In Proceedings of ICCM - 2007- Eighth International Conference on Cognitive Modeling. 103 - 108. Oxford, UK: Taylor & Francis/Psychology Press.

The Emergence of Semantic Topography in a Neurally-Inspired Computational Model

Lee I. Newman (leenewm@umich.edu)
Thad A. Polk (tpolk@umich.edu)
Department of Psychology, 530 Church Street
Ann Arbor, MI 48109 USA

Abstract

Representations in sensory cortices are organized topographically: auditory cortex is organized tonotopically, somatosensory cortex is organized somatotopically, and visual cortex is organized retinotopically. Substantial progress has been made in understanding how topography develops at a neurocomputational level, particularly in the early and middle stages of processing in the visual system. We extend this work to investigate how higher-level semantic representations could develop based on topographic input from sensory maps in the ventral visual pathway. The receptive fields of cells in these maps correspond to the loci of activity within a cortical topography rather than explicitly coded sensory features. Using this model, we show that meaningful semantic representations at increasing levels of abstraction naturally emerge as a result of exposure to a set of visual stimuli. For example, when presented with a set of simple visual features (color, texture, size, and shape) the model develops semantic representations that distinguish basic level categories (e.g., dogs, tables, cars), superordinate categories (e.g., animals, furniture, vehicles), and living versus nonliving things. This work therefore offers a computationally explicit hypothesis about how semantic representations could emerge in the brain. Our results suggest the possibility that high-order concept representations may be encoded topographically much the same way as low-order sensory representations, and that these representations may be learned based on the same principles of neural computation known to be operating in sensory cortex.

Introduction

Topography is an important principle of neural organization and is present in all sensory cortices of the brain (Kandel, Schwartz, & Jessell, 2000). Topography represents a mapping from sensory space (e.g., location in the visual field, wavelength of light) to cortical space (e.g., location in extrastriate visual cortex). In such a mapping, nearby neurons in cortical space have similar receptive fields (they respond to nearby parts of sensory space) and are therefore selective for similar sensory features. Put simply, sensory similarity is reflected in cortical proximity.

Topographic representations have been particularly well-studied in the early processing stages in the ventral visual pathway of the primate brain. This pathway proceeds through a hierarchy of stages beginning with striate cortex (V1) and proceeding through extrastriate cortices (V2, V4) to inferotemporal cortex (IT). In the early stages of processing, receptive fields are small and neurons are highly selective for primitive sensory features (e.g., specific retinal locations, orientation of bars of light, and colors). As processing progresses receptive fields become increasingly large and neurons become selective for more abstract and complex features (e.g., specific configurations of features, and simple objects).

Physiological studies have established that topographies are present in the early and middle stages of the ventral pathway (V1, V2, V4) and the structure of these topographies has been well-described in the literature (Fujita, Tanaka, Ito, & Cheng, 1992; Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Livingstone & Hubel, 1984; Shipp & Zeki, 1989, 2002a, 2002b; Tanaka, 1996; Tootell, Silverman, Hamilton, Devalois, & Switkes, 1988; Tootell, Switkes, Silverman, & Hamilton, 1988; Vanessen & Gallant, 1994). The primary organization is retinotopic, with neighboring neurons coding for sensory information located at nearby positions in the visual field. Embedded within this retinotopy are secondary topographies such as orientation columns and color blobs whose neurons are selective for particular object orientations and colors, respectively. Physiological findings also suggest that a more complex object-based topography may exist in IT cortex, but neither the structure nor learning mechanisms underlying this putative topography are fully understood (Fujita et al., 1992; Sigala & Logothetis, 2002; Tanaka, 2000).

At a computational level, significant progress has been made in understanding how sensory topographies develop in the early and middle stages of the ventral pathway. Biologically plausible models based on self-organizing learning algorithms have simulated topographic development computationally and have been able to successfully reproduce physiological data (Barrow, Bray, & Budd, 1996; Carreira-Perpinan, Lister, & Goodhill, 2005; Goodhill, 1993; Goodhill & Willshaw, 1994; Olson & Grossberg, 1998; Sirosh & Miikkulainen, 1997; Sit & Miikkulainen, 2006). All these models are based on two well-established neural mechanisms, competition and Hebbian learning, that operate throughout neocortex.

In this work we consider what role, if any, these self-organizing mechanisms might play in the learning of higher-order semantics. Specifically, we use a computational model to investigate the type of representation that develops in a later stage cortical map that (i) receives topographically organized sensory inputs, and (ii) self-organizes based on the same mechanisms of neural computation known to operate in the earlier processing stages of the ventral pathway. We find that after exposing the model to a set of visual stimuli, the topography reflects multiple levels of semantic categories more than low-level visual similarities.

An early investigation of the role of topographic structure in semantics was conducted by Ritter & Kohonen (1989). In this work, the authors showed that a SOM could produce a topography of logical word roles (e.g., subject nouns, object nouns, verbs) based on the statistical structure of word context. Applied work in content retrieval (Laaksonen, Koskela, & Oja, 2002; Laaksonen & Viitaniemi, October, 2006) have successfully used SOMs to organize images based on visual similarity. Our model differs from prior work in that it focused on the visual modality, has a hierarchical structure more closely tied to known hierarchical structure of cortex, and has the objective of understanding psychological phenomena associated with semantics.

Self Organizing Maps

Self-organizing maps (SOMs) are a computational abstraction of cortical representation and processing (Kohonen, 1982, 1990). SOMs correspond to a locally connected population of neurons in a contiguous area of cortical tissue. As shown in Figure 1, the basic unit of representation within a SOM is the cell. Cells within a SOM are indexed based on their spatial location and each is modeled as a k-length *weight vector* specifying the preferred k-length input for the cell (i.e., the input that causes the cell to fire maximally). When presented with an input pattern, cells within a SOM compete to represent this pattern. The response of each cell is based on the similarity between its weight vector and the input pattern. The *winning cell* is the cell most similar to the input.

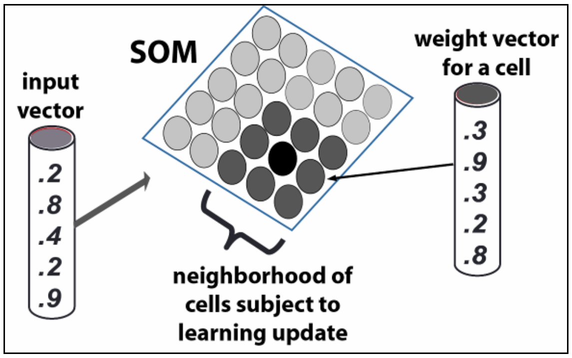

Figure 1: Self organizing map.

SOM learning is accomplished by modifying the weight vector of the winning cell so that it is more similar to the input, therefore making this cell more likely to win again with future presentations. Critically, the weight vectors of cells in close spatial proximity to the winning cells are also updated. As a result, with experience the response of a SOM tends to become spatially organized, learning a mapping from the statistical regularities discovered in its k-dimensional input space to a set of topographically organized neighborhoods within the SOM.

Mathematically, if $\boldsymbol{x}(t)$ is a vector representing the input to a SOM at time t and $\boldsymbol{w}_i(t)$ represents the weight vector of cell i at time t, then the winning cell c is given by

$$ArgMin_c \{\| \boldsymbol{x}(t) - \boldsymbol{w}_c(t) \|\}, \tag{1}$$

where the $\|\cdot\|$ operator denotes vector distance. We use Euclidean distance for all distance computations in our model. The weight vector for each cell i is updated according to the following SOM learning equation

$$\boldsymbol{w}_i(t+1) = \boldsymbol{w}_i(t) + \alpha(t)\, h_{c,i}(t)\{ \boldsymbol{x}(t) - \boldsymbol{w}_c(t) \}, \tag{2}$$

where $\alpha(t)$ is a time-dependent learning rate, and $h_{c,i}(t)$ is a kernel function that is centered on the winning cell c and that computes the magnitude of the update to cell i based on its spatial proximity to the winning cell. In our model we use the the Gaussian kernel function

$$h_{c,i}(t) = \exp\{-\frac{1}{\lambda(t)^2}\left\| \mathrm{i}_{(x,y)} - \mathrm{c}_{(x,y)} \right\|^2\}, \tag{3}$$

where $\lambda(t)$ is a time-dependent parameter that determines the width of the kernel, and $\mathrm{i}_{(x,y)}$ and $\mathrm{c}_{(x,y)}$ denote the map coordinates of cells i and c, respectively. We compute $\alpha(t)$ and $\lambda(t)$ as exponentially decreasing functions of time, with time denoting discrete presentations of training inputs.

While SOMs are not detailed biophysical models of cortex, the core assumptions embodied in this class of models can be mapped directly to their biophysical correlates: The weight vectors of SOM cells correspond biologically to the concept of receptive fields; winning cells are analogous to the peak location of activity within a cortical area, governed by the net result of the competitive interplay between local excitatory and inhibitory activity driven by an input; and the spatially-localized learning algorithm is a computational abstraction of Hebbian learning that occurs between neurons participating in a bump of cortical activity and the active afferent neurons providing their input. SOMs therefore provide a computationally efficient and biophysically plausible method for modeling the development of spatially structured representations in cortex.

Methods

Model Architecture

As shown in Figure 2, our model is a two-level hierarchy of SOMs. The first level consists of a set of four 10x10 *sensory maps* each corresponding to a particular visual feature: color, size, shape, and surface appearance. Each sensory map receives inputs in the form of real-valued sensory vectors, for example the color map receives three-vectors representing the hue, saturation and brightness of the stimuli. These maps are then exposed to a set of visual stimuli and are allowed to self-organize according to the SOM learning equations given by (1), (2) and (3).

The sensory maps capture the type of abstract featural representations found in the later stages of the ventral visual pathway when retinotopy is no longer present. With retinotopy no longer available as a basis of representation, subsequent cortical areas must somehow make sense of a more abstract sensory topography. It is important to note that the sensory maps in our model are not intended to explicitly instantiate the computations performed from early to late stage visual cortex (for one example of such a model, see Riesenhuber & Poggio, 1999; Serre, Oliva, & Poggio, 2007), but rather they serve as plausible summaries of how concrete visual information provided by a stimulus is ultimately encoded in abstract representations in the later stages of visual cortex.

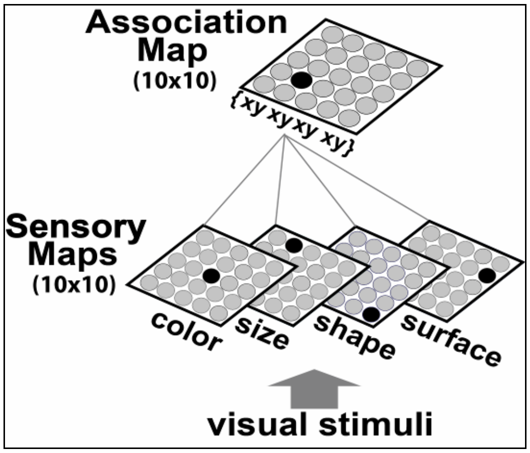

Figure 2: Architecture of the model.

The second level in the model consists of a single 10x10 *association map* that receives convergent inputs from the lower-level sensory maps. Computationally, the input to these maps is a concatenation of the spatial locations of activity in each of the sensory maps. The implication of this method of coding is that the receptive fields of cells in the association map correspond to *cortical coordinates* rather than explicitly encoded sensory-based features. We believe this is a subtle, but important aspect of the model. In the early stages of visual processing, the basis of neural organization is either driven by hardwired anatomical connections (as in the case of retinotopy) or by topography organized around concrete, low-level sensory-based features (as with orientation columns). In the later stages of visual processing, no such representational "boostrapping" is available. These later-stage cortical areas must somehow make sense of more abstract and complex representations based on the spatial location of activity in upstream maps. The association map in our model faces the same challenge: to learn meaningful representations based on spatially encoded inputs from the sensory maps.

Simulation Procedures

Stimuli The model was repeatedly presented with a set of 96 visual stimuli consisting of 8 classes of objects (bicycles, bushes, cars, cats, dogs, chairs, tables, and trees) and 12 variants within each class. The variants captured characteristic within-class featural differences, for example trees of varying color and size. Each stimulus was coded as a [0,1] normalized 10-vector based on its color (hue, saturation, brightness), size (x, y, and z dimension in a typical viewing angle), shape (roundness, complexity) and surface appearance (smoothness, textural uniformity). The vector values were estimated based on images collected from Google™ Images (http://images.google.com).

Model Learning Weight vectors of all cells in the SOMs were initialized to random values in the range [0.1, 0.9]. Learning then proceeded in two phases. First, each of the sensory SOMs were presented with the relevant vector component of the training stimuli (e.g., the color map was presented with hue, saturation and brightness values) and the weight vectors were updated according to equations (1), (2) and (3). In each of 500 learning iterations, the presentation order of the stimuli was randomized to minimize order effects. After training, the stimuli were presented to each of the sensory maps, and the map coordinates of the winning cells for each input were computed for each of the four maps. The four pairs of coordinates for each stimulus were then concatenated and the resulting vectors were then used for the second phase of learning. In this phase, the association map were trained using an identical procedure, with the exception that the input patterns presented to the association map were the concatenated outputs from the trained sensory maps.

Results

Figure 3 shows an example of the representation that is learned by one of the lower-level sensory maps, in this case, the map that was trained using size information for each stimulus[1]. The winning cell for each stimulus is labeled in the figure based on its object class. Inspection of this map reveals that there is no clear object-based topography, as only two of the object classes (cats and bikes) are co-located in the same region of the map. Instead, the map captures an abstract size-based topography in which larger objects are represented in the upper region of the map (smaller objects in the lower region), and taller objects are represented towards the left side of the map (wider objects towards the right). Similar results were found in each of the other sensory maps: an abstract sensory-based topography emerged, but stimuli from the same object class were not co-located in the map.

Figure 4 shows the response[2] of each cell in the size map when presented with one of the tree stimuli. As is evident, the map response to the stimulus is spatially localized to a neighborhood of activity surrounding the winning cell.

---

[1] Many of the variants within each object class share the same values and therefore not all 96 stimuli are visible due to overlap.

[2] Cell responses were computed based on the Euclidian distance between the stimulus and the cell's weight vector. Smaller distances imply greater similarity and larger cell responses.

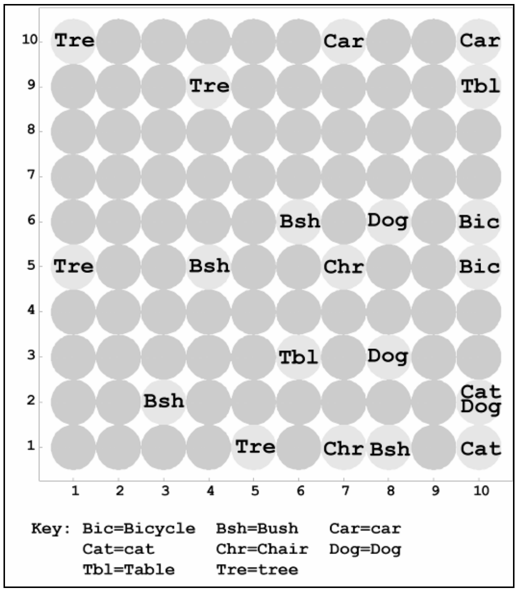

Figure 3: Sensory size map after learning.

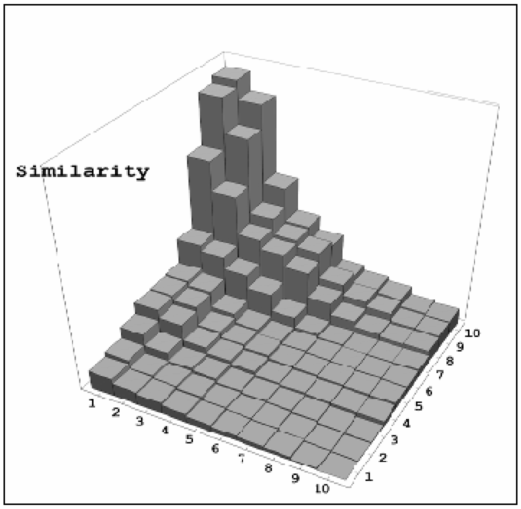

Figure 4: Response of the size map to a *tree* stimulus.

The structure of the association map after learning is shown in Figure 5. The winning cell for each stimulus is labeled in the figure based on object class. In contrast to the sensory maps which learned abstract sensory-based topographies, a clear object-based semantic topography emerged in the association map after learning. For example, a neighborhood of cells responsive to tables emerged in the lower left region of the map, and a set of cells most responsive to trees are spatially co-located in the upper right region. For some object classes, the neighborhoods overlap, as in the case of dogs and cats, and trees and bushes, indicating that the map was unable to distinguish these classes of stimuli based on available visual information. Nevertheless, for all object classes the map learned to represent the class in a spatially co-located region of cells.

Further inspection of the learned association map in Figure 5 reveals a semantic topography that captures superordinate categorical distinctions among the object classes. For example, in the lower-left region of the map there is a neighborhood of cells responsive to chairs and tables (furniture), and similarly there are distinct regions of cells whose receptive fields prefer trees and bushes (plants), bikes and cars (vehicles), and dogs and cats (animals). Furthermore, the topography encoded in the map also learned the semantic distinction between living and non-living stimuli. Cells representing dogs, cats, trees, and bushes (living things) are co-located in the upper region of the map and tables, chairs, cars, and bikes (non-living things) are co-located in neighborhood of cells in the lower region of the map.

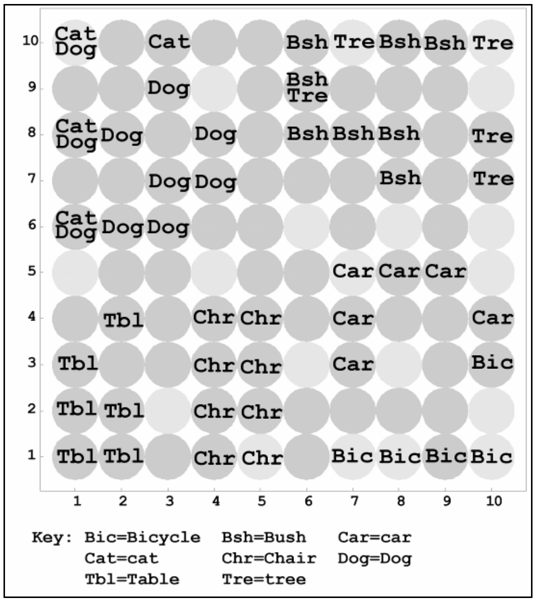

Figure 5: Association map after learning.

The responses of the association map to two sets of similar stimuli are shown in Figure 6. Each class of stimulus shown in the figure (table, chair, tree, and bush) produces a graded, locally organized response around the winning cell. Furthermore, the responses of the map to similar classes of stimuli (table and chair, tree and bush) share a similar subset of active cells, demonstrating the existence of topography in the form of neighborhoods of cells responsive to stimuli at multiple levels of abstraction (object class and superordinate category).

Although these results confirm the hypothesis that high-level semantic topography can be driven by the same principles of neural computation found in the lower levels of the visual pathway, we raise two concerns. First, it is possible that these results are an artifact of the learning procedure due to the random initialization of the maps and/or the random ordering in which the stimuli were presented. To address this concern, we simulated the model over a large number of random seeds and confirmed that the structure of the learned topography in the association map was consistent across all simulations. Although the specific location of the category clusters varied, the clusters themselves reliably emerged as did the superordinate organization (plants, animals, furniture, and vehicles) and the living versus non-living distinction.

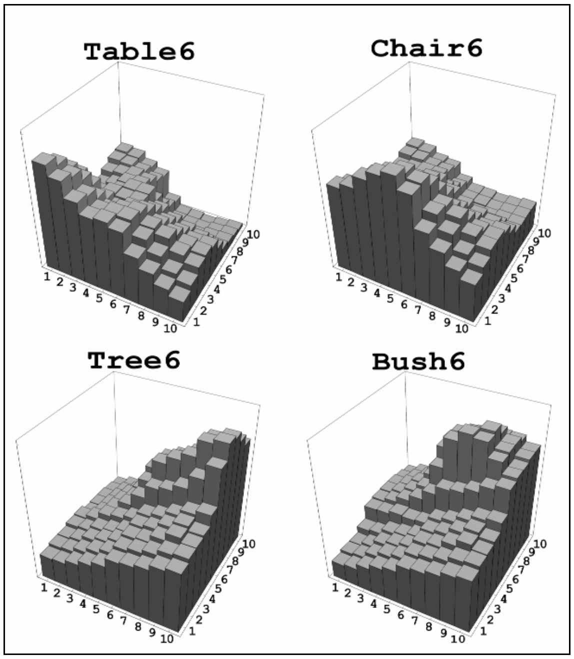

Figure 6: Association map responses.

A second concern is that the semantic organization shown in Figure 5 is an artifact of the way that the stimuli were coded. It is possible that the 96 stimulus vectors were trivially separable at multiple levels of abstraction, and therefore the resultant topography was in some sense predetermined. To address this concern, we performed a hierarchical clustering[3] of the training stimuli, producing the tree shown in Figure 7. As is evident, hierarchical clustering does not produce the same meaningful categorical clusters of stimuli produced in our model: At the level of object classes, only bicycles are distinguished from the other classes (i.e., they share a common branch within the hierarchy); at the level of superordinate category, only plants and animals are independently clustered; and at the highest level, the clustering categorically generates two heterogeneous groups: plants-bicycles and cars-tables-chairs-animals.

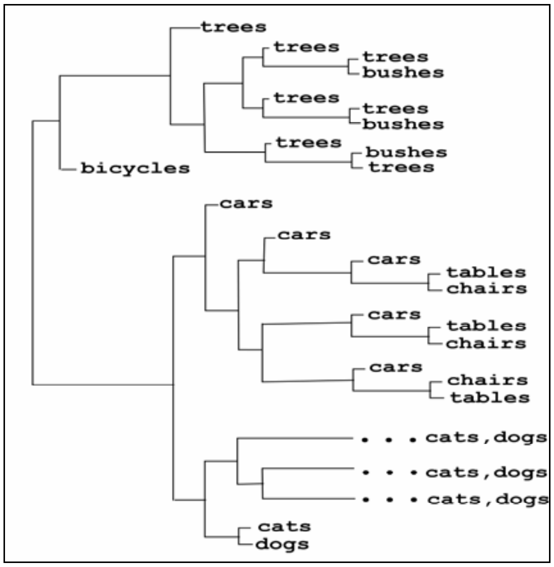

Figure 7: Clustering tree of the training stimuli.

Discussion

The aim of this work was a plausibility proof of the hypothesis that the principles of self-organization and cortical topography at work in lower-level visual processing may also drive the learning and structure of higher-order semantics. We developed an explicit computational model in which a self-organizing map receives inputs from a hierarchy of topographic sensory maps and we found that meaningful semantic representations at increasing levels of abstraction naturally emerge as a result of exposure to a set of sensory stimuli. Specifically, when presented with a set of simple visual features (color, size, texture, shape) the higher-level map develops a topography of semantic representations that distinguishes basic level categories (bicycles, bushes, cars, cats, dogs, chairs, tables, and trees), superordinate categories (plants, animals, furniture, vehicles), and living versus nonliving things.

This work therefore offers a computationally explicit hypothesis about how semantic representations could emerge in the association cortex of the brain. Our results may also be relevant to the ongoing debate about whether conceptual knowledge is primarily organized anatomically by modality as posited by the sensory-functional hypothesis (Farah & McClelland, 1991; Warrington & McCarthy, 1983; Warrington & Shallice, 1984), or by category (Caramazza & Mahon, 2003). The results of our simulations suggest fundamental mechanisms of neural computation could lead to the emergence of topographically organized semantic representations without the need to posit that these representations are innate.

---

[3] The clusters were computed using the hierarchical clustering algorithm in Mathematica® using a Euclidian distance function.

References

Barrow, H. G., Bray, A. J., & Budd, J. M. L. (1996). A self-organizing model of "color blob" formation. *Neural Computation, 8*(7), 1427-1448.

Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends in Cognitive Sciences, 7*(8), 354-361.

Carreira-Perpinan, M. A., Lister, R. J., & Goodhill, G. J. (2005). A computational model for the development of multiple maps in primary visual cortex. *Cerebral Cortex, 15*(8), 1222-1233.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment - modality specificity and emergent category specificity. *Journal of Experimental Psychology-General, 120*(4), 339-357.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature, 360*(6402), 343-346.

Goodhill, G. J. (1993). Topography and ocular dominance - a model exploring positive correlations. *Biological Cybernetics, 69*(2), 109-118.

Goodhill, G. J., & Willshaw, D. J. (1994). Elastic net model of ocular dominance - overall stripe pattern and monocular deprivation. *Neural Computation, 6*(4), 615-621.

Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P., & Tootell, R. B. H. (1998). Retinotopy and color sensitivity in human visual cortical area v8. *Nature neuroscience, 1*(3), 235-241.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science.* New York: McGraw-Hill, Health Professions Division.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*(1), 59-69.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE, 78*(9), 1464-1480.

Laaksonen, J., Koskela, M., & Oja, E. (2002). Picsom - self-organizing image retrieval with mpeg-7 content descriptors. *Ieee Transactions on Neural Networks, 13*(4), 841-853.

Laaksonen, J., & Viitaniemi, V. (October, 2006). *Emergence of ontological relations from visual data with self-organizing maps.* Paper presented at the 9th Scandinavian Conference on Artificial Intelligence, Espoo, Finaland.

Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual-cortex. *Journal of Neuroscience, 4*(1), 309-356.

Olson, S. J., & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks, 11*(2), 189-208.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019-1025.

Ritter, H., & Kohonen, T. (1989). Self-organizing semantic maps. *Biological Cybernetics, 61*(4), 241-254.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the United States of America, 104*(15), 6424-6429.

Shipp, S., & Zeki, S. (1989). The organization of connections between areas v5 and v1 in macaque monkey visual-cortex. *European Journal of Neuroscience, 1*(4), 309-332.

Shipp, S., & Zeki, S. (2002a). The functional organization of area v2, i: Specialization across stripes and layers. *Visual neuroscience, 19*(2), 187-210.

Shipp, S., & Zeki, S. (2002b). The functional organization of area v2, ii: The impact of stripes on visual topography. *Visual neuroscience, 19*(2), 211-231.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*(6869), 318-320.

Sirosh, J., & Miikkulainen, R. (1997). Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Computation, 9*(3), 577-594.

Sit, Y. F., & Miikkulainen, R. (2006). Self-organization of hierarchical visual maps with feedback connections. *Neurocomputing, 69*(10-12), 1309-1312.

Tanaka, K. (1996). Representation of visual features of objects in the inferotemporal cortex. *Neural Networks, 9*(8), 1459-1475.

Tanaka, K. (2000). Mechanisms of visual object recognition studied in monkeys. *Spatial vision, 13*(2-3), 147-163.

Tootell, R. B. H., Silverman, M. S., Hamilton, S. L., Devalois, R. L., & Switkes, E. (1988). Functional-anatomy of macaque striate cortex .3. Color. *Journal of Neuroscience, 8*(5), 1569-1593.

Tootell, R. B. H., Switkes, E., Silverman, M. S., & Hamilton, S. L. (1988). Functional-anatomy of macaque striate cortex .2. Retinotopic organization. *Journal of Neuroscience, 8*(5), 1531-1568.

Vanessen, D. C., & Gallant, J. L. (1994). Review: Neural mechanisms of form and motion processing in the primate visual-system. *Neuron, 13*(1), 1-10.

Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain, 106*(DEC), 859-878.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(SEP), 829-854.