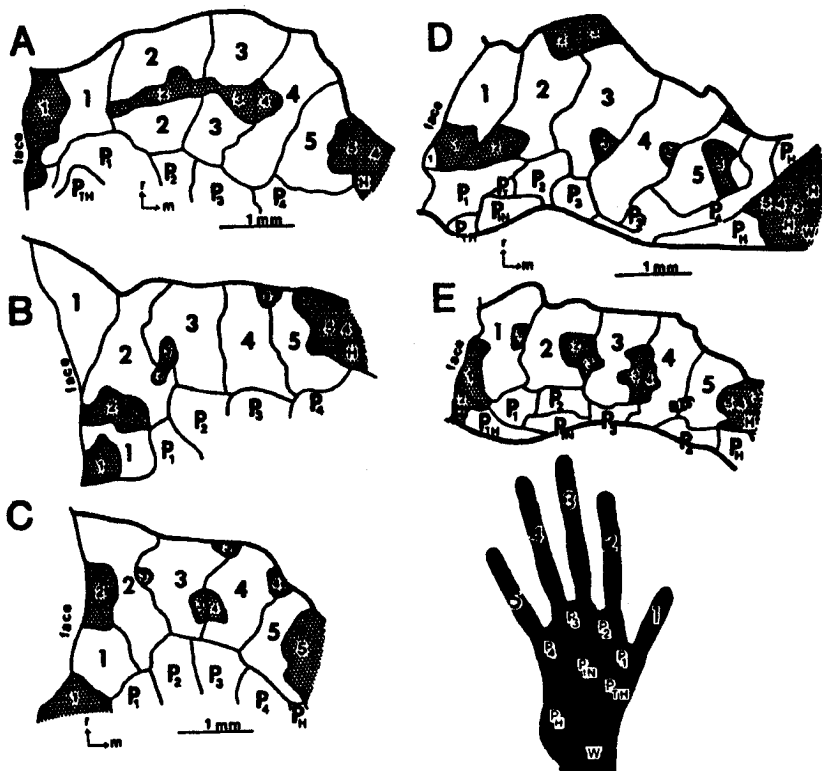
3

Figure 1: "Hand maps" from cortical area 3b of five individual adult squirrel monkeys. (Adapted from [10].) Each region of the cortical maps is labelled to show the location on the hand of the centers of the receptive fields of neurons in that area. Hatched areas represent cortical cells with receptive fields on the dorsal (hairy) surface of the hand.

fingertips are mapped, or at the interface between the representations of the glabrous areas of two adjacent fingers, and rarely occur in the center of the glabrous representation of a digit. The number of maps studied may, however, be too small to justify this generalization.

3   A Model for the Formation of Somatosensory Cortical Maps

3.1   The Two-Dimensional Model

Earlier [14] we studied the activity-driven formation of the somatotopic representation of the hand in the somatosensory cortex, using a model based on the self-organizing feature map algorithm of Kohonen [5, 6, 7]. In the model, cortical *cells*, which represent a small column of neurons, are located at points (i, j) on a two-dimensional lattice, and a number N of sensory receptor cells are distributed at random locations on a two-dimensional area in the shape of a hand with five fingers, Figure 2. Locations on the hand surface are color coded for easy identification.
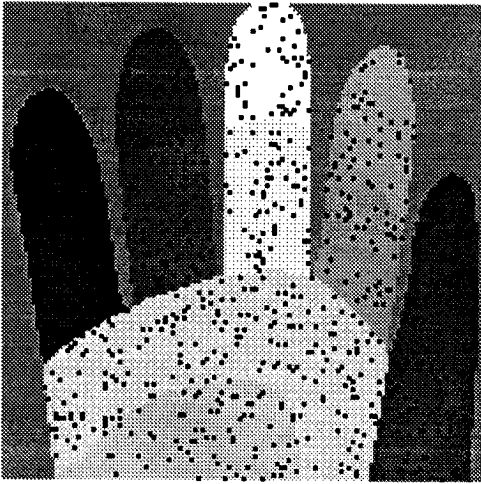
Figure 2: The two-dimensional model hand surface. Small regions of the hand are color-coded for easy identification. The small, black squares represent the locations of the 800 sensory receptors.
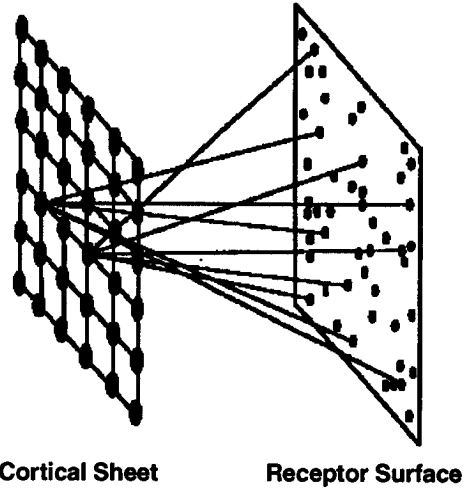

Cortical Sheet   **Receptor Surface**

Figure 3: Each cortical cell in a 128 by 128 two-dimensional lattice receives input from all of the receptor cells. The variable strengths of these connections are given by the weight vectors $\vec{w}_{ij}$.

All cortical cells are assumed to be able to receive input from each of the sensory cells. (Figure 3.) The strength of the connection between a cell at the lattice location (i, j) and the m-th receptor at location $\vec{r}_m$ on the hand surface is denoted by $w_{ij,m}$. The set of all connection strengths to an individual cortical cell are collectively called a weight vector,

$$\vec{w}_{ij} = (w_1, w_2, \ldots, w_m, \ldots, w_N). \tag{1}$$

The weight vector $\vec{w}_{ij}$, together with the receptor locations $\vec{r}_m$, describes the *receptive field* of cortical cell (i, j).

After setting the weight vectors to some initial values, formation of the map proceeds by an iterative process. A localized stimulus, representing a touch, is presented at some random location on the hand surface. The intensity distribution of the stimulus on the receptors may be expressed in the same vector form as the weights:

$$\vec{v} = (v_1, v_2, \ldots, v_m, \ldots, v_N). \tag{2}$$

We chose to use stimuli with Gaussian intensity distributions of a constant width, with the probability of a stimulus being centered at any location on the hand surface made constant. The output o_{ij} of each cortical cell is computed as a sum of the stimulus intensity at each receptor weighted by the connection strengths between the neurons and receptors:

$$o_{ij} = \sum_m w_{ij,m} v_m = \vec{w}_{ij} \cdot \vec{v} \tag{3}$$

The cell with the highest output at each time-step is chosen as the "winner" neuron, and its coordinates are labelled (i_0, j_0). The outputs of all cells are replaced by the values h_{ij} of

a Gaussian output function, the *neighborhood function*, centered at the selected cell (i_0, j_0),

$$h_{ij}(t) = exp[-((i - i_0)^2 + (j - j_0)^2)/\sigma^2(t)]. \tag{4}$$

This neighborhood function represents the "activity bubbles" seen experimentally in the cortex in response to stimuli [3, 7]. Weight values of every cell in the network are next changed according to the *update rule*,

$$w_{ij,m}(t + 1) = w_{ij,m}(t) + \epsilon(t)h_{ij}(t) \cdot r_m / \sqrt{\sum_m (w_{ij,m}(t) + \epsilon(t)h_{ij}(t) \cdot r_m)^2}, \tag{5}$$

which represents a Hebb-like modification of synapses and, due to the definition of h_{ij} (4), makes cells which are nearby in the cortical array tend to respond most strongly to similar inputs. The term in the denominator simply normalizes the weights after each update step so that they remain bounded.

The *learning step width* $\epsilon(t)$ and the width $\sigma(t)$ of the output function h_{ij} may be kept constant or may decrease during a simulation. A decreasing learning step width $\epsilon(t)$ is used to model the decrease in plasticity of the cortical maps during the early, developmental period of an animals life. Likewise, a decreasing width of the output function $\sigma(t)$ allows neurons to gradually specialize to respond to different sub-regions of the input space. However, with the two-dimensional model we were also able to demonstrate that the self-organizing feature-map algorithm, with $\sigma(t)$ and $\epsilon(t)$ set to small constant values, can mimic the plasticity of cortical maps in adult animals. With $\sigma(t)$ and $\epsilon(t)$ set to small values, the map structure is fairly constant, but can change slowly in response to changes in the distribution of sensory inputs, as in our simulation [14] of an amputation experiment performed by Merzenich, et. al. [11].

In our simulations of map formation with the two-dimensional model, we found that an ordered cortical map would form even when the initial weight values were selected at random, as long as the value of $\sigma(t)$ were sufficiently large initially. Figure 4 shows the development of one model map starting with initially random connections. The cortical map is represented by coloring the cell locations in the two-dimensional cortical array with the color of the location of the centroid of the receptive field of that cell, $\sum_m w_{ij,m} \vec{r}_m$. The genetically-specified initial connections between the cortex and skin surface which are set up before any supposed self-organizing process begins are of course not completely random. Initially each neuron is likely to be connected primarily to receptors in only a small area of skin surface, and the activity-based learning process merely refines this initial crude mapping. The initial connections must at least enforce a preference for the map to have a particular orientation, since the overall orientation of the hand-maps is observed to be constant between individuals.

3.2 The Three-Dimensional Model

The two-dimensional model gives insight into how a self-organizing process due to Hebb-like changes in synaptic efficacies could account for the development of topographic somatosensory maps, and the plastic modification of these maps in response to changes in the types of
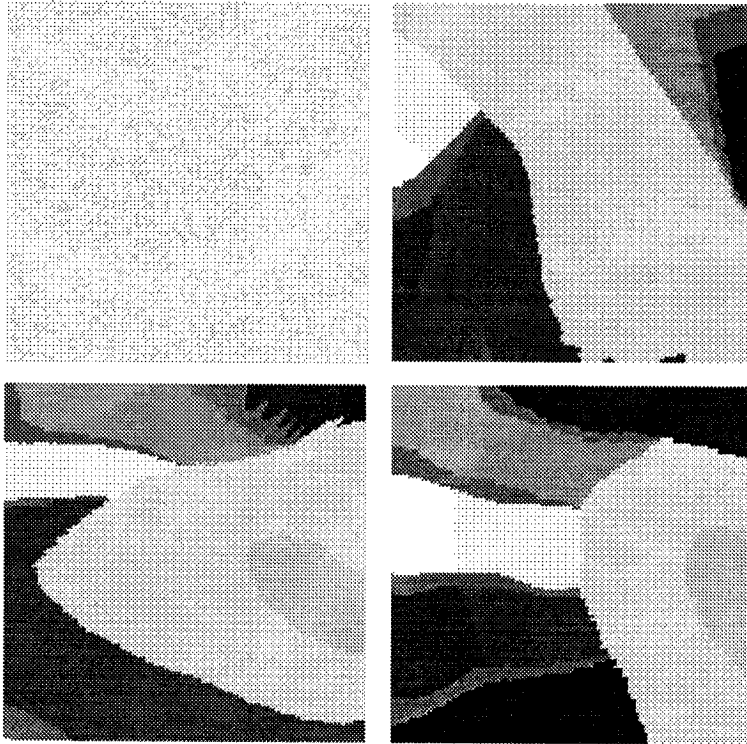
Figure 4: The development of the cortical map in the two-dimensional model, starting from initially random weight values, is show in the sequence of images from upper left to lower right. Each pixel corresponds to one of the 128 by 128 cortical cells, and is given the color of the location of the centroid of the cortical cell's receptive field on the model hand surface.

sensory input being processed. This model cannot, however, predict the structure of these cortical maps in detail. In particular it cannot predict how the inputs from the front and back surfaces of the hand will be segregated in the cortical maps, since the model includes receptors on only a single plane surface. Our success in using the self-organizing feature map algorithm to predict the detailed structure of cortical visual maps led us to consider whether such a simple algorithm could also be sufficient to explain the development of the detailed feature maps observed experimentally in the somatosensory cortex, and if so, how specific the initial pattern of connectivity would need to be. To test the abilities of the algorithm, we extended our earlier model to include a more realistic, three-dimensional skin surface.

The self-organizing map algorithm for the three-dimensional model is identical to the two-dimensional version, except that the vectors, $\vec{r}_m$, have components in three spatial directions and represent the locations of the sensory receptors on the surface of a three-dimensional model hand. In some simulations we employed a collection of blocks and cylinders to represent the shape of a complete hand, but in the simulations described here only the fingers and a small fraction of the upper hand were modeled, thereby allowing for greater receptor density.

It is likely that the experimentally observed disparity in the sizes of the areas of the cortical map representing input from the front and back surfaces of the hand result from the lower receptor density on the back (hairy) surface and the more frequent stimulation of the smooth (glabrous) surface. In our model the front surface of the hand, comprising skin surfaces forward of a plane located approximately one-third of the distance from the most forward points to the rearmost points, was distinguished from the back by a higher density of receptors. Various ratios of the two densities were tested in different simulations. To represent the naturally occurring stimulus distribution would be impossible in such a simple model, but it seems reasonable to assume that the glabrous surface of the hand and fingers is more frequently stimulated during normal hand use. We again used stimuli represented by a Gaussian distribution of intensity at each receptor, with the constant three-dimensional width of the distribution chosen large enough that stimuli falling between two fingers could excite receptors on each finger, but small enough that stimuli falling on the front or back surface of the hand could not excite receptors on the opposite surface. The centers of the stimuli could be located at any point on the hand surface, with the probability density of the stimuli greater on the frontal plane of the model hand.

We intended to investigate with our model how different initial connectivity schemes affected the types of feature maps formed by the self-organizing process, and to determine what patterns of initial connections would lead to maps similar to those observed experimentally. In the two-dimensional model, ordered topographic maps would form even when the initial connections between the cortex and the skin surface were randomly assigned. However, in the full three-dimensional model, simulations with initially random connection strengths did not in general produce maps which closely resembled the experimentally observed maps. Instead maps developed which were locally topologically ordered at most locations, but which contained topological "defects" on a larger scale.

By appropriately specifying the strength of the initial connections between each cortical cell and the receptors, essentially any desired initial map pattern can be set. When initial conditions were specified such that the rough two-dimensional topography was specified, with the third dimension (front to back) suppressed, final maps developed which more closely resembled those seen in real monkeys. An example of such a final map is shown in Figure 5a, using conventions similar to those in Figures 1 and 4. The model map resembles experimental maps in that the front surface is represented relatively fully and continuously, while representation of the back surface is relegated to small, internally continuous "islands". The model map differs from the experimental maps in that these "islands" tend to occur more often in the center of the representations of the fingers, rather than at the borders between the fingers, as in the experimental maps. This correlation was found in all of the model maps developed with widely different model parameters. Although in most of the experimental maps, the "islands" tend not to be located in the center of the representations of the fingers, some of the experimental maps do look similar to the model maps. Compare for instance Figure 5a with the map for monkey "A" in Figure 1.

Since the self-organizing model of map development tended to cause the "islands" of dorsal skin representation to develop in different locations than those seen in experimental maps, the model seems to suggest that the initial connections formed during normal brain development may contain biases favoring development of representations of the hairy surface
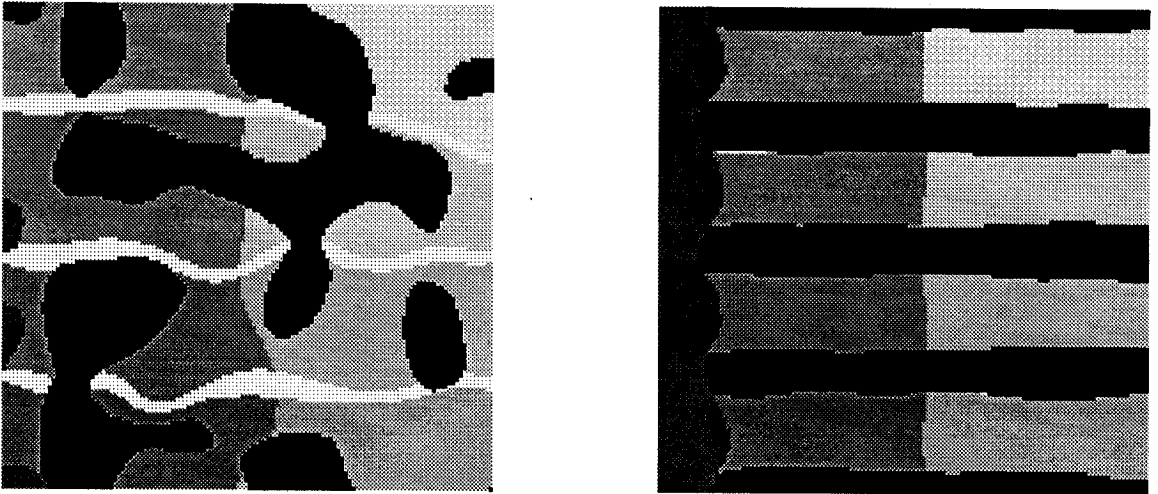
Figure 5: Final cortical maps from the three-dimensional model, starting from different initial connectivity patterns. The 1400 receptors were located on a model hand consisting of cylinders, representing four fingers, and a small rectangular sheet representing the upper part of the palm. Each pixel corresponds to one of the 128 by 128 cortical cells. Those with receptive fields centered on the glabrous hand surface are depicted in various shades. Cortical cells with receptive fields centered on the dorsal surface are colored black. The palm appears at the left edge of each map. In map **a** (left), the initial weights provided the rough two-dimensional topographic order, with the third dimension suppressed. The value of $\sigma(t)$ decreased linearly from 12 to 9 over 8000 iterations. In map **b** (right) the initial connections were specified in detail such that the hairy surface was represented in the areas between the representations of adjacent glabrous surfaces. The final map retains a smoothed form of this structure. For this simulation $\sigma(t)$ decreased linearly from 9 to 8 over 8000 iterations. In both simulations $\epsilon(t)$ was constant at 0.1, the number of receptors on the front and back surfaces was approximately equal, and 85 percent of the stimuli were centered somewhere on the frontmost one-fifth of the model hand.

in certain locations. To test whether the self-organizing algorithm could be influenced by biases in the initial connectivity pattern to generate maps more like the experimentally observed maps, we applied the algorithm to maps with highly specified initial connections. When the initial connections were set up such that the back surfaces of the fingers were represented at the edges of the maps, as in the map in monkey "E" in Figure 1, the maps retained this organization during the self-organizing process. We also set initial connections such that the map was specified to have the glabrous surface of the fingers represented in topological order, with the hairy surface represented in the space between the representation of the glabrous surface of adjacent digits. When the algorithm was started from these initial conditions, the map underwent a rapid rearrangement to smooth out discontinuities, but soon settled into a stable form shown in Figure 5b, maintaining the representation of the dorsal surface between adjacent fingers.

4 Discussion

Feature maps are such a common component of many diverse cortical areas that any attempt to understand the function of the brain must include an explanation of their structure and function. Higher brain functions depend of course on more than simply mapping some features to particular cortical locations, but studying how the brain organizes incoming information into maps, gives an indication of what features are considered most important by the brain for interpretation of the wealth of information with which it is constantly barraged. Studies of the way the visual system breaks down the components of a scene into simple features has already proven useful in constructing models for machine vision, and will certainly continue to do so in the future. Finding the common aspects of different models of the formation of these structures could give clues to the purpose served by the *particular* forms of the maps which are observed and could help us learn how to train artificial neural networks to do the same sorts of input decomposition and adaptive control processes performed by the cortical maps.

We have shown here that the self-organizing feature map model is one model which can successfully predict the formation of the types of maps observed in the somatosensory cortex. This model, which is intended to describe the input-driven refinement of an initially crude map set up by some prior process, successfully predicts the formation of hand maps which agree with the experimentally observed maps in many details, even when starting from a very rough initial connectivity pattern. The model correctly predicts that the representation of the dorsal finger surfaces will be relegated to small "islands" embedded within an overall two-dimensional topographic map. However the model maps and experimental maps disagree in the relative location of the "islands" with respect to other features in the map. The model consistently predicted that the "islands" should be located in the centers of the representations of the glabrous surface of the fingers, while in the experimental maps the "islands" are far more common in the border regions between the representations of adjacent fingers. However, in simulations where the initial connectivity pattern specifies a map where the dorsal finger skin is represented between adjacent fingers, the model predicts that the self-organizing process will lead to a map which maintains a similar organization. Using initial connectivity patterns that contain a bias toward the representation of the dorsal skin in certain areas, but where the bias is implemented in a less extreme manner, would probably lead to formation of maps which agree more fully with the experimental maps. These simulations suggest that the formation of the experimentally observed maps requires initial connections which specify at least the two-dimensional topography of the hand surface and some hint of the three-dimensional structure.

Acknowledgements

References

[1] C. Blakemore and R. C. van Sluyters. Innate and environmental factors in the development of the kitten's visual cortex. *Journal of Physiology*, 248:663–716, 1975.

[2] G. G. Blasdel and G. Salama. *Nature*, 321:579–585, 1986.

[3] D. W. DeMott. *Med Res Eng*, 5:23, 1966.

[4] G. M. Edelman. Cell adhesion molecules. *Science*, 219:450–457, 1983.

[5] T. Kohonen. Analysis of a simple self-organizing process. *Biol Cybern*, 44:135–140, 1982.

[6] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biol Cybern*, 43:59–69, 1982.

[7] T. Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag, New York, 1988.

[8] T. Martinetz, H. Ritter, and K. Schulten. Kohonen's self-organizing map for modeling the formation of the auditory cortex of a bat. In *Connectionism in Perspective*, pages 403–412. North-Holland, 1989.

[9] M. M. Merzenich, J. H. Kaas, J. T. Wall, M. Sur, R. J. Nelson, and D. J. Felleman. Progression of change following median nerve section in the cortical representation of the hand in areas 3b and 1 in adult owl and squirrel monkeys. *Neuroscience*, 10(3):639–665, 1983.

[10] M. M. Merzenich, R. J. Nelson, J. H. Kaas, M. P. Stryker, W. M. Jenkins, J. M. Zook, M. S. Cynader, and A. Schoppmann. Variability in hand surface representations in areas 3b and 1 in adult owl and squirrel monkeys. *Journal of Comparative Neurology*, 258:281–296, 1987.

[11] M. M. Merzenich, R. J. Nelson, M. P. Stryker, M. S. Cynader, A. Schoppmann, and J. M. Zook. Somatosensory cortical map changes following digit amputation in adult monkeys. *Journal of Comparative Neurology*, 224:591–605, 1984.

[12] J. Mittenthal. In A. Baskin and J. Mittenthal, editors, *The Principles of Organization in Organisms — Santa Fe Institute Studies in the Sciences of Complexity*, volume 12. Addison-Wesley, 1991.

[13] K. Obermayer, G. G. Blasdel, and K. Schulten. A neural network model for the formation of the spatial structure of retinotopic maps, orientation- and ocular dominance columns. In T. Kohonen et al., editors, *Artificial Neural Networks*, volume 1, pages 505–511. North Holland, 1991.

[14] K. Obermayer, H. Ritter, and K. Schulten. Large-scale simulations of self-organizing neural networks on parallel computers: Application to biological modelling. *Parallel Computing*, 14:381–404, 1990.

[15] K. Obermayer, H. Ritter, and K. Schulten. A principle for the formation of the spatial structure of cortical feature maps. *Proc Natl Acad Sci USA*, 87:8345–8349, 1990.

[16] K. Obermayer, H. Ritter, and K. Schulten. A model for the development of the spatial structure of retinotopic maps and orientation columns. In A. Baskin and J. Mittenthal, editors, *The Principles of Organization in Organisms — Santa Fe Institute Studies in the Sciences of Complexity*, volume 12. Addison-Wesley, 1991.

[17] H. Ritter and K. Schulten. On the stationary state of Kohonen's self-organizing sensory mapping. *Biol Cybern*, 54:99–106, 1986.

[18] H. Ritter and K. Schulten. Convergence properties of Kohonen's topology conserving mappings: Fluctuations, stability and dimension selection. *Biol Cybern*, 60:59–71, 1988.

[19] C. von der Malsburg and D. J. Willshaw. How to label nerve cells so that they can interconnect in an ordered fashion. *Proc Natl Acad Sci USA*, 74:5176–5178, 1977.