

Dendritic Inhibition Enhances Neural Coding Properties

M. W. Spratling and M. H. Johnson

Centre for Brain and Cognitive Development, Birkbeck College, London. UK.

Abstract

The presence of a large number of inhibitory contacts at the soma and axon initial segment of cortical pyramidal cells has inspired a large and influential class of neural network model which use post-integration lateral inhibition as a mechanism for competition between nodes. However, inhibitory synapses also target the dendrites of pyramidal cells. The role of this dendritic inhibition in competition between neurons has not previously been addressed. We demonstrate, using a simple computational model, that such pre-integration lateral inhibition provides networks of neurons with useful representational and computational properties which are not provided by post-integration inhibition.

Introduction

Lateral inhibition between cortical excitatory cells plays an important role in determining the receptive field properties of those cells. Such lateral inhibition provides a mechanism through which cells compete to respond to the current pattern of stimulation. Inhibitory inputs are concentrated on the soma and axon initial segment of pyramidal cells (Somogyi and Martin, 1985; Mountcastle, 1998) where they can be equally effective at inhibiting responses to excitatory inputs stimulating any part of the dendritic tree.

This observation has formed the basis for many theories of receptive field formation, and is an essential feature of many computational (neural network) models of cortical function (Swindale, 1996; von der Malsburg, 1973; Földiák, 1989, 1990; Marshall, 1995; Sirosh and Miikkulainen, 1994; Oja, 1989; Sanger, 1989; O'Reilly, 1998; Rumelhart and Zipser, 1985; Kohonen, 1997; Grossberg, 1987; Ritter et al., 1992; Hertz et al., 1991; Földiák, 1991; Wallis, 1996). Such neural network algorithms have also found application beyond the neurosciences as a means of data analysis, classification and visualization in a huge variety of fields. These algorithms vary greatly in the details of their implementation. In some, competition is achieved explicitly by using lateral connections between the nodes of the network (Swindale, 1996; von der Malsburg, 1973; Földiák, 1989, 1990; Marshall, 1995; Sirosh and Miikkulainen, 1994; Oja, 1989; Sanger, 1989; O'Reilly, 1998), while in others competition is implemented implicitly through a selection process which chooses the 'winning' node(s) (Rumelhart and Zipser, 1985; Kohonen, 1997; Grossberg, 1987; Ritter et al., 1992; Hertz et al., 1991; Földiák, 1991; Wallis, 1996). However, in all of these algorithms nodes compete for the right to generate a response to the current pattern of input activity. A node's success in this competition is dependent on the total strength of the stimulation it receives and nodes which compete unsuccessfully have their output activity suppressed. This class of models can thus be described as implementing 'post-integration inhibition'.

Inhibitory contacts also occur on the dendrites of cortical pyramidal cells (Kim et al., 1995; Rockland, 1998) and certain classes of interneuron (*e.g.*, double bouquet cells) specifically target dendritic spines and shafts (Mountcastle, 1998; Tamas et al., 1997). Such contacts would have relatively little impact on excitatory inputs more proximal to the cell body or on the action of synapses on other branches of the dendritic tree. Thus these synapses do not appear to contribute to post-integration inhibition. However, such synapses are likely to have strong inhibitory effects on inputs within the same dendritic branch that are more distal to the site of inhibition (Borg-Graham et al., 1998; Koch et al., 1983; Rall, 1964; Segev, 1995; Koch and Segev, 2000). Hence, they could potentially selectively inhibit specific groups of excitatory inputs. Related synapses cluster together within the dendritic tree so that local operations are performed by multiple, functionally distinct, dendritic subunits before integration at the soma (Koch and Segev, 2000; Segev and Rall, 1998; Segev, 1995; Häusser et al., 2000; Häusser, 2001; Mel, 1994, 1999). Dendritic inhibition could thus act to 'block' the output from individual functional compartments. It has long been recognized that a dendrite composed of multiple subunits would provide a significant enhancement to the computational powers of an individual neuron (Mel, 1993, 1994, 1999) and that dendritic inhibition could contribute to this enhancement (Koch et al., 1983; Koch and Segev, 2000; Segev and Rall, 1998). However, the role of dendritic inhibition in competition between cells and its subsequent effect on neural coding and receptive field properties has not previously been investigated.

We introduce a neural network model which demonstrates that competition via dendritic inhibition significantly enhances the computational properties of networks of neurons. As with models of post-integration inhibition we simplify reality by combining the action of inhibitory interneurons into direct inhibitory connections

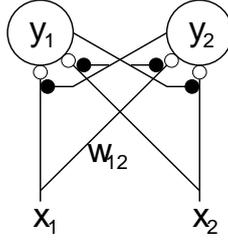


Figure 1: A network competing through pre-integration lateral inhibition. Nodes are shown as large circles, excitatory synapses as small open circles and inhibitory synapses as small filled circles.

between nodes. Furthermore, we group all the synapses contributing to a dendritic compartment together as a single input. Dendritic inhibition is then modeled as (linear) inhibition of this input. The algorithm is described fully in the Methods section, but essentially, it operates by causing each node to attempt to ‘block’ its preferred inputs from activating other nodes. It is thus described as ‘pre-integration inhibition’.

We illustrate the advantages of this form of competition with the aid of a few simple tasks which have been used previously to demonstrate the pattern recognition abilities required by models of the human perceptual system (Marshall, 1995; Marshall and Gupta, 1998; Nigrin, 1993). Although these tasks appear to be trivial, succeeding in all of them is beyond the abilities of single-layer neural networks using post-integration inhibition. These tasks demonstrate that pre-integration inhibition (in contrast to post-integration inhibition) enables a neural network to respond simultaneously to multiple stimuli, to distinguish overlapping stimuli, and to deal correctly with incomplete and ambiguous stimuli.

Methods

A simple, two-node, neural network in which there is pre-integration inhibition is shown in figure 1. The essential idea is that each node inhibits other nodes from responding to the same inputs. Hence, if a node is active and it has a strong synaptic weight to a certain input then it should inhibit other nodes from responding to that input. A simple implementation of this idea for a two-node network would be:

$$y_1 = \sum_{i=1}^m (w_{i1}x_i - \alpha w_{i2}y_2)^+$$

$$y_2 = \sum_{i=1}^m (w_{i2}x_i - \alpha w_{i1}y_1)^+ .$$

Where y_j is the activation of node j , w_{ij} is the synaptic weight from input i to node j , x_i is the activation of input i , α is a scale factor controlling the strength of lateral inhibition, and $(v)^+ = v$ if $v \geq 0$, $(v)^+ = 0$ otherwise. These simultaneous equations are solved iteratively, with the value of α gradually increasing at each iteration, from an initial value of zero. Hence, initially each node responds independently to the stimulus, but as α increases the node activations are modified by competition. Steady-state activity is reached (at large α) when each individual input contributes to the activation of (at most) a single node.

In order to apply pre-integration lateral inhibition to larger networks a more complex formulation was used which is suitable for networks containing an arbitrary number of nodes (n) and receiving an arbitrary number of inputs (m):

$$y_j = \sum_{i=1}^m w_{ij}x_i \left(1 - \alpha \max_{\substack{k=1 \\ (k \neq j)}}^n \left\{ \frac{w_{ik}}{\max_{l=1}^m \{w_{lk}\}} \frac{y_k}{\max_{l=1}^n \{y_l\}} \right\} \right)^+ .$$

This formulation was used to produce all the results presented in this paper. Synaptic weights were normalized such that $\sum_{i=1}^m w_{ij} = 1$. The value of α was increased from zero to ten in steps of 0.25. Activation values reached a steady-state at lower alpha (≈ 2) and remained constant from then on. The step size was found to be immaterial to the final steady-state activation values provided it was less than 0.5.

For the simulation shown in figure 5 a bias was added to the activation of one node. This was implemented by adding 0.1 to the activation of that node during competition. Experiments showed that this bias could occur at any time (and for any duration) prior to α reaching a value of 1.5 to generate the same result.

Although results have not been shown here this method is not restricted to working with binary encodings of input patterns and works equally well with analog encodings.

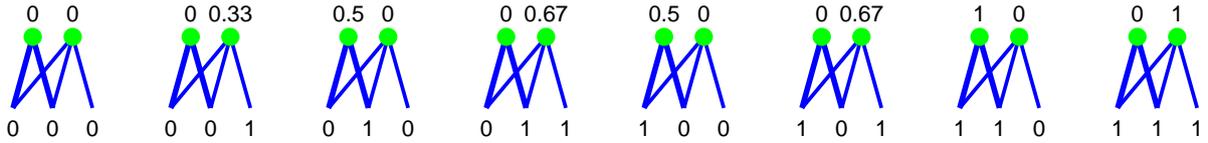


Figure 2: Representing overlapping input patterns. A network consisting of two nodes and three inputs ('a', 'b', and 'c') is wired up so that the first node receives input from 'a' and 'b' (with weight $\frac{1}{2}$ from each) and the second node receives input from all three sources (with weight $\frac{1}{3}$ from each). The response of the network to each possible pattern of inputs is shown. Pre-integration lateral inhibition (lateral weights have been omitted from the figures) enables each node to respond exclusively to its preferred pattern: *i.e.*, either 'ab' (110) or 'abc' (111). Other input patterns cause a weaker response from that node which has the closest matching preferred input.

Results

Overlap

In many situations distinct sensory events will share many features in common. If such situations are to be distinguished it is necessary for different sets of neurons to respond despite this overlap in input features. As a simple example, consider the task of representing two overlapping patterns: 'ab' and 'abc'. A network consisting of two nodes receiving input from three sources (labelled 'a', 'b' and 'c') should be sufficient. However, because these input patterns overlap, when the pattern 'ab' is presented the node representing 'abc' will be partially activated, while when the pattern 'abc' is presented the node representing 'ab' will be fully activated.

When the synaptic weights have certain values both nodes will respond with equal strength to the same pattern. For example, when the weights are all equal, both nodes will respond to pattern 'ab' with equal strength (Marshall, 1995). Similarly, when the total synaptic weight *from* each input is normalized ('post-synaptic normalization') both nodes will respond equally to pattern 'ab' (Marshall, 1995). When the total synaptic weight *to* each node is normalized ('pre-synaptic normalization') both nodes will respond to pattern 'abc' with equal activation (Marshall, 1995). Under all these conditions the response fails to distinguish between distinct input patterns and post-integration inhibition can do nothing to resolve the situation (and will, in general, result in a node chosen at random winning the competition).

Several solutions to this problem have been suggested. Some require adjusting the activations using a function of the total synaptic weight received by the node (*i.e.*, using the Webber Law (Marshall, 1995) or a masking field (Cohen and Grossberg, 1987; Marshall, 1995)). These solutions scale badly with the number of overlapping inputs, and do not work when (as is common practice in many neural network models) the total synaptic weight to each node is normalized. Other suggestions have involved tailoring the lateral weights to ensure the correct node wins the competition (Marshall, 1995; Földiák, 1990). These methods work well (Marshall, 1995), but fail to meet other criteria as discussed below.

The most obvious, but most overlooked, solution would be to remove constraints placed on allowable values for synaptic weights (*e.g.*, normalization) which serve to prevent the input patterns being distinguished in weight space. It is simple to invent sets of weights which unambiguously classify the two overlapping patterns (*e.g.*, if both weights to the node representing 'ab' are 0.5 and each weight to the node representing 'abc' are 0.4 then each node responds most strongly to its preferred pattern and could then successfully inhibit the activation of the other node).

Using pre-integration lateral inhibition overlapping patterns can be successfully distinguished even when normalization is used (either pre- or post-synaptic normalization). Figure 2 shows the response of such a network to all possible input patterns. The two networks on the right show that the correct response is generated to input patterns 'ab' and 'abc'. The other networks show that when partial input patterns are presented the node which represents the most similar pattern is activated in proportion to the degree of overlap between the partial pattern and the preferred input of that node. Hence, when the input is 'a' or 'b', which partially matches both of the training patterns, then the node representing the smallest pattern responds since these partial patterns are more similar to 'ab' than to 'abc'. When the input is 'c' this partially matches only one of the training patterns and hence the node representing 'abc' responds. Similarly, patterns 'bc' and 'ac' most strongly resemble 'abc' and hence cause activation of that node.

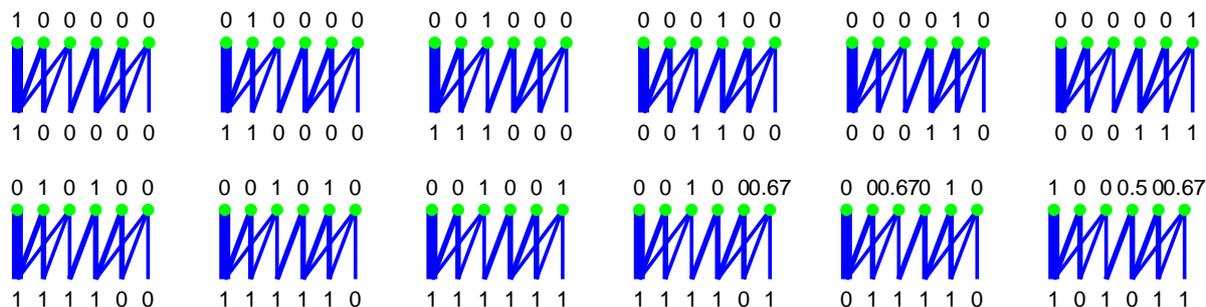


Figure 3: Representing multiple, overlapping, input patterns. A network consisting of six nodes and six inputs ('a', 'b', 'c', 'd', 'e', and 'f') is wired up so that nodes receive input from patterns 'a', 'ab', 'abc', 'cd', 'de', and 'def'. The response of the network to each of these input patterns is shown on the top row. Pre-integration lateral inhibition (lateral weights have been omitted from the figures) enables each node to respond exclusively to its preferred pattern. In addition, the response to multiple and partial patterns is shown on the bottom row. Pattern 'abcd' causes the nodes representing 'ab' and 'cd' to be active simultaneously, despite the fact that this pattern overlaps strongly with pattern 'abc'. Input 'abcde' is parsed as 'abc' together with 'de', and input 'abcdef' is parsed as 'abc' + 'def'. Input 'abcdf' is parsed as 'abc' + two-thirds of 'def', hence the addition of 'f' to the pattern 'abcd' radically changes the representation that is generated. Input 'bcde' is parsed as two-thirds of 'abc' plus pattern 'de'. Input 'acef' is parsed as 'a' + one half of 'cd' + two-thirds of pattern 'def'.

Multiplicity

While it is sufficient in certain circumstances for a single node to represent the input (local coding) it is desirable in many other situations to have multiple nodes providing a factorial or distributed representation. As an extremely simple example consider three inputs ('a', 'b' and 'c') each of which is represented by one of three nodes. Any pattern of inputs can be represented by having zero, one or multiple nodes active. In this particular case the input to the network provides just as good a representation as the output so there is little to be gained. However, this example captures the essence of other, more realistic, tasks in which multiple nodes, each of which represent multiple inputs, may need to be active

Post-integration lateral inhibition can be modified to enable multiple nodes to be active (Marshall, 1995; Földiák, 1990) by weakening the strength of the competition between those pairs of nodes that require to be coactive (the lateral weights need to reach a compromise strength which provides sufficient competition for distinct patterns while allowing multiple nodes to respond to multiple patterns). This either requires *a priori* knowledge of which nodes will be coactive or the ability to learn appropriate lateral weights. However, information locally available at a synapse is insufficient to determine if the correct compromise weights have been reached (Spratling, 1999) and it is thus necessary to add further constraints to derive a learning rule. The proposed constraints require that all input patterns occur with equal probability and that pairs of nodes are coactive with equal frequency (Marshall, 1995; Földiák, 1990). These constraints severely restrict the class of problems that can be successfully represented to those in which all input patterns are mutually exclusive or in which all pairs of input patterns occur simultaneously with equal frequency. As an example of a case for which these networks would fail, consider using a single network to represent the color and shape of an object. At any given time only one node (or group of nodes) representing a single color and one node (or group of nodes) representing a single shape should be active. There thus needs to be strong inhibition between nodes representing properties within the same class, and weak inhibition between nodes representing different properties. This task fails to match the requirements implicitly defined in the learning rules, and application of those rules would lead to weakening of lateral inhibition within each class until multiple color nodes and multiple shape nodes were coactive with equal frequency. Hence, post-integration lateral inhibition, implemented using explicit lateral weights, fails to provide factorial coding except for the exceptional case in which all pairs of patterns co-occur together, or in which external knowledge is available to set appropriate lateral weights.

Networks in which competition is implemented using a selection mechanism can also be modified to allow multiple nodes to be simultaneously active (*e.g.*, *k*-winners-takes-all). However, these networks also place restrictions on the types of task that can be successfully represented to those in which a pre-defined number of nodes need to be active in response to every pattern of stimuli.

In contrast, pre-integration lateral inhibition places no restrictions on the number of active nodes, nor on the

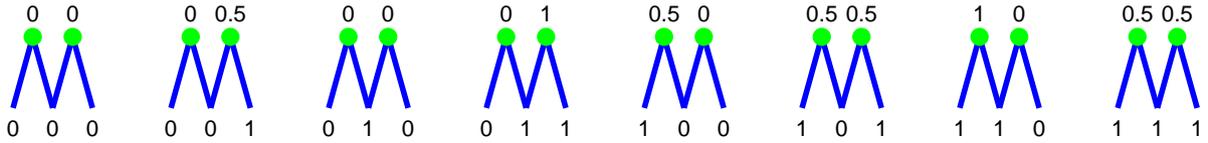


Figure 4: Representing ambiguous input patterns. A network consisting of two nodes and three inputs ('a', 'b', and 'c') is wired up so that the first node receives input from 'ab' and the second node receives input from 'bc' (all weights have a value of $\frac{1}{2}$). The response of the network to each possible pattern of inputs is shown. Pre-integration lateral inhibition (lateral weights have been omitted from the figures) suppresses any response to pattern 'b' (010) which overlaps equally with each node's preferred input pattern. Similarly, when the input is 'abc' the ambiguous contribution from input 'b' is suppressed so that both nodes respond at half strength. It can be seen that in other conditions each node responds at half strength when the input matches half its preferred input, and at full strength when its preferred input is presented.

frequency which which nodes, or pairs of nodes, are active. Such a network can thus respond appropriately to any combination of input patterns; for example, it can directly solve the problem of representing any arbitrary combination of the inputs 'a', 'b' and 'c'. A more challenging problem is shown in figure 3. Here nodes represent six overlapping patterns. The network responds correctly to each of these patterns and to multiple, overlapping, patterns (even in case where only partial patterns are presented).

Ambiguity

In some circumstances there simply is no correct parsing of the input pattern. Consider a neural network with two nodes and three inputs ('a', 'b' and 'c'). If one node represents the pattern 'ab' and the other represents the pattern 'bc' then the input 'b' is ambiguous since it equally matches the preferred input of both nodes. In this situation, most implementations of post-synaptic lateral inhibition would allow one node, chosen at random, to be active at half its normal strength. An alternative implementation (Marshall, 1995) is to use weaker lateral weights to enable both nodes to respond with one-quarter of the maximum response (Marshall and Gupta, 1998). However, this approach is also unsatisfactory since it suggests that one-quarter of each pattern is present, when this is not the case. Neither of these activity patterns seem to provide an appropriate representation. Any response in which both nodes generate equal activity suggests that a single piece of data provides evidence for two interpretations simultaneously. While any response in which one node has higher activity than the other is making an unjustified, arbitrary, selection. Pre-integration lateral inhibition avoids generating responses that are not justified by the available data by preventing any response (Figure 4). It thus produces no representation of the input rather than a potentially misleading representation.

As an example of a situation in which such an approach would be advantageous consider again using a network to represent the color and shape of an object. However, in this situation the network is wired up to generate localist representations of conjunctions of color and shape from a distributed input representation of these separate features. For example, consider a network with four nodes representing 'black-squares', 'white-squares', 'black-triangles' and 'white-triangles' (with the inputs to this network signaling 'black', 'white', 'square' and 'triangle'). In this case the ambiguous situation occurs when multiple objects are presented to the network simultaneously: a black-square and a white-triangle would cause an identical input pattern as a black-triangle and a white-square (Thorpe, 1995). Given such a situation it is important to prevent illusory conjunctions from being represented (Roelfsema et al., 2000), pre-integration lateral inhibition does so by suppressing all responses (Figure 5). One solution to this 'binding' problem would be the action of expectation or attention in disambiguating the situation (Reynolds and Desimone, 1999; Roelfsema et al., 2000). If such modulatory effects are modeled by adding a small increase to the activity of one node during competition then this succeeds in causing a response from those nodes compatible with the biased interpretation, while suppressing activity in the other two nodes (Figure 5). A similar bias applied to a network using post-integration inhibition would cause the biased node to be the most active, but would also suppress the response of the node representing the second object. An alternative solution would be for inputs representing the features of one object to be active simultaneously but out-of-phase with those inputs representing the other object (von der Malsburg, 1981; Gray, 1999; Singer, 1999). In this case the network succeeds (as would a network using the standard method of competition) by responding alternately to the non-ambiguous patterns generated by each individual object presented in isolation.

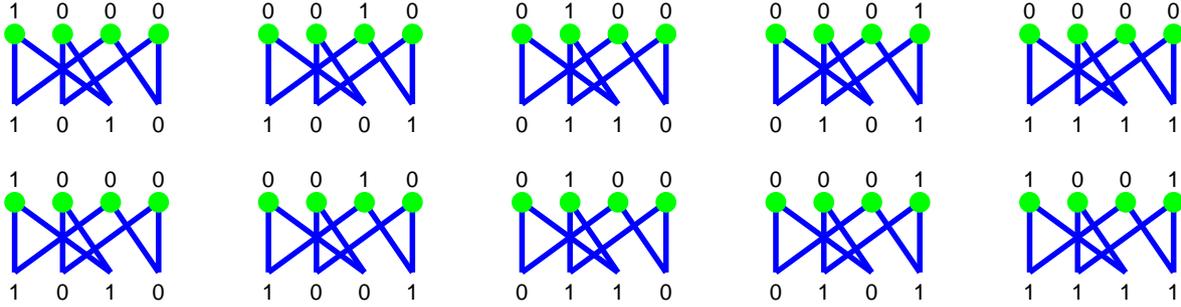


Figure 5: Representing feature conjunctions. A network consisting of four nodes and four inputs (‘black’, ‘white’, ‘square’ and ‘triangle’) is wired up so that the first node receives input from ‘black-square’, the second from ‘white-square’ the third from ‘black-triangle’ and the fourth from ‘white-triangle’ (all weights have a value of $\frac{1}{2}$). The first four figures in the top row show the response of the network to valid conjunctions of features from a single object. The last figure in the top row shows the response to an ambiguous input that could either be caused by the presentation of a black-square and a white-triangle, or by a black-triangle and a white-square. The second row shows responses to the same inputs as used in first row, but with the first node (which represents ‘black-squares’) receiving a small bias input during competition. It can be seen that for input patterns where activation of the first node is not justified by the input the bias has no effect on the outcome. However, for the ambiguous case the bias causes a parsing of the input into ‘black-square’ + ‘white-triangle’.

Discussion

The above examples have shown that pre-integration lateral inhibition provides useful computational capacities that can not be generated using post-integration lateral inhibition. A network of neurons competing through pre-integration lateral inhibition is thus capable of generating correct representations based on the ‘knowledge’ stored in the synaptic weights of the neural network. Specifically, it is capable of generating a local encoding of individual input patterns as well as responding simultaneously to multiple patterns, when they are present, in order to generate a factorial or distributed encoding. It can produce an appropriate representation even when patterns overlap. It is able to respond to partial patterns such that the response is proportional to how well that input matches the stored pattern, and it can detect ambiguities and suppress responses to them.

Our algorithm simplifies reality by assuming that the role of inhibitory cells can be approximated by direct inhibitory weights from excitatory cells, and that these lateral weights have the same strength as corresponding afferent weights. The latter simplification can be justified since weights that have identical values also have identical pre- and post-synaptic activation values and hence could be learnt independently. Such a learning mechanism would require inhibitory synapses contacting the dendrite to be modified as a function of the local dendritic activity rather than the output activity of the inhibited cell. More complex models, which include a separate inhibitory cell population, and which use multi-compartmental models of dendritic processes could relate our proposal more directly with physiology. We hope that our demonstration of the computational and representational advantages that could arise from dendritic inhibition will serve to stimulate such more detailed studies.

Computational considerations have led us to suggest that competition via dendritic inhibition could significantly enhance the information processing capacities of networks of cortical neurons. This claim is anatomically plausible since it has been shown that cortical pyramidal cells innervate inhibitory cell types which in turn form synapses on the dendrites of pyramidal cells (Buhl et al., 1997; Tamas et al., 1997). However, determining the functional role of these connections will require further experimental evidence. Our model predicts that it should be possible to find pairs of cortical pyramidal cells for which action potentials generated by one cell induce inhibitory post-synaptic potentials within the dendrites of the other. Independent of such experimental support, the algorithm we have presented could have immediate advantages for a great number of neural network applications in a huge variety of fields.

Acknowledgements

This work was funded by MRC Research Fellowship number G81/512.

References

- Borg-Graham LT, Monier C, Fregnac, Y (1998) Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* 393(6683):369–373.
- Buhl EH, Tamas G, Szilagyí T, Stricker C, Paulsen O, Somogyi P (1997) Effect, number and location of synapses made by single pyramidal cells onto aspiny interneurons of cat visual cortex. *J Physiol* 500(3):689–713.
- Cohen MA, Grossberg S. (1987) Masking fields: a massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Appl Optics* 26:1866–1891.
- Földiák P (1989) Adaptive network for optimal linear feature extraction. In: *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks*, volume 1, pp. 401–405. New York: IEEE Press.
- Földiák P (1990) Forming sparse representations by local anti-Hebbian learning. *Biol Cybern* 64:165–170.
- Földiák P (1991) Learning invariance from transformation sequences. *Neural Comput* 3:194–200.
- Gray CM (1999) The temporal correlation hypothesis of visual feature integration: still alive and well. *Neuron* 24(1):31–47.
- Grossberg S (1987) Competitive learning: from interactive activation to adaptive resonance. *Cognitive Sci* 11:23–63.
- Häusser M (2001) Synaptic function: dendritic democracy. *Curr Biol* 11:R10–12.
- Häusser M, Spruston N, Stuart GJ (2000) Diversity and dynamics of dendritic signalling. *Science* 290(5492):739–744.
- Hertz J, Krogh A, Palmer RG (1991) *Introduction to the Theory of Neural Computation*. Redwood City, California: Addison-Wesley.
- Kim HG, Beierlein M, Connors BW (1995) Inhibitory control of excitable dendrites in neocortex. *J Neurophysiol* 74(4):1810–1814.
- Koch C, Poggio T, Torre V (1983) Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc Natl Acad Sci USA* 80(9):2799–2802.
- Koch K, Segev I (2000) The role of single neurons in information processing. *Nature Neurosci [Suppl]* 3:1171–1177.
- Kohonen T (1997) *Self-Organizing Maps*. Berlin: Springer.
- Marshall JA (1995) Adaptive perceptual pattern recognition by self-organizing neural networks: context, uncertainty, multiplicity, and scale. *Neural Networks* 8(3):335–362.
- Marshall JA, Gupta VS (1998) Generalization and exclusive allocation of credit in unsupervised category learning. *Network - Comput Neural Syst* 9(2):279–302.
- Mel BW (1993) Synaptic integration in an excitable dendritic tree. *J Neurophysiol* 70(3):1086–1101.
- Mel BW (1994) Information processing in dendritic trees. *Neural Comput* 6:1031–1085.
- Mel BW (1999) Why have dendrites? A computational perspective. In: *Dendrites* (Stuart G, Spruston N, Häusser M, eds.), pp. 271–289. Oxford: Oxford University Press.
- Mountcastle VB (1998) *Perceptual Neuroscience: The Cerebral Cortex*. Cambridge, Massachusetts: Harvard University Press.
- Nigrin A (1993) *Neural Networks for Pattern Recognition*. Cambridge, Massachusetts: MIT Press.
- Oja E (1989) Neural networks, principle components, and subspaces. *Int J Neural Syst* 1:61–68.
- O’Reilly RC (1998) Six principles for biologically based computational models of cortical cognition. *Trends Cog Sci* 2(11):455–462.
- Rall W (1964) Theoretical significance of dendritic trees for neuronal input-output relations. In: *Neural Theory and Modeling* (Reiss RF, ed.), pp. 73–97. Stanford, California: Stanford University Press.
- Reynolds JH, Desimone R (1999) The role of neural mechanisms of attention in solving the binding problem. *Neuron* 24(1):19–29.
- Ritter H, Martinez T, Schulten K (1992) *Neural Computation and Self-Organizing Maps. An Introduction*. Reading, Massachusetts: Addison-Wesley.
- Rockland KS (1998) Complex microstructures of sensory cortical connections. *Curr Opin Neurobiol* 8:545–551.
- Roelfsema PR, Lamme VAF, Spekreijse H (2000) The implementation of visual routines. *Vision Res* 40:1385–1411.
- Rumelhart DE, Zipser D (1985) Feature discovery by competitive learning. *Cognitive Sci* 9:75–112.
- Sanger TD (1989) Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks* 2:459–473.
- Segev I (1995) Dendritic processing. In: *The Handbook of Brain Theory and Neural Networks* (Arbib MA, ed.), pp. 282–289. Cambridge, Massachusetts: MIT Press.
- Segev I, Rall W (1998) Excitable dendrites and spines: earlier theoretical insights elucidate recent direct observations. *Trends Neurosci* 21(11):453–460.

- Singer W (1999) Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24(1):49–65.
- Sirosh J, Miikkulainen R (1994) Cooperative self-organization of afferent and lateral connections in cortical maps. *Biol Cybern* 71:66–78.
- Somogyi P, Martin KAC (1985) Cortical circuitry underlying inhibitory processes in cat area 17. In: *Models of the Visual Cortex* (Rose D, Dobson VG, eds.), chpt 54. Chichester: Wiley.
- Spratling MW (1999) *Artificial Ontogenesis: A Connectionist Model of Development*. PhD thesis, Department of Artificial Intelligence, University of Edinburgh.
- Swindale NV (1996) The development of topography in the visual cortex: a review of models. *Network - Comput Neural Syst* 7(2):161–247.
- Tamas G, Buhl EH, Somogyi P (1997) Fast IPSPs elicited via multiple synaptic release sites by different types of GABAergic neurone in the cat visual cortex. *J Physiol* 500(3):715–738.
- Thorpe SJ (1995) Localized versus distributed representations. In: *The Handbook of Brain Theory and Neural Networks* (Arbib MA, ed.), pp. 549–552. Cambridge, Massachusetts: MIT Press.
- von der Malsburg C (1973) Self-organisation of orientation sensitive cells in the striate cortex. *Kybernetik* 14:85–100.
- von der Malsburg C (1981) *The correlation theory of brain function*. Technical Report 81-2, Max-Planck-Institute for Biophysical Chemistry.
- Wallis G (1996) Using spatio-temporal correlations to learn invariant object recognition. *Neural Networks* 9(9):1513–1519.